



Dorota SIKORA

SÉMANTIQUE LEXICALE ET DIDACTIQUE DES LANGUES

Volume 2 : Présentation des publications

en vue de l'obtention de l'Habilitation à diriger des recherches

Table des matières

Présentation générale des travaux publiés

3

Liste complète de publications

5

I. Sémantique de l'aspect et de la temporalité

11

1. Ce qui change avec le temps (grammatical)... Substitution passé composé / imparfait et imparfait / passé composé dans les phrases plurielles en français (avec J. Sypnicki), 2010, *page 13*
2. Portée des préfixes et complexité en français et en polonais. Le cas de RE- versus PO-, PRZY-, DO-, 2012, *page 37*
3. Is the Polish verb *isé* an auxiliary to be?, 2013, *page 55*
4. Quand le parfait est une affaire de prédication seconde : analyse de la construction polonaise N_1 *mieć participe passé* N_2 , 2015, *page 70*
5. Compléments aspectuels et verbes de manière de déplacement en français : entre bornage et télélicité, 2012, *page 94*
6. L'adverbial *en X temps*, est-il vraiment un test de télélicité?, 2013, *page 116*
7. D'achèvement à état : vers une modélisation de quelques structures de polysémie aspectuelle, 2016, *page 127*

II. Sémantique lexicale et modélisation du sens

143

Étude du sens des unités du lexique *146*

8. Pourquoi revient-on à ses moutons? Regard croisé sur la sémantique du retour en français et en polonais, 2006, *page 146*
9. Éléments d'une sémantique pour la description d'un sentiment. Approche contrastive franco-polonaise, 2008, *page 166*

10. *Arriver et venir* – quand la deixis fait (ou ne fait pas) la différence, 2009 *page 185*
11. Le verbe COURIR et ses extensions sportives, 2012, *page 197*
12. NATION et NARÓD : analyse sémantique d'un couple de faux-amis franco-polonais, 2016, *page 209*
13. Peut-on formaliser l'excentricité ? Quand une description sémantique s'organise en réseau. *Cahiers du Littoral*, à paraître, *page 233*
14. Combinatoire lexicale et polysémie : le cas de manquer. *Les Cahiers du Dictionnaire*, Classiques Garnier, à paraître, *page 245*

Sens des unités phraséologiques *page 255*

15. Identification d'unités phraséologiques et équivalence sémantique dans la traduction, 2016, *page 255*
16. Polysémie et homonymie dans un modèle lexical (avec M.-S. Pausé), 2016, *page 269*

Description et modalisation du sens lexical *page 288*

17. Définir le sens dans un réseau lexical, 2012, *page 288*
18. Modélisation de connaissances lexicales dans une ressource, 2017, *page 305*
19. Introduction du numéro thématique *La définition* (avec A. Polguère), 2016, *page 319*

III. Sémantique lexicale et didactique de vocabulaire

323

20. Modèle lexicographique de croissance du vocabulaire fondé sur un processus aléatoire, mais systématique (avec A. Polguère), 2013, *page 326*
21. Brève histoire d'une erreur lexicale. Polysémie et liens lexicaux dans l'enseignement du vocabulaire, 2014, *page 344*
22. Activité(s) de définition dans l'apprentissage d'une langue seconde, 2016, *page 360*
23. Locutions en apprentissage de langue seconde, 2018, *page 384*
24. Entre sens et compétences métalinguistiques : locutions en phraséodidactique du FLE, 2018, *page 408*
25. De quelques aspects de compétence phraséologique en L2 : locutions en test, 2019, *page 415*
26. L'espace dans le roman : bilan et perspectives (avec C. Haman), à paraître, *page 435*
27. Partition de l'espace dans «Les Âmes grises» de Philippe Claudel, à paraître, *page 445*

Présentation générale des travaux publiés

Ce volume réunit mes études réalisées et publiées pendant et après ma recherche doctorale, sans inclure les articles écrits au début des années 90, ni les textes philosophiques traduits au début des années 2000. Ces travaux figurent dans la liste complète de mes publications, mais paraissent aujourd’hui trop anciens pour rendre compte de mes intérêts scientifiques et de mes recherches actuelles en linguistique.

Plutôt que d’enchaîner mes travaux dans un ordre chronologique, j’ai choisi de les rassembler dans des sections dégagées selon le domaine linguistique dans lequel ils s’inscrivent. Ceux de la première partie sont consacrés à la sémantique de l’aspect et de la temporalité. Ces travaux réalisés dans le prolongement de ma recherche doctorale sont suivis des études, réunies dans la Partie 2, dédiées essentiellement à la réflexion sur le sens des unités du lexique, leur description et leur modélisation. Le troisième thème de mes recherches couvre les applications didactiques développées à partir d’un modèle lexical à base sémantique.

Ce découpage s’avère à la fois nécessaire et problématique. Nécessaire pour montrer des convergences thématiques et des préoccupations intellectuelles qui se développent, tout au long de mon parcours, dans une diversité des questions traitées. Problématique dans les choix et dans l’intégration de tel ou tel article dans une section précise, alors qu’il pourrait très bien être inclus dans une autre. C’est le cas de mon article sur le verbe *courir* publié en 2012, qui traite, certes, du sens lexical, mais qui, pour ce faire, interroge les structures argumentales des différentes acceptations de ce vocable, ce qui très clairement a trait à l’analyse aspectuelle.

Enfin, l’assemblage des articles de ce volume ne s’est pas fait sans quelques difficultés techniques liées aux normes et aux habitudes éditoriales qui changent à

la fois selon le temps et le lieu. En effet, d'une part, en dix ans, la présentation même des articles et des ouvrages publiés a sensiblement évolué. De l'autre, les feuilles de style appliquées par les éditeurs continuent à varier même en synchronie. Les fichiers pdf réunis dans ce volume sont de diverses origines. À côté des tirés à part, il y a ceux qui ont été numérisés par mes soins. Parfois, il m'a fallu simplement enregistrer le fichier source en format pdf. Il en résulte un certain éclectisme de la mise en page, qui – comme je l'espère – n'entravera pas la lecture. Pour faciliter la navigation dans les textes, j'ai introduit une numérotation des pages propre à ce volume. Les numéros de pages repris dans la table des matières sont positionnés de manière centrée en bas de chaque page.

Liste complète des publications (30 mars 2020)

Articles et chapitres d'ouvrages

En préparation

- Phrases préfabriquées des annonces ferroviaires : actes (de discours) avoués et inavoués. *Lexique*, numéro thématique sur les phraséologismes pragmatiques co-dirigé avec Gaétane Dostie, publication prévue en 2021.

À paraître

- Combinatoire lexicale et polysémie : le cas de manquer. *Les Cahiers du Dictionnaire*, Classiques Garnier.
- (avec Catherine Haman) L'espace dans le roman : bilan et perspectives. Dans : C. Haman, D. Sikora et G. Kleiber (éds), *L'espace dans le roman français contemporain*, Presses universitaires de Rennes, collection *La Licorne*.
- Partition de l'espace dans «Les Âmes grises» de Philippe Claudel. Dans : C. Haman, D. Sikora et G. Kleiber (éds), *L'espace dans le roman français contemporain*, Presses universitaires de Rennes, collection *La Licorne*.
- Peut-on formaliser l'excentricité ? Quand une description sémantique s'organise en réseau. *Cahiers du Littoral*.

2019

- (avec A. Tsedryk), De quelques aspects de compétence phraséologique en L2 : locutions en test. *Linguistica Atlantica*, 37 (2),
<en ligne : journals.lib.unb.ca/index.php/la/issue/view/2113>

2018

- Entre sens et compétences métalinguistiques : locutions en phraséodidactique du FLE. *Lettre de l'AIRDF*, 64, 60-66.
- Locutions en apprentissage de langue seconde. 6^e Congrès Mondial de Linguistique Française, *SHS Web of Conferences*, 46 (2018), 05002.

2017

- Modélisation de connaissances lexicales dans une ressource, *Revue Française de Linguistique Appliquée*, XXII, 1, 61-74.

2016

- D'achèvement à état : vers une modélisation de quelques structures de polysémie aspectuelle. *Faits de langue*, 48, 147-162.
- (avec A. Polguère) Introduction du numéro thématique *La définition*. Dans : Polguère, A., Sikora, D. (dir) «La définition», *Cahiers de Lexicologie*, 109, 2, 9-11.
- Activité(s) de définition dans l'apprentissage d'une langue seconde. Dans : Polguère, A., Sikora, D. (dir) «La définition», *Cahiers de Lexicologie*, 109, 2, 117-143.
- NATION et NARÓD : analyse sémantique d'un couple de faux-amis franco-polonais. Dans : Degani, M. Frassi, P. et M. I. Lorenzetti (éds) *The Languages of Politics/La politique et ses langages*, vol. 1, Newcastle upon Tyne : Cambridge Scholars Publishing, 209-232.
- (avec M.-S. Pausé) Polysémie et homonymie dans un modèle lexical. 5^e Congrès Mondial de Linguistique Française, *SHS Web of Conferences*, 27 (2016) 05012.

- Identification d'unités phraséologiques et équivalence sémantique dans la traduction. Dans : Corpas Pastor, Gloria (ed.) *Computerised and Corpus-based Approaches to Phraseology : Monolingual and Multilingual Perspectives/Fraseología computacional y basada en corpus : perspectivas monolingües y multilingües*. Geneva : Tradulex. [ISBN 978-2-9700736-8-0], 218-231.

2015

- Quand le parfait est une affaire de prédication seconde : analyse de la construction polonaise N1 mieć participe passé N2. Dans : Giancarli, P.-D., Fryd, M. (éds) «Aoristes et parfaits. En français, latin, corse, estonien et polonais», *Cahiers Chronos*, 28, 119-142.

2014

- Brève histoire d'une erreur lexicale. Polysémie et liens lexicaux dans l'enseignement du vocabulaire. Dans Neveu F., Blumenthal, P. Hriba, L., Gerstenberg, A., Meinschaefer, J. et S. Prévost (éds.), *4^e Congrès Mondial de Linguistique Française – CMLF 2014*, Paris : Institut de Linguistique Française, 1157-1172.

2013

- L'adverbial *en X temps*, est-il vraiment un test de télicité ? Dans : Mirto, M. I. et J. Radimsky (éds) «Adverbs and Adverbial Complements», *Linguisticæ Investigationes*, 36, 2, 276-286.
- Is the Polish verb *ić* an auxiliary to be ? Dans : Kor Chahine, I. (éd.) *Current Studies in Slavic Linguistics*, Amsterdam, Philadelphia : John Benjamins Publishing Company, 123-137.
- (avec A. Polguère) Modèle lexicographique de croissance du vocabulaire fondé sur un processus aléatoire, mais systématique. Dans : Masseron, C., Garcia-Debanc, C. et Ch. Ronveaux (éds) *Enseigner le lexique. Pratiques sociales, objets à enseigner et pratiques d'enseignement*. Namur : Presses universitaires de Namur, 35-61.

2012

- Le verbe COURIR et ses extensions sportives. *Études de linguistique appliquée*, 165, 95-106.
- Portée des préfixes et complexité en français et en polonais. Le cas de RE- versus PO-, PRZY-, DO-. Dans : Paprocka-Piotrowska, U., Martinot, C. et S. Geolimich, S. (éds) *La Complexité en langue et son acquisition*, Lublin : Towarzystwo Naukowe KUL, 131-147.
- Compléments aspectuels et verbes de manière de déplacement en français : entre bornage et télicité. *Études Cognitives/Cognitive Studies*, 12, 57-78.
- Définir le sens dans un réseau lexical. *Studia Romanica Posnaniensia*, 39, 3, 63-79.

2010

- avec (J. Sypnicki) Ce qui change avec le temps (grammatical)... Substitution passé composé/imparfait et imparfait/passé composé dans les phrases plurielles en français. *Neophilologica* 22 (numéro spécial in memoriam du Professeur Stanisław Karolak), 150-173.

2009

- *Arriver et venir* – quand la deixis fait (ou ne fait pas) la différence. *Pratiques*, 141/142, 138-149.

2008

- Éléments d'une sémantique pour la description d'un sentiment. Approche contrastive franco-polonaise. *Arena Romanistica*, 1, 124-142.

2006

- Pourquoi revient-on à ses moutons? Regard croisé sur la sémantique du retour en français et en polonais. *Studia Romanica Posnaniensia*, 33, 199-218.

1997

- Remarks on a Linguistic Readings of Modalities. *Lingua ac communicas*.
- Quelques remarques sur la polysémie des verbes modaux et ses conséquences. Dans : Sypnicki, J. (éd.) *Polysémie, synonymie, antonymie. Relations dans le lexique : aspects théoriques et applicatifs*. Łódź : Wydawnictwo Uniwersytetu Łódzkiego.

1996

- Pourquoi s'intéresser aux modalités logiques quand on fait de la linguistique ? *Studia Romanica Posnaniensia*, 21, 73-89.

1994

- Antonymie interne : cas particulier de polysémie. *Studia Romanica Posnaniensia*, 19, 75-81.

1993

- Adverbe et modalités : pour une esquisse de typologie. Dans : *Tradition et modernité. Actes du colloque Centenaire de la philologie romane à l'Université Jagellonne (24-26.09.1992)*, Kraków : Universitas, 281-286.

Traductions

Livre

2019

- Łukasiewicz, Jan. *Du principe de contradiction chez Aristote. (O zasadzie sprzeczności u Arystotelesa)*, Paris : Édition de l'Éclat. 2^e édition entièrement revue et corrigée.

2000

- Łukasiewicz, Jan. *Du principe de contradiction chez Aristote. (O zasadzie sprzeczności u Arystotelesa)*, Paris : Édition de l'Éclat, 2000[1910].

Articles

2006

- Salamucha, Jan, Le thomisme comme philosophia perennis. *Revue des Sciences Religieuses*, 80, 1, 57-65.

2000

- Drozdowicz, Zbigniew, Les Méditations cartésiennes d'Henri Poincaré. Dans : Bourgeois, B. et J. Hevet (éd.) *L'Esprit cartésien, Actes du XXVI^e Congrès de l'association des Sociétés de Philosophie de langue française*, 1097-1101.

1999

- W.K. Juan Gris dans Almanach de l'art nouveau. *Catalogue de l'exposition Juan Gris, Peintures et dessins 1887-1927*, Musée Cantini, Marseille (17.09.1998-3.01.1999). Marseille : Éditions des Musées nationaux.

1998

- Drozdowicz, Zbigniew, La philosophie polonaise de la première moitié du XX^e siècle. *Zeszyty Filozoficzne*, 6, 150-163.

I. Sémantique de l'aspect et de la temporalité

Sept études publiées entre 2010 et 2016 sont réunies dans cette première partie :

- [1] Ce qui change avec le temps (grammatical)... Substitution passé composé / imparfait et imparfait / passé composé dans les phrases plurielles en français (avec J. Sypnicki). *Neophilologica* 22 (numéro spécial in memoriam du Professeur Stanisław Karolak), 2010, p. 150-173.
- [2] Portée des préfixes et complexité en français et en polonais. Le cas de RE- versus PO-, PRZY-, DO-. Dans : Paprocka-Piotrowska, U., Martinot, C. et S. Geolimich, S. (éds) *La Complexité en langue et son acquisition*, Lublin : Towarzystwo Naukowe KUL, 2012, 131-147.
- [3] Is the Polish verb *ić* an auxiliary to be? Dans : Kor Chahine, I. (éd.) *Current Studies in Slavic Linguistics*, Amsterdam, Philadelphia : John Benjamins Publishing Company, 2013, p. 123-137.
- [4] Quand le parfait est une affaire de prédication seconde : analyse de la construction polonaise N_1 *mieć participe passé* N_2 . Dans : Giancarli, P.-D., Fryd, M. (éds) «Aoristes et parfaits. En français, latin, corse, estonien et polonais», *Cahiers Chronos*, 28, 2015, 119-142.
- [5] Compléments aspectuels et verbes de manière de déplacement en français : entre bornage et télicité. *Études Cognitives/Cognitive Studies*, 12, 2012, 57-78.
- [6] L'adverbial *en X temps*, est-il vraiment un test de télicité? Dans : Mirto, M. I. et J. Radimsky (éds) «Adverbs and Adverbial Complements», *Linguisticae Investigationes*, 36, 2, 2013, 276-286.
- [7] D'achèvement à état : vers une modélisation de quelques structures de polysémie aspectuelle. *Faits de langue*, 48, 2016, 147-162.

Les deux premiers articles sont consacrés aux problèmes des structures situationnelles et à leurs modifications liées directement aux opérateurs temporels [1] et aux

préfixes verbaux [2] en polonais et en français. La perspective comparée adoptée dans ce dernier texte porte essentiellement sur le degré de complexité qu'implique le procédé de préfixation dans sa dimension tant morphologique que sémantique.

Les études [3] et [4] examinent la valeur temporelle et aspectuelle de deux constructions polonaises : *ić + infinitif* dans [3] et N_1 *mieć participe passé* N_2 dans [4], avec l'objectif de vérifier si leur degré de grammaticalisation (si grammaticalisation il y a !) permet de les inclure dans le système des temps verbaux polonais. En effet, du point de vue de leur structure lexico-syntaxique, ces deux constructions polonaises s'inscrivent dans un modèle attesté dans plusieurs langues (par exemple *aller + infinitif* ou le passé composé en français). Néanmoins, pour ce qui est du polonais, dans les deux cas, la réponse apportée dans les articles est négative, même si c'est pour des raisons différentes. Pour le premier, le processus n'a pas suffisamment avancé ; pour le second, l'analyse tant sémantique que syntaxique fait conclure à un cas de prédication seconde.

Enfin, les articles [5] et [6] se focalisent sur le sens verbal dans ses relations avec les compléments qu'il sélectionne ou qu'il autorise, en interrogeant les conséquences aspectuelles de ces combinatoires. La question de la polysémie aspectuelle, qui associe des modifications dans l'organisation du temps interne des situations aux développement de sens multiples d'un vocable, étudiée dans l'article [7], assure une transition vers les travaux de la Partie II, consacrés au sens des unités lexicales et à sa modélisation.

Dorota Sikora

ATILF — CNRS UMR 7118

Józef Sypnicki

Institut Filologii Romainskiej UAM

Ce qui change avec le temps (grammatical)... Substitution passé composé / imparfait et imparfait / passé composé dans les phrases plurielles en français

Abstract

The goal is that our article is not accountable or even less, to take sides in the debate on substitution PC / IMP and IMP / PC. Quite the contrary, we are proposing to clarify some issues related to the substitution of temporal operators in the descriptions of multiple events, we will search in the current work tracks and tools for our investigation.

Throughout this paper, we try to show the elements for the replacement PC / IMP and IMP / PC in sentences referring to many. It is, indeed, they are from different levels of analysis, although strictly interrelated.

Keywords

Aspect, iterativity, tense, event structure, semantics and discourse relations.

Introduction

Si les nombreux travaux de Stanisław Karolak témoignent de la diversité de ses intérêts en linguistique, les problèmes aspectuels y occupent une place toute particulière. L'aspect, envisagé en tant que « catégorie conceptuelle à caractère universel » doit être distingué de ses manifestations formelles, insistait-il encore dans son dernier article (S. Karolak, à paraître). Il définit ce que l'on pourrait également qualifier de *primitifs aspectuels*, en partant du postulat que suivant : à catégorie universelle, des outils capables de la caractériser, quels que soient les moyens linguistiques de l'exprimer. Deux concepts, ceux de continuité temporelle (durée) et de non-continuité temporelle (momentanéité), forment la base du système descriptif qu'il élabore pour expliquer le mécanisme de dérivation des aspects complexes à partir d'aspects simples.

Un solide accord est aujourd'hui établi dans la communauté des linguistes au sujet du caractère compositionnel de l'aspect et de son statut de catégorie sémantique. Tout en souscrivant à ces conceptions générales, nous nous proposons, dans cet article, de porter un regard attentif sur la référence verbale plurielle en français. L'on sait depuis G. Kleiber (1987) qu'un traitement aspectuo-temporel de phrases dites habituelles est de loin préférable à des approches quantificationnelles¹ proposées, entre autres, par Ö. Dahl (1975) ou G. Carlson (1982). Pour notre part, deux raisons ont motivé le choix de cette thématique. La pluralité événementielle et ses manifestations dans la langue impliquent à la fois l'aspect et la sémantique des temps verbaux, c'est-à-dire les deux problématiques que Stanislaw Karolak n'a eu de cesse d'explorer. Deuxièmement, la possibilité de substituer, dans les phrases plurielles, le passé composé à l'imparfait et inversement est un problème épineux pour les apprenants polonophones étudiant le français. Pour modeste qu'elle soit, toute tentative de comprendre les raisons de cette difficulté nous semble présenter un certain intérêt, voire un intérêt certain.

Notre démarche consistera à observer des énoncés attestés pour essayer de déterminer quels sont les éléments qui favorisent ou qui, au contraire, empêchent la substitution entre le passé composé et l'imparfait. Mais avant d'exposer nos analyses, il convient d'explicitier le cadre théorique et les outils conceptuels qui seront les nôtres tout au long de ces pages.

1. Présentation d'outils conceptuels

1.1. Trois types d'événements multiples (pluriels)

Un premier point consiste donc à admettre que les événements décrits par des énoncés d'une langue, en l'occurrence le français, diffèrent par ce que L. Talmy (2000) qualifie de *plexité* (*plexity*). Ce terme renvoie à l'articulation d'éléments équivalents dans la structure d'une entité, c'est-à-dire celle d'un événement pour ce qui est de notre propos. Un événement peut être une simple occurrence ou bien une série d'occurrences itérées, c'est-à-dire un événement multiple².

Nous nous intéresserons aux énoncés qui sont des descriptions d'événements multiples. En suivant les distinctions proposées par G. Kleiber (1987), nous admettons qu'ils peuvent recevoir une lecture habituelle, fréquentative ou itérative.

¹ Pour un argumentaire, voir G. Kleiber (1987 : 103–105).

² L. Talmy (2000 : 48) parle d'événement *simplex*, lorsqu'il s'agit d'une simple occurrence et d'événement *multiple*, lorsque celui-ci en comprend plusieurs. Pour un exposé complet, voir L. Talmy (2000, vol. 1 : 47 et suiv.).

Observons, en effet, les exemples (1) et (2) empruntés à G. Kleiber (1987), renumérotés par nos soins :

- (1) *Dans sa jeunesse, Paul allait à l'école à pied*³.
 (2) *Paul est allé à l'école à pied, le mois dernier*⁴.

Dans les deux cas, il s'agit d'énoncés décrivant une série d'événements du même type (*ipse eo*, d'une série d'occurrences événementielles e_1, e_2, e_3, \dots) réalisés par *Paul*. En d'autres termes, *Paul* a plusieurs fois réalisé un même déplacement à pied pour rejoindre l'établissement scolaire dont il relevait. En résumé, (1) et (2) réfèrent à des événements multiples e composés d'une série d'occurrences e_1, e_2, e_3, \dots itérées pendant un certain temps. La période pendant laquelle elles se reproduisent, c'est-à-dire l'intervalle de référence, « constitue en même temps une spécification de la durée de la situation dénotée » (G. Kleiber, 1987 : 112). Cet intervalle n'est pas nécessairement spécifié dans l'énoncé. Explicite ou implicite, il est situé dans le temps par rapport au moment d'énonciation qu'il peut précéder, recouvrir et/ou suivre totalement ou partiellement⁵.

Comme le montre G. Kleiber (1987), outre un positionnement dans le *continuum* temporel, l'intervalle de référence peut posséder une structuration interne. Dans nos exemples (1) et (2) ci-dessus, on remarque immédiatement que l'intervalle de référence diffère par son organisation : *dans sa jeunesse* a un caractère duratif, dépourvu de bornes temporelles claires, alors que *le mois dernier* fixe une période précise, comprise soit en tant que nombre de jours, soit en tant que laps de temps allant du premier au dernier jour de l'une des douze séquences de l'année.

La structuration de l'intervalle de référence et la répartition des occurrences itérées permettent de distinguer entre trois types de pluralité événementielle décrits par les énoncés français⁶. Les phrases habituelles sont vraies pour un intervalle de référence dont la durée n'est pas forcément contenue entre deux bornes temporelles précises, initiale et finale. Autrement dit, la structuration de l'intervalle de référence est d'une moindre importance. Celui-ci peut être borné à droite et/ou à gauche, sans qu'une telle délimitation soit obligatoire ; c'est la raison pour laquelle tout en différant par la structuration de l'intervalle de référence, (1) et (2) sont bien des phrases habituelles.

Il en est autrement des phrases fréquentatives qui, en plus de marquer la pluralité du référent événementiel, véhiculent un jugement sur la distribution des occurrences qui le constituent. Or, un tel jugement, exprimé le plus souvent par des

³ G. Kleiber (1987 : 112, exemple 9).

⁴ G. Kleiber (1987 : 115, exemple 21).

⁵ Pour une présentation de cette problématique, voir G. Kleiber (1987 : 109—112).

⁶ Il nous est impossible, faute de place, de discuter ici la généricité et la gnomique dans leur rapport avec la pléxité des événements.

adverbiaux tels que *souvent*, *fréquemment*, *rarement*, etc., ne peut pas être dissocié de l'intervalle de référence. En effet, si l'on affirme (3) :

(3) *Paul est souvent allé à l'école à pied le mois dernier.*

on ne peut considérer le déplacement de Paul comme fréquent que par rapport à une période définie, bien délimitée à gauche, comme à droite. Il est possible, en effet, que pour un temps plus long, mettons un trimestre, le même nombre de déplacements soit estimé comme peu élevé.

Dans le cas de phrases itératives, la structuration de l'intervalle de référence, ainsi que la répartition des occurrences itérées importent peu, dans la mesure où leur quantité est simplement chiffrée, comme dans le cas de l'exemple (4) emprunté à G. Kleiber (1987 : 115)⁷ :

(4) *Paul est allé dix fois / plusieurs fois à l'école à pied, le mois dernier.*

En effet, la quantification numérale permet de ne pas prendre en compte l'intervalle de référence, puisque le nombre d'occurrences reste le même que l'on connaisse ou non la période pour laquelle il est valide.

Il en ressort que les énoncés référant à des événements multiples véhiculent les informations au sujet d'au moins deux types d'entités : les occurrences e_1, e_2, e_3, \dots itérées et l'intervalle de référence pendant lequel leur itération se produit. Dans les sections qui suivent nous observerons quelques énoncés à référent pluriel pour vérifier s'il est possible de détecter des corrélations entre les opérateurs temporels employés et le mode de description des éléments qui forment un événement multiple. Avant de le faire, il convient cependant d'exposer les éléments qui nous ont servi de repères dans notre réflexion sur les temps verbaux.

1.2. Du côté de la sémantique des temps verbaux

En abordant les problématiques des temps verbaux français et de leurs valeurs sémantiques, nous sommes pleinement conscients d'avancer sur un terrain meuble, même si celui-ci se limite au passé composé et à l'imparfait. En effet, de nombreuses recherches se poursuivent dans ce domaine, en nourrissant une réflexion et une littérature abondantes constamment renouvelées. De même, pour ce qui est de l'aspect, force est de constater que d'innombrables ouvrages consacrés depuis plus d'un siècle à la perfectivité et à l'imperfectivité verbales sont quelque peu intimidants. Comment ne pas paraître réducteurs, en touchant à une problématique dont des générations de linguistes ont montré la complexité ?

⁷ G. Kleiber (1987 : 115, exemple 20).

Cependant, notre objectif n'est pas d'en rendre compte ou, encore moins, de prendre position dans ces débats passionnants. Bien au contraire, en nous proposant d'élucider quelques questions liées à la substitution des opérateurs temporels dans les descriptions d'événements multiples, nous rechercherons dans les travaux actuels des pistes et des outils pour notre investigation. Dans la présente section, nous fixons les points de repère qui ont organisé notre recherche.

1.2.1. Types de situations et points de vue aspectuels

Pour décrire un événement particulier réel ou fictif, le locuteur dispose dans sa langue de représentations idéalisées associées à des prédicats. Et C. Tenny (1994 : 3) de préciser : «[...] les prédicats des langues naturelles imposent une certaine structure sur les événements qu'ils décrivent, structure qui est en grande partie temporelle ou aspectuelle »³.

Ces objets linguistiques sont des abstractions construites de manière intensionnelle, à partir de perceptions individuelles, mais partagées dans une communauté langagière. Depuis E. Bach (1986), on les désigne sous le nom d'*éventualités* (*eventualities*) ou de *situations*, que l'on répartit en un certain nombre de catégories. Ainsi, en français *peindre* est-il, selon le classement que l'on adopte, un achèvement (dans la tradition vendliérienne), une transition (C. Vet, 1980 ; P. Caudal, 2000) ou culmination (E. Bach, 1986), alors que *marcher* doit-il être classé parmi les activités (Z. Vendler, 1967 ; C. Vet, 1980) ou processus (P. Caudal, 2000).

«Lorsque les locuteurs parlent d'une situation particulière, ils la présentent comme un exemplaire d'un type idéalisé, en recourant aux formes linguistiques associées avec lui», explique C. Smith (1986 : 99). Elle opte ainsi pour une conception énonciative de l'aspect (*speaker-based approach*), partagée en linguistique française entre autres par A. Borillo (1991), L. Gosselin (1996), P. Caudal (2000, 2006), P. Caudal et C. Vettiers (2007). Dans cette perspective, le locuteur choisit d'abord le type de situation à laquelle il associe, dans son discours, l'événement du monde (réel ou fictif) dont il construit la description. Les énoncés *Il n'est plus là* et *Il est parti il y a une heure* peuvent être parfaitement co-référentiels, en renvoyant à un même état de choses. Cependant, dans le premier cas, le locuteur l'associe à un état, c'est-à-dire une situation durative et stative, alors que dans le second il le désigne par recours à un événement ponctuel et dynamique.

C'est également dans la lignée de C. Smith (1991) que nous abordons les notions de perfectivité, d'imperfectivité et de résultativité. Une fois que le locuteur associe le référent extra-linguistique à un type de situation disponible dans la langue, il le présente selon un point de vue. À ce propos, L. Gosselin (1996) parle

³ C. Tenny (1994 : 3) : "[...] predicates of natural language impose a certain structure on the events they describe, a structure which is to a large extent temporal or aspectual".

de *monstration*, alors que P. Caudal (2000) ainsi que P. Caudal et C. Vetters (2007) recourent à la notion de *focalisation*. Il s'agit cependant d'une opération similaire, qui « donne à voir » la situation d'une certaine façon. En français, une situation peut être saisie selon un point de vue imperfectif, perfectif ou résultatif⁹. Dans cette langue, l'imperfectivité, la perfectivité et la résultativité sont des propriétés d'énoncés, liées avant tout à l'emploi d'un temps grammatical. Tout en restant conscients que ces notions font aujourd'hui objet de nombreux débats, il nous est indispensable, pour mener nos analyses, de nous en remettre à un état de la question. On admettra donc que l'imperfectivité se caractérise par sa valeur sécante : elle donne à voir un procès en déroulement, seulement « une partie de l'éventualité » (P. Caudal, 2006), sans sa borne finale. Le point de vue aspectuel imperfectif est véhiculé entre autres par l'imparfait, alors que le passé simple, le plus-que-parfait et le passé composé permettent de montrer cette même situation d'un point de vue perfectif, c'est-à-dire dans sa totalité et dans ses limites temporelles. Les temps grammaticaux français souvent qualifiés de *parfaits* tels que le passé composé et le plus-que-parfait sont en mesure de mettre en focus l'état résultant respectivement d'un événement antérieur au moment d'énonciation ou à un autre point de référence situé dans le passé.

On comprend dès lors pourquoi il est possible, du moins dans une langue comme le français, de traiter les temps verbaux non seulement de tiroirs temporels, mais également d'opérateurs de points de vue aspectuels. P. Caudal et C. Vetters (2007 : 137) l'expliquent ainsi : « un temps verbal a pour fonction illocutoire de marquer la perspective adoptée par l'énonciateur vis-à-vis d'une situation décrite par un énoncé ». Autant dire qu'il permet d'effectuer une opération de mise en focus.

1.2.2. Pragmatique des temps verbaux et relations discursives

Dans le passage cité, P. Caudal et C. Vetters mettent en avant un point important. En évoquant la fonction illocutoire des opérateurs temporels, ils soulignent la contribution des facteurs pragmatiques dans le calcul des valeurs aspectuelles d'un énoncé. Nous verrons que l'analyse des phrases plurielles en français doit bien prendre en compte cet apport pragmatique, car il est indispensable pour identifier

⁹ Nous avons essayé de montrer ailleurs (D. Sikora, 2009) qu'une approche en termes de points de vue aspectuels est particulièrement intéressante pour des analyses comparatives, notamment entre les langues romanes et les langues slaves. Pour ce qui est de ces dernières, elle autorise à considérer que le point de vue perfectif ou imperfectif est lexicalisé, donc indissociable de la forme verbale. Ce qui nous conduit à admettre deux points : i. l'aspect imperfectif dans une langue comme le polonais est lié à la forme verbale elle-même, alors qu'en français il est l'effet d'un calcul sémantique effectué sur le plan de l'énoncé ; ii. le français possède un point de vue résultatif qui, en polonais, n'apparaît que sur le plan pragmatique (D. Apothéloz, M. Nowakowska, à paraître).

les relations discursives qui s'avèrent plus d'une fois décisives quant à la substitution des opérateurs temporels *PC/IMP* et *IMP/PC*, qui nous intéresse tout particulièrement dans cette étude. Ajoutons qu'en parlant de relations discursives, nous nous référons à celles qui ont été définies et modélisées par N. Asher (1993), A. Lascarides et N. Asher (1993), N. Asher et A. Lascarides (2003) dans le cadre de la *Discourse Representation Theory*, mais que nous en proposons une approche peu formelle. En effet, N. Asher et A. Lascarides proposent de distinguer cinq principales relations discursives, à savoir Explication (*Explanation*), Elaboration (*Elaboration*), Narration (*Narration*), Arrière-Plan (*Background*) et Résultat (*Result*). Leur présentation dépasse le cadre d'un article, aussi nous limiterons-nous à définir succinctement celles que nous sommes amenés à traiter à une étape de notre réflexion.

En nous appuyant sur les éléments discutés dans cette section, nous croyons possible de repérer, dans celles qui suivent, les facteurs essentiels responsables du jeu des opérateurs temporels *PC/IMP* et *IMP/PC* dans les phrases à référent événementiel multiple en français.

Tout au long de ce travail, nous avons constamment pris deux précautions. Premièrement, dans le souci de parvenir à des résultats tant soit peu cohérents, nous nous sommes concentrés sur deux opérateurs temporels du riche paradigme verbal du français : le *PC* en tant que porteur du point de vue aspectuel perfectif et résultatif, ainsi que l'*IMP* qui parmi ses valeurs possibles, véhicule celle d'imperfectif. La seconde précaution que nous avons eue est inhérente aux choix d'exemples. En effet, nous avons construit notre raisonnement uniquement sur des énoncés authentiques attestés. Il va sans dire que, contrairement aux exemples construits, de nombreux facteurs discursifs y interfèrent avec les informations sémantiques, pour construire la signification visée par le locuteur dans un contexte d'énonciation spécifique. Prétendre dresser un répertoire complet de facteurs favorisant ou, au contraire, bloquant l'interchangeabilité entre le *PC* et l'*IMP* reviendrait à faire preuve d'une ambition démesurée. Sans présumer de nos forces, il nous semble néanmoins possible d'identifier certains d'entre eux.

2. Qu'en est-il de la substitution *PC/IMP* et *IMP/PC* ?

Lorsque, dans un énoncé pluriel, le *PC* vient remplacer l'*IMP* ou inversement, plusieurs résultats doivent être envisagés. Il n'est pas rare, en effet, que le changement d'opérateur, *ipse eo* de point de vue aspectuel, produise un énoncé mal formé, rejeté par nos informateurs natifs. Ainsi, l'énoncé (5a) a-t-il été unanimement qualifié d'agrammatical :

- (5) *J'ai nagé deux fois par semaine de septembre à décembre, 3000 mètres par séance environ. (www.velo101.com)*
- (5a) * *Je nageais deux fois par semaine de septembre à décembre, 3000 mètres par séance environ.*

De même, (6a) a été considéré comme déviant :

- (6) *Les îles du Nicaragua, Utilita, génial pour la plongée, j'ai nagé deux fois avec des requins baleines. (www.lonelyplanet.fr)*
- (6a) ? *Les îles du Nicaragua, Utilita, génial pour la plongée, je nageais deux fois avec des requins baleines.*

Lorsque la substitution PC / IMP est autorisée, il convient de distinguer plusieurs types de situations. Il se peut que l'énoncé original et sa version remaniée restent co-référentiels, avec toutefois des différences de signification. Les exemples (7) et (7a) illustrent une telle situation :

- (7) *C'est difficile de trouver des restos pas trop chers dans le Kansai, enfin surtout dans les grandes villes, on a eu moins de mal à Hiroshima. On a souvent marché pendant des heures avant de trouver le plat pas cher qui n'avait pas l'air d'avoir trempé dans l'eau pendant des mois. (http://cacahuèteaujapon.hautetfort.com)*
- (7a) *C'est difficile de trouver des restos pas trop chers dans le Kansai, enfin surtout dans les grandes villes, on a eu moins de mal à Hiroshima. On marchait souvent pendant des heures avant de trouver le plat pas cher qui n'avait pas l'air d'avoir trempé dans l'eau pendant des mois.*

Ailleurs, la substitution conduit à un énoncé parfaitement acceptable dans sa forme grammaticale, mais référant à un tout autre événement. C'est ce que l'on observe en (8) et en (8a) : le premier décrit un événement multiple de type habituel, alors que le second renvoie à une occurrence.

- (8) *Avant le lever du soleil, on avait chargé nos affaires et attelé Madeleine : il fallait profiter au maximum des courtes journées d'automne. Nous marchions jusqu'à midi, progressant lentement, trois kilomètres à l'heure à peine — avant de nous arrêter dans un nouveau champ que Nathalie et Pierre avaient repéré la veille ou le matin même. (Frantext, Mathews, H. Ma vie dans la CIA : une chronique de l'année 1973).*
- (8a) *Avant le lever du soleil, on avait chargé nos affaires et attelé Madeleine : il fallait profiter au maximum des courtes journées d'automne. Nous avons marché jusqu'à midi, progressant lentement, trois kilomètres à l'heure à peine — avant de nous arrêter dans un nouveau champ que Nathalie et Pierre avaient repéré la veille ou le matin même.*

Dans les sections qui suivent nous allons porter un regard plus attentif sur ces énoncés et sur les effets produits tant sur le plan sémantico-aspectuel, que sur celui de la référence. Chemin faisant, on aura l'occasion d'observer comment ces changements affectent les relations discursives entre les événements décrits et, par là-même, la représentation des réalités rapportées.

3. Blocage de substitution PC / IMP

Observons les exemples (5) et (6) dans lesquelles le PC ne peut pas être remplacé par l' IMP sans provoquer d'agrammaticalité. Il convient de noter cependant que si les énoncés (5a) et (6a) ont en commun leur forme boiteuse, l'anomalie tient, dans les deux cas, à des raisons différentes. Nous discutons ces deux cas dans les sections (3.1) et (3.2).

3.1. Opérateurs PC-IMP et structuration de l'intervalle de référence

Le regard rapide porté sur les exemples (1) et (2) au début de cette étude autorise à soupçonner quelque relation entre le choix de l'opérateur temporel et la présentation de l'intervalle de référence. En effet, dans (1), l'adverbial *dans sa jeunesse* le saisit de manière durative, sans limite temporelle claire, en coïncidant avec l'opérateur IMP . Dans (2), l'emploi du PC s'accompagne de l'adverbial *le mois dernier*, auquel il convient d'accorder une lecture inclusive.

Avant de promouvoir le soupçon au rang d'hypothèse, analysons l'exemple (5) et sa version modifiée (5a) sous l'angle d'un éventuel rapport entre l'opérateur temporel et la structuration de l'intervalle de référence.

(5) *J'ai nagé deux fois par semaine de septembre à décembre, 3000 mètres par séance environ.* (www.velo101.com)

L'énoncé (5) doit être classé parmi les phrases habituelles. Les occurrences e_p , e_p , e_p ... sont vérifiées pour un intervalle de référence clairement délimité par le complément *de septembre à décembre* et leur répartition s'organise selon le mode spécifié par l'adverbial *deux fois par semaine*. Une telle organisation interne de l'événement multiple bloque la substitution du PC par l' IMP , comme le montre (5a), qualifié de déviant par tous nos informateurs francophones :

(5a) * *Je nageais deux fois par semaine de septembre à décembre, 3000 mètres par séance environ.*

La question que l'on se pose face à (5a) est celle de l'élément responsable du blocage. Avec l'exemple (1), nous avons pu constater que l'*IMP* n'est pas par principe incompatible avec la pluralité habituelle. Il faut noter cependant que dans (1), l'intervalle de référence est présenté dans sa durée, sans borne temporelle précise. De son côté, (5) comporte deux adverbiaux délimitatifs. Le premier, *de septembre à décembre* est de nature temporelle et il spécifie les bornes initiale et finale de l'intervalle de référence. Le second, *3000 mètres par séance environ*, a un caractère spatial et il permet de délimiter chacune des occurrences e_1, e_2, e_3, \dots qui constituent l'habitude assertée de x (*je*). Dans (5b) et (5c), nous avons introduit l'*IMP* à la place du *PC*, en supprimant à tour de rôle l'un des modificateurs délimitatifs :

(5b) *Je nageais deux fois par semaine, 3000 mètres par séance environ.*

(5c) * *Je nageais deux fois par semaine de septembre à décembre.*

Dans (5b), le modifieur délimitatif *3000 mètres par séance environ*, spécifiant l'étendue spatiale de chacune des occurrences e_1, e_2, e_3, \dots , a été préservé. L'énoncé (5b) est pleinement acceptable, tout en référant à un événement multiple à caractère habituel. Il s'avère donc qu'une configuration, conjuguant la délimitation de chacune des occurrences et l'opérateur temporel *IMP* demeure possible, sans altérer la lecture habituelle.

Il en est autrement de (5c). En présence du modifieur délimitatif *de septembre à décembre*, qui marque les bornes temporelles de l'intervalle de référence pendant lequel l'habitude attribuée à x (*je*) reste valide, l'*IMP* ne peut pas remplacer le *PC* sans provoquer d'agrammaticalité.

Les exemples (5b) et (5c) mettent en avant un premier facteur qui exclut la substitution du *PC* par l'*IMP*. Il s'agit très clairement de l'intervalle de référence. Nous avons admis, avec G. Kleiber (1987), que sa structuration compte peu dans le cas des phrases habituelles qui restent valides pour une période dont les contours temporels peuvent être explicitement marqués ou, au contraire, rester flous. Néanmoins, lorsque l'intervalle de référence est nettement délimité, comme celui de l'exemple (5), le *PC* (ou bien un autre opérateur temporel véhiculant le point de vue perfectif) est de rigueur, car les bornes qu'il impose correspondent à celles fixées par le modifieur délimitatif. *A contrario*, l'*IMP*, qui présente l'événement sans référer à ses bornes, conduit à un énoncé déviant, puisqu'un hiatus s'installe entre les informations véhiculées par cet opérateur temporel et le modifieur délimitatif : l'un présente l'intervalle comme dépourvu de limite, pendant que l'autre indique la présence d'un début et d'une fin.

3.2. Opérateur *rc* et pluralité itérative

Il est intéressant de noter que nous ne disposons, dans notre corpus de travail, d'aucun exemple de phrase itérative dont le verbe soit employé à l'imparfait. Seuls les temps à point de vue perfectif sont attestés, comme dans l'exemple (6) :

- (6) *Les îles du Nicaragua. Utilla, génial pour la plongée. J'ai nagé deux fois avec des requins baleines. (www.lonelyplanet.fr)*

Face à l'absence de cas attestés, nous avons essayé de remplacer le *rc* de l'exemple (6) par un *msr*. Le résultat, présenté sous (6a), apparaît comme clairement défectueux :

- (6a) **Les îles du Nicaragua. Utilla, génial pour la plongée. Je nageais deux fois avec des requins baleines.*

Si l'*msr* s'accorde mal avec la pluralité itérative, c'est probablement dû à la nature même de cette dernière. Rappelons que, conformément à G. Kleiber (1987 : 115) : « Une phrase simplement itérative est une phrase qui présente une situation comme étant vérifiée à deux, trois, ..., plusieurs reprises à l'intérieur d'un intervalle temporel ». Cependant, la quantification numérale, qui consiste à fournir le nombre exact d'occurrences itérées, réduit sensiblement le rôle que l'intervalle joue dans la représentation de l'événement multiple. Imaginons, en effet, que le locuteur de (6) n'ait plus jamais l'occasion, sa vie durant, de renouveler son expérience nautique. Entouré de ses petits enfants, il pourra toujours énoncer (6b) sans mentir :

- (6b) *J'ai nagé deux fois avec des requins baleines.*

Deux fois désigne une valeur absolue, quelle que soit l'étendue de la période pendant laquelle les occurrences sont reproduites : un séjour de vacances ou une vie entière.

Si, comme nous le pensons à la suite de G. Kleiber (1987), l'intervalle de référence a une saillance très faible en cas de quantification numérale, il convient d'admettre que le *rc* (ou tout autre opérateur temporel) compte alors pour ce qui est du mode de présentation des occurrences e_1, e_2, e_3, \dots . Considérons de ce point de vue l'exemple (6). L'événement décrit par le prédicat *nager_{rc}* comprend deux occurrences e_1 et e_2 , vérifiées pour la période de référence implicite qui correspond à la durée du séjour sur l'île d'Utilla. En simplifiant, on peut constater que la pluralité itérative apparaît, lorsqu'il s'agit de compter les occurrences e_1, e_2, e_3, \dots qui forment l'événement multiple e . Or, on ne peut compter que ce qui est comptable. Le terme de *comptable* que l'on utilise avant tout pour caractériser la référence

nominale, peut être transféré dans le domaine verbal, si l'on accepte une approche davidsonienne¹⁰ autorisant à réifier les événements.

Nous pensons, en effet, que contrairement à (5), le point de vue perfectif véhiculé par le *PC* dans (6) porte non pas sur l'intervalle de référence, mais sur les occurrences itérées. Deux arguments viennent étayer cette hypothèse. Premièrement, la quantification numérale, condition *sine qua non* de l'itérativité telle qu'elle est définie par G. Kleiber (1987), ne peut être effectuée que si les entités sur lesquelles elle porte sont saisies comme comptables. Pour des événements, cela exige une conceptualisation avec des limites temporelles précises. Deuxièmement, la quantification numérale introduit un ordre chronologique nettement plus marqué que celui d'une phrase simplement habituelle et fréquentative où la succession n'est que présupposée. Or, pour pouvoir entrer dans une séquence chronologique, un événement doit être présenté dans son intégralité¹¹. On comprend dès lors que l'opérateur temporel *IMP*, en tant que porteur d'imperfectivité, ne peut qu'être rejeté dans une phrase itérative¹², dans la mesure où il donne à voir l'événement en cours de son déroulement, sans fixer ses limites temporelles des occurrences e_1, e_2, \dots, e_n . En résumé, l'incompatibilité entre l'opérateur *IMP* et l'expression de pluralité itérative s'explique, à notre avis, par l'articulation de deux facteurs :

- (i) l'intervalle de référence n'ayant que peu d'importance en cas de quantification numérale, l'opérateur temporel d'une phrase itérative détermine le mode de présentation des occurrences e_1, e_2, \dots, e_n qui forment e et qui
- (ii) doivent être saisies en tant qu'entités comptables, c'est-à-dire avec des limites temporelles nettes. Dans ce cas, seuls les temps à point de vue perfectif, dont le *PC*, sont autorisés.

L'observation de nos exemples permet de formuler deux constats. Premièrement, nous avons pu remarquer que les opérateurs temporels portent, dans un énoncé pluriel, sur deux types d'entités : l'intervalle de référence ou les occurrences itérées. Les facteurs responsables du blocage imposé à la substitution des opérateurs temporels *PC* / *IMP* et *IMP* / *PC* n'en sont pas indépendants.

Lorsque, dans une phrase habituelle, l'intervalle de référence est borné, seul le *PC* semble de mise. C'est ce que nous avons pu observer dans (5), où les bornes imposées par l'opérateur temporel *PC* doivent correspondre à celles fixées par le modifieur délimitatif *de septembre à décembre*. L'introduction de l'*IMP* aboutit à une contradiction entre l'information [-BORNE] qu'il véhicule et celle [+BORNE] apportée par ce modifieur délimitatif qui explique, à notre avis, le caractère inacceptable de (5a).

Deuxièmement, avec un exemple comme (6), il est possible de repérer un lien entre le type de pluralité événementielle et la saisie des occurrences, notamment

¹⁰ Cf. D. Davidson (1967) et sa traduction française (1993) par Pascal Engel.

¹¹ Cf. M. Kozłowska (1988b).

¹² Sauf pour ce qui est de l'imparfait dit *narratif* qui ne semble pas exclu *a priori*. Nous n'en avons cependant trouvé aucun exemple attesté.

dans les énoncés itératifs. On verra, dans la section suivante, que les phrases habituelles et fréquentatives sont nettement plus disponibles pour la substitution des opérateurs rc/IMF et IMF/rc du fait, entre autres, de contraintes moins lourdes sur la saisie des occurrences e_p, e_p, e_p, \dots .

4. Substitution d'opérateurs temporels et ses effets sémantico-aspectuels

4.1. Opérateurs rc/IMF et le mode de présentation d'occurrences

Les exemples (1) et (2) que nous avons empruntés à G. Kleiber (1987) et discutés au début de nos investigations, laissent penser que les deux opérateurs, le rc et l' IMF , peuvent apparaître dans les phrases habituelles. Il serait cependant illégitime de conclure que la substitution rc/IMF ou IMF/rc s'y effectue sans aucune contrainte. Le cas de (5) présenté dans la section 3.1 indique clairement que celle-ci ne peut pas se faire systématiquement et que le caractère inclusif de l'intervalle de référence peut imposer le recours à un opérateur temporel à point de vue perfectif tel que le rc .

Afin de dégager quelque régularité, nous nous sommes tournés vers notre corpus de travail. On observe, en effet, que le va-et-vient entre le rc et l' IMF se produit avant tout dans les phrases habituelles (9 et 9a) et dans les fréquentatives (7 et 7a).

- (9) ... ce sont les régions de Paris où j'ai **marché des nuits** entières quand j'étais une petite touriste étudiante, apprenant par cœur les noms des rues et des places jusqu'à en rêver. (Frantext, Bastide, F.-R., *Les adieux*)
- (9a) ... ce sont les régions de Paris où je **marchais des nuits** entières quand j'étais une petite touriste étudiante, apprenant par cœur les noms des rues et des places jusqu'à en rêver.
- (7) (a) C'est difficile de trouver des restos pas trop chers dans le Kansai, enfin surtout dans les grandes villes, on a eu moins de mal à Hiroshima. (b) **On a souvent marché** pendant des heures avant de trouver le plat pas cher qui n'avait pas l'air d'avoir trempé dans l'eau pendant des mois. (<http://cacahué-teaujapon.hautetfort.com>)
- (7a) C'est difficile de trouver des restos pas trop chers dans le Kansai, enfin surtout dans les grandes villes, on a eu moins de mal à Hiroshima. **On marchait souvent** pendant des heures avant de trouver le plat pas cher qui n'avait pas l'air d'avoir trempé dans l'eau pendant des mois.

(9), ainsi que sa version remaniée (9a) sont des phrases habituelles. De même, dans le cas de (7) et de sa forme modifiée (7a), le changement d'opérateur temporel n'affecte pas le caractère fréquentatif de la phrase. S'agit-il pour autant, dans les deux cas, d'énoncés synonymes ? Pour répondre à cette question, il faut d'abord en poser une autre : celle de savoir ce qui change avec l'opérateur temporel dans les deux paires d'exemples ?

Dans (9) comme dans (9a), les occurrences e_1, e_2, e_3, \dots qui forment l'habitude se reproduisent tout au long de la période spécifiée dans la subordonnée temporelle *quand j'étais une petite touriste étudiante*, présentée dans sa durée, sans aucune borne temporelle. Le mode de présentation de l'intervalle de référence semble donc commun à (9) et à (9a). On peut donc légitimement admettre que, dans ces énoncés, l'opérateur temporel n'interagit pas avec ce paramètre. Si cette observation est exacte, l'opérateur temporel doit avoir trait à la présentation des occurrences.

Effectivement, le πc de l'énoncé (9) présente chacune des occurrences itérées de marche à travers Paris comme une séquence délimitée, équivalent d'une nuit. Les bornes initiale et finale de e_1, e_2, e_3, \dots correspondent ainsi au début et à la fin de chaque période comprise entre le coucher et le lever du jour. On pourrait ainsi parler de saisie perfective de chacune des occurrences itérées.

Cependant, l'adverbial *des nuits entières* autorise également une autre lecture, actualisée en (9a) sous l'effet de l'opérateur *imr*. Les occurrences e_1, e_2, e_3, \dots sont saisies dans leur étendue temporelle. Autrement dit, le locuteur donne à voir leur phase interne en déroulement. Du coup, l'adverbial *des nuits entières* désigne ces périodes de manière durative, et non plus inclusive, comme cela avait lieu dans (9).

Si notre interprétation est valide, elle permet de constater que le point de vue perfectif ou imperfectif lié à l'emploi d'un opérateur temporel ne porte pas exclusivement sur l'intervalle de référence. Il existe des cas, comme ceux de (9) et (9a), où il est à l'origine du mode de présentation des occurrences e_1, e_2, e_3, \dots dont l'itération forme l'événement multiple.

Observons à présent la configuration des éléments qui forment la description de l'événement multiple dans (7). Conformément aux critères établis par G. Kleiber (1987), il s'agit d'une phrase fréquentative : l'itération des occurrences *marcher*_(march) est vérifiée pour une période, certes, présentée dans sa durée, mais dont les bornes temporelles sont clairement fixées. Cet intervalle de référence n'est pas explicitement déterminé, mais le calcul des relations discursives permet de l'identifier comme celui qui correspond à la durée du séjour dans la région du Kansai¹⁷.

¹⁷ Il s'agit de la relation d'Élaboration que N. Asher et A. Lascarides (2003 : 440) définissent ainsi : "β's event is part of α's event". Effectivement, l'événement multiple que nous avons marqué de β (*On a souvent marché / On marchait souvent pendant des heures...*) fait partie de l'événement α identifié à travers ... *trouver des restos pas trop chers dans le Kansai* qui en pré suppose un autre, à savoir être dans le Kansai.

Rappelons que dans les phrases fréquentatives, l'intervalle de référence possède toujours des contours temporels nets, qui ne peuvent pas être modifiés sous peine d'invalider le jugement fréquentatif. En effet, on peut supposer qu'en prolongeant le séjour dans le Kansai, le locuteur/la locutrice aurait mieux connu les lieux, ce qui écarterait les déplacements à la recherche d'un restaurant adapté. Par conséquent, les touristes ne pourraient plus affirmer avoir souvent effectué de longues marches : rapportées à un intervalle plus long, celles-ci ne seraient plus considérées comme fréquentes.

En comparant (7) avec sa version (7a) dans laquelle *imP* remplace *rc*, on constate que le changement d'opérateur temporel n'a pas atteint la structuration de l'intervalle de référence. De même, la relation discursive d'Élaboration entre les énoncés α et β reste intacte. Pas plus que dans (9) et (9a), les opérateurs temporels ne semblent donc pas porter sur l'intervalle de référence. Ce qui change, c'est le mode de présentation des occurrences itérées : dans (7), le *rc* marque le point de vue perfectif que le locuteur porte sur elles, alors que l'*imP* de (7a) adopte une perspective imperfective.

Nous tenons à présent les éléments qui permettent de répondre aux deux questions posées au début de cette section. La première consistait à nous interroger sur la nature du changement produit par la substitution. À ce sujet, nous avons pu constater que les opérateurs temporels *rc* et *imP* introduisent un point de vue différent sur les occurrences e, e_p, e_p, \dots qui forment l'événement multiple e .

La seconde question est celle de l'éventuelle synonymie entre les énoncés originaux au *rc* (9 et 7) et leurs versions modifiées par substitution de l'opérateur *imP* (9a et 7a). Or, avec les deux énoncés de chaque paire, le locuteur décrit incontestablement un même événement multiple e , respectivement habituel et fréquentatif. Par conséquent, (9) et (9a), tout comme (7) et (7a) restent co-référentiels. En revanche, la modification de l'opérateur temporel change le point de vue aspectuel porté sur les occurrences e_p, e_p, e_p, \dots . Elle s'accompagne donc clairement de différence de signification entre les énoncés de chaque couple. Par conséquent, la relation synonymique n'est que partielle.

Remarquons que, contrairement à la pluralité itérative, l'habitualité et la fréquentivité n'exigent pas une saisie globale des occurrences itérées. Celles-ci peuvent, au contraire, être présentées dans la durée, sans limite temporelle finale. La possibilité d'une double conceptualisation de e_p, e_p, e_p, \dots autorise la substitution des opérateurs *rc* / *imP*.

4.2. Substitution *imP* / *rc* et modifications référentielles

Dans plusieurs exemples, le changement d'opérateur temporel produit des modifications sur le plan de la référence. Rappelons à ce titre l'exemple (8) pour le comparer avec sa version modifiée (8a) :

- (8) *Avant le lever du soleil, on avait chargé nos affaires et attelé Madeleine : il fallait profiter au maximum des courtes journées d'automne. Nous marchions jusqu'à midi, progressant lentement, trois kilomètres à l'heure à peine — avant de nous arrêter dans un nouveau champ que Nathalie et Pierre avaient repéré la veille ou le matin même.* (Frantext, Mathews, H. Ma vie dans la CIA : une chronique de l'année 1973).
- (8a) *Avant le lever du soleil, on avait chargé nos affaires et attelé Madeleine : il fallait profiter au maximum des courtes journées d'automne. Nous avons marché jusqu'à midi, progressant lentement, trois kilomètres à l'heure à peine — avant de nous arrêter dans un nouveau champ que Nathalie et Pierre avaient repéré la veille ou le matin même.*

Il convient de remarquer que la lecture habituelle de *nous marchions jusqu'à midi* semble tenir essentiellement à l'opérateur *MP*. Certes, le cotexte gauche évoque *de courtes journées d'automne* dont la forme plurielle renforce l'interprétation selon laquelle les séquences de marche ont eu lieu plus d'une fois, mais il n'empêche pas la référence spécifique qui apparaît en (8a) avec le *PC*.

Pour saisir le mécanisme à l'œuvre, observons d'abord le rôle assumé dans les deux versions de notre exemple, par le modifieur temporel *jusqu'à midi*. Dans (8), ce modifieur permet de fixer la borne finale de chacune des occurrences itérées, alors qu'en (8a) il marque la limite d'un événement spécifique. Or, force est de constater que dans ce second cas, le *PC* fait correspondre la borne finale fixée par le modifieur et celle due à l'opérateur temporel. Dans une telle configuration, la lecture par défaut est donc spécifique ; la construction de la référence multiple n'y est pas exclue *a priori*, mais elle nécessite le recours à une expression quantifiante telle que *généralement, d'habitude, souvent*, etc. marquant l'itération des séquences de marche, comme dans (8b).

- (8b) *Avant le lever du soleil, on avait chargé nos affaires et attelé Madeleine : il fallait profiter au maximum des courtes journées d'automne. Généralement, d'habitude, souvent nous avons marché jusqu'à midi, progressant lentement, trois kilomètres à l'heure à peine — avant de nous arrêter dans un nouveau champ que Nathalie et Pierre avaient repéré la veille ou le matin même.*

Le rôle assumé par l'opérateur *MP* apparaît désormais plus clairement. Les éléments dégagés lors de la comparaison avec (8a) portent à croire qu'appliqué à un prédicat qui renvoie une occurrence bornée, l'*MP* indique qu'il s'agit d'événement multiple. Or, chacune des occurrences itérées étant délimitée, le point de vue imperfectif qu'il véhicule porte inévitablement sur l'intervalle de référence, présenté sans borne temporelle. Il est intéressant de noter un effet évidemment corollaire de la continuité associée à l'*MP* : appliqué à des occurrences d'événement délimitées et, de ce fait, conceptualisées comme entités comptables, l'opérateur temporel *MP*

marque leur itération. Il est important par ailleurs de souligner que, dans notre exemple (8), l'intervalle ne possède aucune expression explicite sous forme de circonstant. Seul l'imparfait l'indique dans sa continuité.

On pourrait difficilement ne pas rapprocher¹⁴ le cas de (8) et de (8a) d'un autre phénomène bien connu en aspectologie, à savoir l'effet multiplicatif exercé par l'IMP, lorsqu'il est appliqué aux achèvements et aux accomplissements. Les exemples (10) et (10a) en sont une illustration :

- (10) *L'essentiel : j'ai perdu mon sac [...] Signe pur de malaise, inquiétude, non signe de perdre ma féminité...* (Franz, Ernaux A., *Se perdre*)
 (10a) *L'essentiel : je perdais mon sac [...] Signe pur de malaise, inquiétude, non signe de perdre ma féminité...*

Employé au PC, *perdre mon sac* s'interprète par défaut comme une occurrence d'événement spécifique, sauf si une expression quantifiante telle que *quelquefois*, *souvent*, *plusieurs fois*, etc. indique la plexité multiple. Dans le cas d'un achèvement, tel que *perdre mon sac* dans (10), l'IMP impose sa valeur continue imparfective en exigeant une réinterprétation en termes de lecture habituelle, comme le montre la version (10a) de notre exemple. L'on pourrait ainsi avancer l'hypothèse selon laquelle la continuité liée à la valeur imparfective de l'IMP signale, dans le cas des achèvements (10) ou d'occurrences bornées (8), la nécessité de compter avec un intervalle de référence dont les bornes temporelles restent implicites, sauf indication spécifique par modifieurs adverbiaux.

4.3. Points de vue aspectuels et relations discursives

Dans plusieurs exemples étudiés, la perte de co-référentialité due au changement d'opérateur temporel se répercute sur le plan discursif. Il se peut par conséquent que tout en restant grammaticalement possible, la substitution PC/IMP ou IMP/PC ne puisse se faire sans modifier les relations discursives qui s'établissent entre les événements décrits. Tel est le cas de l'exemple (11) et de sa version modifiée (11a). Dans (11), *les ouvriers marchaient des heures pour se rendre au travail* a une lecture habituelle qui disparaît suite à la modification effectuée en (11a) :

- (11) *Lorsque Nike s'est implanté il y a dix ans, les ouvriers marchaient des heures pour se rendre au travail ; trois ans plus tard, ils avaient les moyens*

¹⁴ Bien évidemment, *marcher* et *perdre* possèdent des propriétés aspectuelles différentes et désignent des types de procès différents. Cependant, dans les deux cas, la situation à laquelle ils réfèrent est saisie comme une entité comptable, c'est-à-dire pourvue de contours temporels nets. Nous ne pouvons que signaler ce parallélisme qui mérite une étude approfondie.

d'acheter des vélos ; trois ans de plus, et ils y allaient tous en mobylette.
(www.cerclesliberaux.com)

- (11a) *Lorsque Nike s'est implanté il y a dix ans, les ouvriers ont marché des heures pour se rendre au travail ; trois ans plus tard, ils avaient les moyens d'acheter des vélos ; trois ans de plus, et ils y allaient tous en mobylette.*

Les deux exemples ci-dessus sont particulièrement intéressants dans la mesure où ils illustrent les rapports complexes qui relient la valeur aspectuelle associée à un temps et la construction des relations discursives (N. Asher, 1993 ; A. Lascarides, N. Asher, 1993 ; N. Asher, A. Lascarides, 2003) entre les événements décrits dans une séquence de texte. Comme le remarquent W. De Mulder et C. Vetters (1999), la valeur constante des opérateurs temporels ne peut pas être décrite par recours aux relations discursives. De même, on ne peut qu'adhérer à la position de N. Asher (1993), selon laquelle la valeur temporelle et aspectuelle des temps verbaux est l'un des facteurs à prendre en compte dans le calcul des relations discursives¹⁵. Pour ce qui est de notre propos, les énoncés (11) et (11a) montrent clairement qu'un changement d'opérateur temporel entraîne des conséquences sur le plan discursif et que, par conséquent, des contraintes discursives peuvent constituer un obstacle pour la substitution dans un co(n)texte particulier.

Il a été maintes fois souligné que le *rc* est susceptible de recevoir deux lectures qui correspondent à deux valeurs aspectuelles (C. Vikner, 1985 ; D. Creissels, 2000 ; P. Caudal, 2000, 2006 ; P. Caudal, C. Vetters, 2007) : perfective et résultative. Observons comment elles s'articulent dans la subordonnée temporelle *lorsque Nike s'est implanté il y a dix ans*. Isolé de son contexte, le verbe au *rc* peut désigner un événement ponctuel situé avant le moment d'énonciation (*J'implantation de Nike a eu lieu dix ans auparavant*) ou bien renvoyer à un état consécutif à l'installation de l'usine, valide au moment d'énonciation (*elle est implantée depuis dix ans*). Dans le premier cas, on considérera que le *rc* permet d'adopter un point de vue perfectif sur l'événement, alors que dans le second, il s'agit d'un point de vue résultatif. Qu'en est-il dès lors de nos exemples (11) et (11a) ?

En (11), une relation d'Arrière-Plan¹⁶ s'établit entre e_1 (*Nike s'est implanté il y a dix ans*) saisi de manière ponctuelle dans cette configuration discursive, et l'événement

¹⁵ Cf. N. Asher (1993 : 265) : "A number of factors determine discourse relations. Clearly the content of the constituents is relevant. Sometimes tense and aspect are also important in determining discourse structure. [...] The semantics of discourse relations is often complex".

¹⁶ A. Lascarides et N. Asher (1993 : 440) définissent ainsi la relation d'Arrière-Plan (*Background*) entre α et β : "The state described in β is the 'backdrop' or circumstances under which the event in α occurred (no causal connections, but event and state temporally overlap)". Dans notre discussion, α est l'énoncé décrivant e_1 , alors que β désigne e_2 . Pour une définition formelle de la relation d'Arrière-Plan (*Background*), voir : A. Lascarides, N. Asher, 1993 ; N. Asher, A. Lascarides, 2003.

ment multiple e_2 (*les ouvriers marchaient des heures pour se rendre au travail*) qui recouvre un intervalle de référence non borné. Nous admettons que l'opérateur rc le présente dans sa continuité, avec un point de vue imperfectif. Cette période pendant laquelle les ouvriers, faute d'autres moyens, se déplacent à pied est celle d'une pauvreté affichée. En d'autres termes, la phrase habituelle qui décrit e_2 spécifie les circonstances ou le décor dans lequel e_1 a lieu¹⁷. Il n'y pas de relation causale, mais un recouvrement temporel de e_1 par e_2 . Pour ce qui est du cotexte droit, des relations de Narration¹⁸ relient successivement e_2 (*les ouvriers marchaient des heures pour se rendre au travail*) à la situation e_3 (*ils avaient des moyens d'acheter des vélos*) et à e_4 (*ils y allaient tous en mobylette*). Nous proposons de visualiser cette configuration de relations discursives par la figure 1.

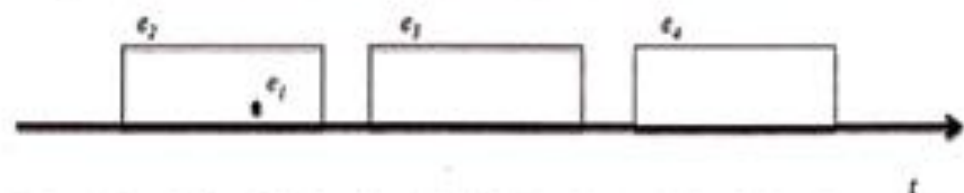


Figure 1. Organisation de relations discursives d'Arrière-Plan et de Narration dans l'exemple (II)

L'introduction du rc dans l'exemple (11a) est à l'origine de nombreux changements. Dans un premier temps, on remarque que l'événement e_2 (*les ouvriers ont marché des heures pour se rendre au travail*) n'y a plus un caractère multiple. Il s'agit, au contraire, d'une simple occurrence d'événement. En d'autres termes, le e_2 de (11) et celui de (11a) diffèrent quant à leur plexité. En effet, les bornes temporelles imposées par le rc présentent e_2 selon un point de vue perfectif, ce qui permet, en (11a) de raconter une tout autre histoire. Le déplacement des ouvriers y apparaît comme une réaction à l'implantation de l'usine. Si l'on se limitait à la relation entre e_1 et e_2 , on conclurait à celle de Narration. En (11a), e_2 (*les ouvriers ont marché des heures pour se rendre au travail*) est présenté comme une conséquence de e_1 (*Nike s'est implanté il y dix ans*), bien qu'il n'y ait pas de relation causale à proprement parler¹⁹. Cependant, les relations discursives en (11a) ne se limitent pas à celle entre e_1 et e_2 . Pour saisir toute leur complexité, il convient d'étendre l'analyse sur le cotexte droit. Il s'avère dès lors que l'implantation de l'usine a produit

¹⁷ Par souci de précision, rappelons que le problème qui nous occupe ici n'est pas celui de relations syntaxiques entre les parties d'une proposition complexe. Il s'agit d'identifier les rapports entre les événements rapportés dans le discours.

¹⁸ A. Lascarides et N. Asher (1993 : 440) caractérisent ainsi la relation de Narration (α , β): "The event described in β is a consequence of but not strictly speaking caused by the event described in α ".

¹⁹ Il s'agit d'une conséquence qui n'est pas strictement impliquée. Elle a donc un caractère non nécessaire de ce que les logiciens appellent *defeasible consequence* (pour une présentation complète, voir N. Asher et A. Lascarides, 2003).

d'autres conséquences, à savoir l'achat de cycles, puis de mobylettes, signes ostensibles d'enrichissement progressif des ouvriers. On constate donc que les relations de Narration s'établissent entre l'événement multiple à caractère habituel e_1 (*les ouvriers ont marché des heures pour se rendre au travail*), l'état e_2 (*ils avaient des moyens d'acheter des vélos*) et e_3 (*ils y allaient tous en mobylette*) qui a également un référent pluriel.

Or, si la situation des ouvriers s'améliore, ce n'est pas parce qu'il y a eu installation de l'usine, mais parce que celle-ci existe toujours au moment d'énonciation (t_1 dans la figure 2). Autant dire que, dans la configuration discursive qu'offre (11a), le PC de Nike s'est implanté autorise une focalisation sur l'état résultant. Autrement dit, cet opérateur temporel permet de porter un regard, un point de vue résultatif sur la situation. Nous l'envisageons dès lors en tant que conséquence toujours valide d'un événement passé, c'est-à-dire de l'installation de l'usine. Avec une telle valeur aspectuelle, les relations discursives sont radicalement différentes. L'événement e_1 correspond à une situation stative que l'on peut résumer par l'énoncé : *l'usine Nike existe et fonctionne*. Dans ce cas, e_2 , e_3 et e_4 entretiennent avec e_1 une relation d'Élaboration : il y a un recouvrement temporel par e_2 des situations e_2 , e_3 et e_4 ²⁰. Le schéma suivant représente dès lors les relations discursives entre les événements rapportés dans (11a) :

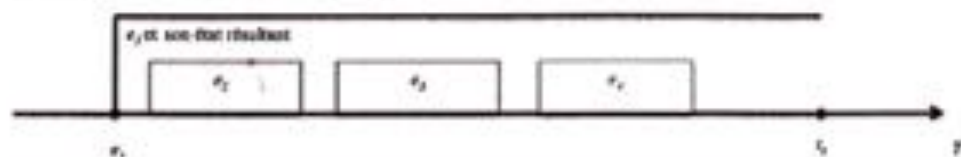


Figure 2. Organisation de relations discursives d'Élaboration et de Narration dans l'exemple (11a)

5. Marquer une étape plutôt que de conclure

Tout au long de ces pages, nous avons essayé de faire apparaître les éléments déterminants pour la substitution PC/IMP et IMP/PC dans les phrases à référent multiple. Il s'avère, en effet, qu'ils relèvent de plans d'analyse différents, bien que strictement interdépendants.

Il a été nécessaire, dans un premier temps, d'identifier les entités qui forment un événement multiple. C'est ainsi que nous avons pu constater que les opérateurs temporels PC et IMP peuvent porter sur l'intervalle de référence ou sur les occurrences itérées e_1 , e_2 , e_3 , ... En ce qui concerne la présentation de ces deux

²⁰ Selon A. Lascarides et N. Asher (1993 : 440), nous avons affaire à la relation d'Élaboration entre α et β , quand " β 's event is part of α 's event".

types d'éléments dans un énoncé, plusieurs contraintes doivent être prises en compte.

Dans la section 3.2, nous avons remarqué que la nature de la pluralité événementielle impose le choix d'un opérateur perfectif dans les phrases itératives. D'une part, l'intervalle de référence y compte peu, de l'autre la quantification numérique semble exiger une saisie globale des occurrences e_p, e_p, e_p, \dots . Les énoncés habituels autorisent une plus grande liberté, pour ce qui est de l'opérateur temporel, sauf si des contraintes particulières pèsent sur la structuration de l'intervalle de référence, comme dans l'exemple (5). Les phrases fréquentatives apparaissent comme les plus tolérantes à l'égard des substitutions PC/IMP et IMP/PC . La raison en est à chercher, nous semble-t-il, dans la nature même de la pluralité fréquentative. De par le caractère *sine qua non* inclusif de l'intervalle de référence, les opérateurs temporels ne peuvent agir que sur les occurrences, en les présentant soit d'un point de vue perfectif ou imperfectif.

Plus d'une fois, les exemples discutés font ressortir le rapport étroit qui s'établit entre les opérateurs temporels PC ou IMP et des phénomènes aspectuels tels que bornage (temporel et/ou spatial) et télicité. Enfin, comme le montre l'exemple (11), il est quelquefois nécessaire de composer avec les relations discursives pour lesquelles une substitution PC/IMP ou IMP/PC n'est pas indifférente.

Face à une telle multitude et diversité de facteurs, il nous semble prudent d'éviter un exposé de résultats définitifs. Tout au plus avons-nous marqué une étape, en identifiant un certain nombre de problèmes que sous-tend la substitution d'opérateurs temporels dans des phrases à référent événementiel pluriel. Autant dire qu'il s'agit plutôt de points de départ pour l'exploration de ces pistes entremêlées.

Références

- Apothélos D., Nowakowska M., à paraître : « Note sur la résultativité et la valeur de parfait leur expression en polonais ». *Cahiers Chronos*.
- Asher N., 1993 : *Reference to Abstract Objects in Discourse*. Dordrecht, Kluwer.
- Asher N., Lascarides A., 2003 : *Logics of Conversation*. Cambridge, Cambridge University Press.
- Bach E., 1986 : "The Algebra of Events". *Linguistics and Philosophy*, 9, 5—16.
- Borillo A., 1991 : « De la nature compositionnelle de l'aspect ». *Travaux de linguistique et de philologie*, 29, 97—102.
- Carlson G., 1982 : "Generic Terms and Generic Sentences". *Journal of Philosophical Logic*, II, 145—181.
- Carlson G., 1995 : *The Generic Book*. Chicago, The University Press of Chicago.
- Caudal P., 2000 : « La polysémie aspectuelle — contraste français / anglais ». [Thèse de doctorat]. Université Paris VII — Denis Diderot.

- Caudal P., 2006 : « Aspect ». Dans : D. Godard, L. Roussarie, F. Corblin, eds : *Sémantico-pédagogie : dictionnaire de sémantique*. GDR Sémantique & Modélisation, CNRS.
- Caudal P., Vetters C., 2007 : « Passé composé et passé simple : Sémantique diachronique et formelle ». *Cahiers Chronos*, 16, 121—151.
- Combettes B., François J., Noyau C., Vet C., 1993 : « Introduction à l'étude de l'aspect dans le discours narratif ». *Verbum*, 4, 5—48.
- Creissels D., 2000 : « L'emploi résultatif de 'être + participe passé' en français ». *Cahiers Chronos*, 6, 133—142.
- Dahl Ö., 1975 : « Remarques sur le générique ». *Langages*, 79, 55—60.
- Dahl Ö., 1981 : "On the Definition of the Telic-Atelic (Bounded-Unbounded) Distinction". In : P.J. Tedeschi, A. Zaenen, eds : *Syntax and Semantics. Tense and Aspect*, 14. New York, Academic Press, 79—90.
- Davidson D., 1967 : "The Logical Form of Action Sentences". In : N. Rescher, ed.: *The Logics of Decision and Action*. Pittsburgh, University of Pittsburgh Press.
- Davidson D., 1993 : « Forme logique des phrases d'action ». Dans : D. Davidson : *Actions et événements*, Presses universitaires de France.
- De Mulder W., Vetters C., 1999 : « Temps verbaux, anaphores (pro)nominales et relations discursives ». *Travaux de linguistique*, 39, 37—58.
- De Swart H., 1987 : « Phrases habituelles et sémantique des situations ». Dans : G. Kleiber, éd. : *Rencontres(s) avec la générique*. Paris, Klincksieck, 261—279.
- De Vogüe S., 1990 : « Valeurs de l'imparfait : pour une solution plus modale que temporelle ». *Studia Romanica Posnaniensia*, 14, 175—193.
- De Vogüe S., 1993 : « Des temps et des modes ». *Le Gré des Langues*, 6, 65—91.
- Gosselin L., 1996 : *Sémantique de la temporalité en français : un modèle calculatoire et cognitif du temps et de l'aspect*. Louvain-la-Neuve, Duculot.
- Grzegorzczkowska R., Zaron Z., red., 1997 : *Semantyczna struktura tekstu i wypowiedzi*. Warszawa, Wydawnictwo Uniwersytetu Warszawskiego.
- Karolak S., 1984 : „Składnia wyrażen predykatywnych”. W: Z. Topolińska, M. Grochowski, S. Karolak: *Gramatyka współczesnego języka polskiego. Składnia*. Warszawa, PWN, 11—172.
- Karolak S., 1997 : *Semantika i struktura sinjanskiego vida*. Kraków, Wydawnictwo Naukowe Wyższej Szkoły Pedagogicznej.
- Karolak S., 2001 : *Od semantyki do gramatyki : wybór rozpraw*. Warszawa, Slawistyczny Ośrodek Wydawniczy.
- Karolak S., 2005 : *Struktura i semantyka aspektu w językach naturalnych*. Kielce, Wyższa Szkoła Umiejętności.
- Karolak S., à paraître : « L'aspect dans une langue : le français ». *Travaux de linguistique*.
- Kleiber G., 1987 : *Du côté de la référence verbale. Les phrases habituelles*. Berne, Francfort/M, New York, Paris, Peter Lang.
- Kozłowska M., 1998a : « Aspect, mode d'action et classes aspectuelles ». Dans : J. Moeschler, dir. : *Le temps des événements. Pragmatique de la référence temporelle*. Paris, Editions Kimé, 101—122.
- Kozłowska M., 1998b : « Bornage, télécité et ordre temporel ». Dans : J. Moeschler, dir. : *Le temps des événements. Pragmatique de la référence temporelle*. Paris, Editions Kimé, 221—244.

- Lascarides A., Asher N., 1993: "Temporal Interpretation, Discourse Relations and Commonsense Entailment". *Linguistics and Philosophy*, 16, 437—493.
- Mindak J., 1988: « Perfectivité ». Dans: M. Bracops, éd.: *Équivalences. Généricité, spécificité et aspect. Revue de l'Institut supérieur de Traducteurs et Interprètes de Bruxelles*. Vol. 17/1-2-3, vol. 18/1, 151—162.
- Moens M., Steedman M., 1988: "Temporal Ontology and Temporal Reference". *Computational Linguistics*, 14, 2, 15—28.
- Moeschler J., dir.: 1998: *Le temps des événements. Pragmatique de la référence temporelle*. Paris, Editions Kimé.
- Partee B., 1999: "Some Remarks on Linguistic Uses of the Notion of Event". In: C. Tenny, J. Pustejovsky, eds: *Events as Grammatical Objects. The Converging Perspectives of Lexical Semantics and Syntax*. Stanford, CSLI Publications, 493—495.
- Sikora D., 2009: *Les Verbes de manière de mouvement en polonais et en français. Éléments pour une étude comparée des propriétés structurelles des prédicats*. [Thèse de doctorat] Nancy, Université Nancy 2 — ATILF — CNRS (UMR 7118).
- Smith C., 1986: "A Speaker-based Approach to Aspect". *Linguistics and Philosophy*, 9, 97—115.
- Smith C., 1991: *The Parameter of Aspect*. Dordrecht, Boston, London, Kluwer.
- Sypnicki J., Vetulani G., 1996: « Sur l'aspect en français et en polonais ». *Studia Romanica Posnaniensia*, 21, Poznań, UAM, 115—122.
- Talmy L., 2000: *Toward a Cognitive Semantics*. Vol. 1: *Concept Structuring System*. Vol. 2: *Typology and Process in Concept Structuring*. Harvard, MIT Press.
- Tenny C., 1994: *Aspectual Roles and the Syntax-Semantics Interface*. Dordrecht, Boston, Londres, Kluwer Academic Publishers.
- Vendler Z., 1967: "Verbes and Times". *Linguistics in Philosophy*, 97—121.
- Verkuyl H.J., 1972: *On the Compositionnal Nature of Aspect*. Dordrecht, Reidel.
- Verkuyl H.J., 1993: *A Theory of Aspectuality. The Interaction between Temporal and Atemporal Structure*. [Cambridge Studies in Linguistics, 64]. Cambridge, Cambridge University Press.
- Verkuyl H.J., 1999: *Aspectual Issues: Studies on Time and Quantity*. Stanford, CA, CSLI Publications.
- Verkuyl H.J., Vet C., Borillo A., Bras M., Le Draoulec A., Molendijk A., De Swart H., Vetter C., Vieu L., 2004: "Tense and Aspect in Sentences". In: F. Corblin, H. De Swart, eds: *Handbook of French Semantics*. Stanford, CSLI Publications, 233—270.
- Vet C., 1980: *Temps, aspect et adverbies de temps en français contemporain*. Genève, Droz.
- Vet C., 1992: « Le passé composé: contextes d'emploi et interprétation ». *Cahiers de pragmatique*, 19, 37—59.
- Vet C., 1993: « Peut-on calculer l'interprétation des syntagmes nominaux et des prédications verbales ». Dans: A. Hulk, F. Melka, J. Schrotten, réd.: *Du lexique à la morphologie: de côté de chez Zwaaen*. Amsterdam—Atlanta, Rodopi.
- Vet C., 1994: « Petite grammaire de l'Aktionsart et de l'aspect ». *Cahiers de grammaire*, 19, 1—17.
- Vet C., 1995: « Structures discursives et interprétation du discours ». *Modèles linguistiques*, 16, 2, 111—122.

Vetters C., 1996: *Temps, aspect et narration*. Amsterdam, Rodopi.

Vetters C., De Mulder W., 2000: « Passé simple et imparfait : contenus conceptuel et procédural », *Cahiers Chronos*, 6, 13—36.

Vikner C., 1985: « L'aspect comme modificateur du mode d'action : à propos de la construction 'être + participe passé' », *Langages*, 67, 95—113.

Portée des préfixes et complexité en français et en polonais. Le cas de *RE-* versus *PO-*, *PRZY-*, *DO-*

0. Complexité et préfixes

À l'origine de cette réflexion¹ se trouve le constat d'une difficulté récurrente dans un cours de langue polonaise : même les étudiants francophones familiers avec la distinction entre les formes perfectives et imperfectives du verbe évitent systématiquement les premières au profit des secondes dans leurs productions spontanées. Ainsi, *przeczytać* (*lire*) est remplacé par *czytać* (*lire*), *zjeść* (*manger*) ou *najeść się* (*manger à sa faim*) par *jeść* (*manger*), alors que *pójść* (*partir à pied*), *przyjść* (*venir à pied*) se voient régulièrement préférer leur base, c'est-à-dire *iść* (*aller à pied, marcher*). Face à ce constat, il est naturel de s'interroger sur les raisons de ce rejet.

Il est peu probable que ces raisons soient uniquement de nature morphologique dans la mesure où la dérivation verbale par préfixes s'effectue de manière similaire en français et en polonais. Cela correspond à la définition générale de la dérivation proposée par Apothéloz (2002 : 73) : « c'est une opération consistant à composer une base et un affixe ». Mais au-delà du parallélisme formel des processus eux-mêmes aboutissant, en français comme en polonais, à des formes composées, réalise-t-on effectivement

* Dorota SIKORA – ATILF – CNRS (UMR 7118).

¹ Je tiens à remercier chaleureusement Hanne Leth Andersen, Eva Havu, Amr Helmy Ibrahim, Agnieszka Kaliska, Claire Martinot et Urszula Paprocka-Piotrowska pour les discussions passionnantes lors du colloque international « La complexité en langue et son acquisition » (Paris, 5-7 juillet 2010, Université Paris VII Descartes) qui m'ont permis d'explorer de nouveaux chemins. Les éventuels errements sont entièrement de mon fait.

la/les même(s) opération(s) sur le plan sémantique et grammatical ? La difficulté qu'éprouvent les étudiants francophones, à faire usage de dérivés polonais ne serait-elle pas liée aux effets produits par les préverbes – effets bien différents de ceux qu'ils observent dans leur propre langue ?

Nous avons admis d'entrée de jeu que la difficulté de maniement des préfixes verbaux est un cas de complexité. En suivant Dahl (2004 cité dans Andersen 2010), nous nous proposons de mesurer la complexité d'un objet, en l'occurrence linguistique, par sa description la plus courte. L'approche de Dahl implique la marche à suivre. Les éléments indispensables pour décrire un phénomène complexe doivent être identifiés et articulés. Dans un travail à visée contrastive, il s'agira ensuite de comparer des descriptions d'objets soumis à l'analyse dans le but de vérifier si elles impliquent les mêmes éléments descriptifs et si, le cas échéant, ceux-ci sont organisés de manière similaire. En observant une poignée de préfixes, nous nous efforcerons de répondre à la question de savoir si la complexité liée à l'action de préfixes français et polonais est de même nature ou si, au contraire, elle implique des éléments différents.

Il serait présomptueux de prétendre traiter, dans le cadre d'un article, l'ensemble de préverbes français et polonais, d'autant qu'il ne s'agit aucunement d'étudier chacun d'entre eux dans la multitude de ses combinaisons possibles avec des bases verbales. Nous nous sommes imposé une double limitation. D'une part, nous nous intéresserons à un nombre fort réduit de préfixes disponibles dans les deux langues (*RE-* en français, *PO-*, *PRZY-*, *DO-* en polonais). De l'autre, seuls les verbes appartenant à la classe de verbes de manière de mouvement² seront pris en compte en tant que bases susceptibles

² Nous sommes consciente des désaccords au sujet de cette classe verbale. Considérés comme une sous-classe de verbes locatifs (Gross 1975, Guillet et Leclère 1992) à valeur médiane (Boons 1985, 1987), ils sont classés par d'autres auteurs (Borillo 1998) parmi les verbes de déplacement, dans la mesure où ils forment des constructions *Nocible Vdyn Prép NIsite*. Ailleurs (Asher et Sablayrolles 1995, Aunargue 2008), ils sont qualifiés de verbes de changement d'emplacement, puisqu'ils « dénotent un changement d'emplacement par rapport au cadre de référence terrestre » (Aunargue 2008 : 1906). L'auteur montre néanmoins que le changement d'emplacement se conjuguant, dans des cas fréquents, avec celui de relation locative, il n'est pas impossible de recourir, dans le label accordé à cette classe verbale, au terme de *déplacement* (cf Aunargue 2008). Notons également les divergences s'agissant de déterminer l'extension de cette classe : certains y intègrent *avancer*, *zigzaguer* (Aunargue 2008, Stosic 2009), alors que pour d'autres (Herslund 1990, Sikora 2009), ils relèvent d'une catégorie différente. Enfin, il convient de signaler la compatibilité de certains d'entre ces verbes, du moins en

de sélectionner les affixes susmentionnés. Nous commencerons, dans la première section, par présenter notre démarche, en espérant justifier nos choix, notamment en ce qui concerne les éléments retenus ou non pour analyse. La section 2 sera consacrée à la préfixation en *RE-* et aux modifications qu'elle produit. Il s'agira de comparer le dérivé et sa base. La même démarche sera adoptée, dans la troisième section, pour les verbes polonais. Chemin faisant, nous espérons montrer que les préfixes verbaux français et polonais opèrent sur des plans différents et permettent de véhiculer des informations de nature non similaire, du moins en ce qui concerne *RE-* d'un côté, *PO-*, *PRZY-*, *DO-* de l'autre. Dans la quatrième et dernière section, nous réunissons les éléments dégagés au fil de notre étude afin de montrer que, du fait de leur portée différente, le préfixe *RE-* d'un côté, *PO-*, *PRZY*, *DO-* de l'autre, illustrent des cas de complexités propres à chacun des deux systèmes linguistiques.

1. Choix terminologiques et *modus operandi*

1.1. Délimitation du champ d'analyses

Face à l'abondante littérature consacrée aux préfixes slaves, dont polonais, notre démarche paraît inévitablement modeste. D'une part, nous traitons un nombre réduit de préfixes, de l'autre nous nous sommes limitée à des verbes relevant des classes équivalentes dans les deux langues. Ainsi, allons-nous prendre en considération *marcher*, *rouler*, *courir*, *nager*, *voler*, *ramper*, verbes qui sont rendus en polonais respectivement par les verbes déterminés *ić*, *jechać*, *płynąć*, *biec*, *lecieć*, *pełznąć* ou par *chodzić*, *jeździć*, *pływać*, *biegać*, *latać*, *pełzać*, connus sous la dénomination de verbes indéterminés.

La limitation dans le choix aussi bien des préfixes que des verbes retenus pour analyse tient à deux raisons. Premièrement, compte tenu de notre objectif, nous ne prétendons pas fournir une description exhaustive des préfixes français et polonais, tâche à laquelle s'étaient déjà attelés plusieurs auteurs et dont l'ampleur serait ici démesurée³. Notre but consiste à comparer le verbe-base et son/ses dérivé(s) afin de repérer les éléments modi-

français, avec l'adverbial *sur place* qui conduit quelquefois à contester leur statut de verbes de déplacement, voire de mouvement, pour parler plutôt de verbes d'activité corporelle (Asnès et Kupferman 2008).

³ A titre d'exemple, la description sémantique du seul préfixe *DO-* proposée par Przybylska (2006) comprend 35 pages. Par ailleurs, l'ouvrage n'est pas consacré à l'ensemble de préfixes, mais à cinq d'entre eux.

fiés sous l'effet du préfixe, en déterminant quels éléments forment leurs descriptions les plus courtes. Conformément aux propositions de Dahl (2004), elles formeront par la suite une base permettant d'évaluer si une même opération morphologique de préfixation verbale aboutit à des formes dont la complexité s'organise de manière identique. Deuxièmement, la focalisation sur les classes de verbes correspondantes en français et en polonais est, à notre avis, la condition *sine qua non* d'une certaine cohérence dans les résultats. Si leurs membres s'apparentent par leur sens lexical et, par conséquent, par le type de procès qu'ils dénotent, on peut raisonnablement s'attendre à des régularités dans les interactions entre les préfixes et leur sémantisme.

1.2. Dénominateur aspectuel commun

Aussi bien *RE-* en français que l'ensemble des préfixes polonais sont d'un commun accord qualifiés d'aspectuels. D'une part, il est courant d'analyser *RE-* en termes d'itération ou de duplication, de l'autre *PO-*, *PRZY-*, *DO-* servent, entre autres, à dériver les formes perfectives des bases imperfectives. Dans les deux cas, il s'agit incontestablement de caractéristiques aspectuelles, mais de phénomènes bien distincts. Dans un travail à visée comparative comme le nôtre, le premier pas consiste à fixer un dénominateur aspectuel commun que l'on pourra décrire avec des outils valides pour les deux langues. C'est la raison pour laquelle il est nécessaire d'éclaircir quelques points terminologiques liés aux paramètres aspectuels.

Ainsi, parlerons-nous d'aspect lexical (*Aktionsart*), lorsqu'il s'agira de caractériser le schéma temporel associé au sens véhiculé par un verbe (*time schemata* Vendler 1967 : 98-99). Ces schémas se laissent décrire par des traits tels que dynamité, (a)télicité, durativité et répartir en nombre limité de classes ou types (Smith 1991). Pour désigner de manière générique les états, les activités, les accomplissements, les achèvements identifiés par Vendler (1967), ainsi que les autres types proposés dans les classements établis par différents auteurs, nous recourons au terme de *situation* introduit par Smith (1991)⁴.

⁴ La typologie proposée par Vendler (1967) a été largement repensée et retravaillée. Padučeva (1996), entre autres, en a démontré l'inadéquation aux systèmes verbaux slaves. Dans la lignée de Padučeva (1996) et de Grzegorzczkova (1998), Laskowski (1998) propose une typologie de situations propre au polonais. Cependant, dans un travail comparatif, il nous faudrait une typologie adéquate aux deux langues étudiées,

Nous nous abstenons, en revanche, de parler d'aspect grammatical pour désigner l'imperfectivité ou la perfectivité d'une forme verbale, que ce soit en français ou en polonais. Nous pensons qu'il est plus judicieux de mener une analyse comparative selon une approche énonciative de l'aspect proposée par Smith (1991). Ainsi, en choisissant parmi les moyens que la langue met à sa disposition, le locuteur adopte un point de vue – imperfectif ou perfectif⁵ – sur la situation associée à la forme verbale. Or, il est fréquent que les langues diffèrent entre elles par les moyens dont elles disposent pour exprimer le point de vue aspectuel. En français, celui-ci est véhiculé essentiellement par les temps verbaux, alors qu'en polonais, il est codé dans la forme du verbe dont il détermine les caractéristiques grammaticales (par exemple, son paradigme flexionnel).

Malgré un exposé inévitablement sommaire, ces distinctions nous permettront, dans les sections suivantes, d'observer les éléments qui tombent dans la portée d'un préfixe et nommer les effets produits.

2. RE- et verbes de manière de mouvement français

Le préfixe *RE-* est quasi unanimement qualifié d'aspectuel (Apothéloz 2005, 2007, Wilmet 2007, Mascherin 2007)⁶. Selon Wilmet (2007 : 343), il assume une fonction duplicative intermittante. Remarquons au passage que *RE-* est un préfixe particulièrement productif en français actuel. Certains dérivés des verbes de manière de mouvement étudiés ici sont répertoriés dans le Trésor de la Langue Française informatisé (*recourir, remarquer, revoler*). D'autres (*rerouler, renager, reramper*), bien qu'absents de ce dictionnaire de référence, se trouvent largement attestés, notamment dans les écrits électroniques qui nous ont fourni les exemples (1) à (3) ci-dessous :

ce qui constitue un sujet d'étude bien plus large que ne l'autorise la taille d'un article. C'est la raison pour laquelle nous préférons insister sur les paramètres aspectuels (par exemple la durativité) qui changent avec les préfixes.

⁵ Nous ne traiterons pas ici du point de vue résultatif discuté, entre autres, dans Caudal et Veters (2007).

⁶ Il convient cependant de signaler les objections formulées par Amiot (2002) soulignant que, dans certains cas, *RE-* ne porte pas sur la situation dans sa totalité, mais sur l'un des actants. Or, si tel est le cas, le préfixe n'a pas un caractère aspectuel.

- (1) Bonjour, enfin je peux rerouler avec mon bébé, suite à mon accident le 30 août 2004. (www.forum-auto.com)
- (2) J'ai renagé hier après presque 3 semaines d'arrêt. Je crois qu'il va me falloir une bouée. (sptri.forumactif.com)
- (3) Elle s'assoit, se retourne des deux côtés, se met sur ses quatre pattes, et avance deux pas, puis elle rampe et se remet à quatre pattes et tombe et rrampe, etc ... (www.etrenceinte.com)

Les dérivés font preuve d'une grande régularité en ce qui concerne les effets de sens produits par l'interaction de *RE-* avec le sémantisme des verbes de manière de mouvement. Amiot (2002) en propose la paraphrase suivante : « à nouveau, avec retour à la situation antérieure ». L'auteur de (1) pourrait, en effet, dire sans trahir sa pensée : « Je peux rouler à nouveau après l'arrêt dû à l'accident », alors que celui/celle de (2) déclare, en effet, avoir nagé à nouveau. De même, « elle rampe à nouveau après être tombée » reprend fidèlement la signification de *reramper* dans (3). Ainsi, la valeur sémantique des dérivés en *RE-* est calculable, ce qui conduit Apothéloz (2005) à proposer le traitement de ce préfixe en termes de foncteur lexical représenté par la formule « à nouveau x » où la variable x n'est autre que la base de dérivation.

En réunissant les résultats mis en évidence par les auteurs cités, essayons à présent d'identifier les éléments qui, pris dans la portée du préfixe (foncteur) *RE-*, changent sous son effet. Comparons l'énoncé (4a) extrait de Frantext et sa version obtenue par dérivation pour les besoins de notre exposé.

- (4) a. Nous avons marché vers la gare (Frantext, Schneider, B. *Un silence d'environ une demi-heure*, 1996)
- b. Nous avons remarqué vers la gare.

La paraphrase proposée par Amiot (2002) est valide pour *remarquer* de (4b). En ce qui concerne *marcher* et *remarquer*, on constate des différences importantes dans leurs fonctionnements respectifs sur le plan sémantique et, par conséquent, référentiel. La base renvoie à une séquence de marche à pied qui, dans (4a), se déroule selon une trajectoire définie, *ipse eo* à un déplacement, et qui couvre une séquence temporelle t_1 . Or, l'interprétation du dérivé en *RE-* repose sur un ensemble de présupposés. Ainsi, *remarquer* dans (4b) dénote l'événement e_2 , précédé par au moins un autre e , mettons e_1 , de même type. Sa signification repose donc sur le présupposé selon lequel un déplacement e_1 s'était achevé avant que e_2 survienne. La séquence

temporelle t_1 , qui correspond à e_1 , est donc suivie d'une autre période, mettons t_{1+n} , qui marque l'arrêt de la marche.

Il est important de remarquer qu'en produisant un effet d'itération, *RE-* n'atteint aucunement l'aspect lexical, pas plus qu'il n'affecte le point de vue aspectuel. *Marcher* de (4a) et *remarcher* de (4b) se voient associer un seul et même type de situation que l'on qualifiera d'activité (Vendler 1967, Dowty 1979, Vet 1980, Borillo 1987) ou de procès (Mourelatos 1978, Bach 1986) selon le classement adopté. Par ailleurs, le point de vue perfectif ou imperfectif sur la situation ne change pas avec le préfixe. Ce point mérite d'être souligné car il s'avère crucial pour les différences observées entre la portée de *RE-* et celle des préfixes polonais.

Si comme le propose Dahl (2004, cité dans Andersen 2010), la complexité d'un objet, en l'occurrence celle d'une forme linguistique, se mesure à l'aune de sa description la plus courte, il va sans dire que le paradigme dérivationnel qui nous occupe ici produit des formes complexes. La comparaison du verbe-base (*marcher*) et de son dérivé (*remarcher*) montre que la complexification formelle de ce dernier se caractérise par une densité d'informations plus élevée. Pour décrire un dérivé de verbes de manière de mouvement en *RE-*, il faut rendre compte des présupposés sur lesquels repose l'effet itératif.

Observons à présent comment un mécanisme morphologiquement similaire opère sur les verbes de manière de mouvement en polonais.

3. Préfixes polonais et verbes de manière de mouvement

3.1. Trois préfixes, deux modes de fonctionnements et deux sous-classes verbales

Les équivalents polonais de *marcher*, *rouler*, *courir*, *nager*, *voler*, *ramper* se trouvent répartis en deux classes. Nous avons proposé ailleurs (Sikora 2009) un examen détaillé du sémantisme des deux sous-classes polonaises, aussi limiterons-nous à présent à une présentation sommaire.

Les verbes *isć*, *jechać*, *płynąć*, *biec*, *lecieć*, *pełznąć*, qualifiés de déterminés, lexicalisent les mêmes manières de déplacement que leurs correspondants français ci-dessus, tout en signifiant que la progression du mouvement est organisée en une trajectoire⁷. La seconde classe réunit les verbes

⁷ Pour un exposé complet, voir Sikora 2009.

dits indéterminés *chodzić, jeździć, pływać, biegać, latać, pełzać* également capables de rendre le sens de *marcher, rouler, courir, nager, voler, ramper*, mais le mouvement dont la manière est spécifiée dans leur sémantisme ne suppose aucune organisation linéaire en trajectoire.

Le traitement des exemples que nous proposons dans cette partie diffère de celui que nous avons pratiqué dans le cas de (1), (2), (3) et (4) dans la section précédente. Il est en effet difficile, voire impossible, de trouver des exemples d'un seul et même verbe-base successivement préfixé par *PO-*, *DO-*, *PRZY-* dans un co-texte strictement identique à chaque fois. Choisir au contraire des énoncés authentiques formés par les dérivés d'une même base implique des co-textes différents et entraîne la nécessité d'analyser systématiquement les interférences spécifiques à chaque exemple.

Afin de rendre les éléments soumis à l'analyse plus saillants, en évitant des interférences contextuelles, nous avons choisi de retenir un exemple de phrase simple. Le verbe déterminé *jechać* (exemple 5a) nous a servi, dans un premier temps, de base pour en dériver des formes en *PO-*, *PRZY-*, *DO-* (exemples 5b, 5c, 5d). Dans (6), nous lui avons substitué son correspondant indéterminé *jeździć* pour lui adjoindre ensuite, dans (6b), (6c) et (6d), les mêmes préfixes.

- (5) a. ... kierowca **jechał** do Lublina (nkjp.pl, texte de presse)
 ... le conducteur *JECHAĆ*_[passé masculin imperfectif] à Lublin
 ... *le conducteur allait à Lublin (à bord d'un moyen de transport terrestre)*⁸
- b. ... kierowca **pojechał** do Lublina
 ... le conducteur *POJECHAĆ*_[passé masculin perfectif] à Lublin
 ... *le conducteur est parti pour Lublin (à bord d'un moyen de transport terrestre)*
- c. ... kierowca **przyjechał** do Lublina
 ... le conducteur *PRZYJECHAĆ*_[passé masculin perfectif] à Lublin
 ... *le conducteur est venu à Lublin (à bord d'un moyen de transport terrestre)*
- d. ... kierowca **dojechał** do Lublina
 ... le conducteur *DOJECHAĆ*_[passé masculin perfectif] à Lublin
 ... *le conducteur a fini par arriver à Lublin (à bord d'un moyen de transport terrestre)*

⁸ Les traductions françaises fournies en petites italiques peuvent à elles seules constituer un sujet de débat que nous ne sommes malheureusement pas en mesure d'entreprendre ici. Elles sont destinées uniquement à aider le lecteur non polonophone.

- (6) a. ... kierowca **jeździł** do Lublina
 ... le conducteur JEŹDŹIĆ_[passé masculin *imperfectif*] à Lublin
 ... *le conducteur allait à Lublin (à bord d'un moyen de transport terrestre)*
- b. ... kierowca **pojeździł** do Lublina
 ... le conducteur POJEŹDZIĆ_[passé masculin *perfectif*] à Lublin
 ... *pendant une certaine période, le conducteur allait à Lublin (à bord d'un moyen de transport terrestre)*
- c. ... kierowca **przyjeźdzał** do Lublina
 ... le conducteur PRZYJEŹDŹAĆ_[passé masculin *imperfectif*] à Lublin
 ... *le conducteur venait à Lublin (à bord d'un moyen de transport terrestre)*
- d. ... kierowca **dojeźdzał** do Lublina
 ... le conducteur DOJEŹDŹAĆ_[passé masculin *imperfectif*] à Lublin
 ... *le conducteur finissait par arriver à Lublin (à bord d'un moyen de transport terrestre)*

Les trois préfixes analysés ici représentent deux modes de fonctionnement différents. D'une part, *PO-* opère d'une manière que nous qualifions d'externe, car il a avant tout une portée délimitative sur la situation désignée par la base. De l'autre, *PRZY-* et *DO-* interagissent avec le sémantisme du verbe, en modifiant les rapports entre les actants. C'est la raison pour laquelle nous proposons de considérer qu'il s'agit d'un mode de fonctionnement interne.

3.2. *PO-* et le mode de fonctionnement externe

Commençons par observer l'effet produit par *PO-* sur les verbes déterminés dans les exemples (5a) et (5b). Dans (5a), le verbe base *jechać* (*aller à bord d'un moyen de transport terrestre, rouler*), de par son imperfectivité, dénote un procès en cours de déroulement, progressant vers Lublin, terminus désigné par le complément de lieu. Le chauffeur *x* avance selon une trajectoire qui, si rien n'interrompt sa progression, le conduira à Lublin (*y*) de manière qu'une relation locative *x est à y* pourra être établie. Ce procès suppose donc une certaine durée. Avec le dérivé en *PO-* dans (5b), nous avons affaire, au contraire, à une forme qui véhicule le point de vue perfectif sur le procès. Dans (5b), le préfixe impose des bornes qui « découpent » la partie initiale du déplacement, celle qui correspond au début

de la trajectoire, en conférant au dérivé une valeur inceptive. Si nous avons proposé de qualifier le fonctionnement de *PO-* d'externe, c'est parce que son action consiste avant tout à imposer des bornes à une situation qui en est initialement dépourvue. La perfectivisation du point de vue aspectuel s'accompagne néanmoins d'une modification importante en ce qui concerne le contenu notionnel du prédicat. En (5a), le verbe déterminé *jechać* vise une relation locative finale entre *x* et *y*. Dans (5b), la trajectoire se trouve réduite par *PO-* à son point initial, ce qui modifie sensiblement la relation établie par le prédicat entre les actants : il relie désormais *x* et le point initial quitté au moment où le déplacement commence. Si notre analyse est exacte, une hiérarchie de modifications réalisées par *PO-* semble se dégager. L'introduction des bornes transforme les relations entre les actants et, par conséquent, le type de situations associé au prédicat. *Jechać* de (5a) est bien caractérisé par une durée, trait dont *pojechać* dans (5b) est dépourvu. On y retrouve une organisation particulière des trois parmi les quatre⁹ fonctions que Klemensiewicz (1951) identifie en analysant les préfixes polonais : i. aspectuelle perfectivisante (perfectivisation par bornage), ii. sémantique consistant, dans notre exemple, en une modification de relation entre les actants et, de ce fait, de sens lexical qui, à son tour, iii. entraîne le changement d'aspect lexical.

C'est également une fonction aspectuelle bornante qu'assume le préfixe *PO-*, lorsqu'il se combine avec les verbes indéterminés *chodzić, jeździć, pływać, biegać, latać, pełzać*. Reprenons nos exemples (6a) et (6b).

- (6) a. ... kierowca **jeździł** do Lublina
 ... le conducteur JEŹDŹIĆ_[passé masculin imperfectif] à Lublin
 b. ... kierowca **pojeździł** do Lublina
 ... le conducteur POJEŹDŹIĆ_[passé masculin perfectif] à Lublin

La seule interprétation possible de (6a) est itérative, habituelle conformément aux distinctions établies par Kleiber (1987). (6a) signifie que le conducteur a plusieurs fois effectué le déplacement jusqu'à Lublin. La forme verbale est imperfective, mais l'imperfectivité concerne, dans le cas des verbes indéterminés sélectionnant un complément de lieu¹⁰, non pas

⁹ La quatrième fonction, assez éloignée de nos préoccupations, consiste à dériver des verbes dénominaux. Cf. Klemensiewicz (1951).

¹⁰ Il est important de préciser cette configuration syntaxique. La lecture plurielle de certains verbes imperfectifs polonais est, certes, bien connue, mais dans la plupart des cas, elle a besoin d'adverbiaux quantifiants (*souvent, toujours, rarement*, etc) pour

l'un des déplacements itérés, mais l'habitude elle-même. Son intervalle de référence (Kleiber 1987), c'est-à-dire la période pendant laquelle le chauffeur réitère ses voyages, est situé dans le passé par rapport au moment d'énonciation et présenté sans borne initiale ni finale. En d'autres termes, le point de vue imperfectif porte sur l'intervalle de référence.

Avec la forme dérivée *pojeździć* dans (6b), la lecture habituelle est bel et bien préservée, alors que le verbe est perfectif. De fait, *PO-* délimite l'intervalle de référence en marquant ses bornes initiale et finale. Les déplacements successifs effectués par le chauffeur se reproduisent tout au long d'un intervalle de référence situé dans le passé, comme en (6a), mais saisi globalement, fermé à gauche comme à droite.

Si le préfixe *PO-* opère, dans le cas de verbes indéterminés, de manière externe en imposant des bornes à une situation qui en est dépourvue, la perfectivisation qui en résulte diffère de celle que nous avons observée en (5b). Dans (6b), la perfectivité suppose la durativité sur le plan de l'aspect lexical. Il convient de souligner qu'en ce qui concerne les verbes indéterminés, la portée de *PO-* est limitée au point de vue aspectuel. Le préfixe n'affecte pas les relations entre les actants, pas plus qu'il ne change les propriétés caractérisant l'aspect lexical.

On terminera cette section en concluant que malgré un fonctionnement que l'on peut, dans les deux cas, qualifier d'externe et bornant, la portée de *PO-* est nettement plus étendue lorsque ce préfixe se combine avec les verbes déterminés. Si, en (5b) et en (6b), *PO-* assume bien la fonction aspectuelle consistant à imposer des bornes et en créant des formes perfectives, la perfectivité prend des formes différentes. Avec les verbes déterminés (l'exemple *pojechać* dans 5b), la fonction aspectuelle bornante produit des modifications jusque dans le sens et dans la nature du procès associé au prédicat. *Pojeździć* de (6b) montre, au contraire, qu'appliqué aux indéterminés, tout en imposant des bornes, *PO-* n'affecte point le sémantisme du verbe, pas plus qu'il ne transforme la nature habituelle et durative de la situation.

s'actualiser. Tel est également le cas des verbes indéterminés qui apparaissent avec un seul actant sujet. *Jeździł* (il roulait) peut avoir aussi bien une signification particulière à valeur cursive que plurielle. Or, un complément de lieu, comme *do Lublina* (à Lublin) suffit pour imposer une interprétation plurielle à caractère habituel. Le recours à une expression adverbiale quantifiante n'est pas nécessaire.

3.3. PRZY-, DO- : deux cas de modification interne

Si, dans les deux classes examinées, *PO-* forme des dérivés perfectifs, il n'est pas possible d'en dire autant de *PRZY-* et de *DO-*. Observons les modifications opérées par ces préfixes sur une base déterminée telle que *jechać* (exemple 5a *versus* 5c et 5d) pour les comparer ensuite à celles que l'on repère dans (6c) et (6d) dans les dérivés formés par l'indéterminé *jeździć* de (6a).

(5) a. ... kierowca **jechał** do Lublina (nkjp.pl, texte de presse)

... le conducteur *JECHAĆ*_[passé masculin *imperfectif*] à Lublin

... *le conducteur allait à Lublin (à bord d'un moyen de transport terrestre)*

c. ... kierowca **przyjechał** do Lublina

... le conducteur *PRZYJECHAĆ*_[passé masculin *perfectif*] à Lublin

... *le conducteur est venu à Lublin (à bord d'un moyen de transport terrestre)*

d. ... kierowca **dojechał** do Lublina

... le conducteur *DOJECHAĆ*_[passé masculin *perfectif*] à Lublin

... *le conducteur a fini par arriver à Lublin (à bord d'un moyen de transport terrestre)*

Les préfixes *PRZY-* et *DO-* combinés aux bases déterminées forment des dérivés perfectifs, mais l'opération de perfectivisation se déroule autrement qu'en (5b) et (6b) dans la mesure où il ne s'agit pas de bornage. Les deux préfixes ont, dans un premier temps, une fonction sémantique : ils modifient les relations entre les participants de l'événement, en l'occurrence du déplacement. En comparant (5a) et (5c), on remarque que *przyjechać* relie *x* (*le conducteur*) et le terminus *y* spécifié par le complément *do Lublina*, en établissant entre eux une relation de localisation *x est à y*¹¹. À la différence de *PO-*, *PRZY-* opère de manière interne, en modifiant les relations entre les actants. En d'autres termes, le contenu relationnel de *przyjechać* n'est pas le même que celui de *jechać*. L'aspect lexical du dérivé se trouve également changé par rapport à celui du verbe-base. Le premier (*przyjechać*) dénote une situation dépourvue de durée, téléique, atomique. Si l'on suivait la

¹¹ Il convient de préciser que cette relation est établie par le prédicat même dans des phrases minimales de type *Przyjechał (Il est arrivé)*, dépourvues de spécification explicite de *y* sous forme de complément.

typologie de Vendler (1967)¹², *jechać* serait une activité [+ durée, - télicité], alors que son dérivé en *PRZY-* appartiendrait à la classe des accomplissements [- durée, + télicité]. La perfectivité des dérivés de forme *PRZY* + base déterminée est donc tributaire du sens lexical et, par conséquent, du type de situation désignée, non pas de bornage comme en cas de *PO-*.

De même, le préfixe *DO-* combiné aux verbes déterminés change, dans un premier temps, le contenu sémantique du prédicat pour aboutir à la modification du point de vue aspectuel. Dans (5d), l'adjonction de *DO-* à *jechać* permet de signifier la relation de localisation établie, une fois la partie finale de la trajectoire parcourue par *x* (*le conducteur*). Contrairement à *PRZY-*, *DO-* implique donc une certaine durée, fût-elle courte, nécessaire pour parachever le déplacement et atteindre son terme. Du point de vue de l'aspect lexical, les dérivés en *DO-* formés sur des bases déterminés sont à la fois duratifs et téliques. Le point de vue perfectif est, dans ce cas encore, fonction des rapports entre les actants, c'est-à-dire du contenu lexical du dérivé, et non pas d'une opération de bornage.

Lorsque le verbe-base est indéterminé, comme dans le cas de (6c), les dérivés formés à l'aide des préfixes à portée interne tels que *DO-* et *PRZY-*, restent imperfectifs.

(6) a. ... kierowca **jeździł** do Lublina

... le conducteur *JEŹDZIĆ*_[passé masculin imperfectif] à Lublin

... *le conducteur allait à Lublin (à bord d'un moyen de transport terrestre)*

c. ... kierowca **przyjeżdżał** do Lublina

... le conducteur *PRZYJEŹDŹAĆ*_[passé masculin imperfectif] à Lublin

... *le conducteur venait à Lublin (à bord d'un moyen de transport terrestre)*

d. ... kierowca **dojeżdżał** do Lublina

... le conducteur *DOJEŹDŹAĆ*_[passé masculin imperfectif] à Lublin

... *le conducteur finissait par arriver à Lublin (à bord d'un moyen de transport terrestre)*

Tout comme (6a), (6c) et (6d) sont des phrases habituelles. Dans ce paradigme dérivationnel, les préfixes assument uniquement leur fonction

¹² Nous faisons cette référence à Vendler à titre d'exemple pour illustrer les changements effectués. Avec toutes les réserves que nous avons émises (note 4) et la prudence qui s'impose dans l'application d'un classement établi pour une langue à une autre.

sémantique (Klemensiewicz 1951). Ainsi, *PRZY-* signifie la relation locative consécutive à chacun des déplacements, alors que *DO-* marque en plus le parcours de la partie finale de la trajectoire. Aucun des préfixes à fonctionnement interne combinés aux verbes indéterminés n'atteint le point de vue aspectuel : l'intervalle de référence pendant lequel l'itération des déplacements se produit n'a pas de limites temporelles. Le dérivé, ainsi que sa base, sont imperfectifs.

Il est intéressant de remarquer que contrairement aux dérivés français en *RE-*, ceux que l'on forge notamment à l'aide de *PO-*, *PRZY-* et *DO-* en polonais sont co-référentiels avec leurs bases. Ils permettent de renvoyer à un seul et même événement *e*, en en présentant les différentes facettes ou étapes. Dans la section suivante, nous nous efforcerons de dresser un bilan plus large de différences qui caractérisent le processus de dérivation par préfixation en français et en polonais.

4. Bilan : vers une description complète la plus courte

En analysant les effets produits sur les verbes de manière de mouvement par le préfixe *RE-* du côté français, et ceux que l'on reconnaît à *PO-*, *PRZY-* et *DO-* en polonais, nous avons admis qu'ils sont aisément comparables sur le plan morphologique. Envisagés sous cet angle, les dérivés sont plus complexes que leurs bases puisque leur description la plus courte est plus élaborée. En effet, les formalisations qui caractérisent les dérivés prennent les formes suivantes : *RE + V*, *PO + V*, *PRZY + V*, *DO + V*. Une même formule générale *PRÉF + V* est donc valable pour les deux langues en comparaison.

Au-delà de cette similitude morphologique, les préfixes que nous avons étudiés en français et en polonais diffèrent par leur portée. En français, *RE-* porte sur l'ensemble de l'événement désigné par la base, en introduisant des présupposés sur lesquels repose le sens du dérivé. Dans la mesure où *RE-* modifie ainsi le schéma temporel associé au verbe, le changement qui s'opère par rapport au verbe-base a un caractère aspectuel.

Du côté polonais, les préfixes examinés entraînent des modifications sur le plan aspectuel ou plutôt, devrions-nous dire, sur les plans aspectuels. Le bornage opéré par *PO-* affecte la trajectoire ou l'intervalle de référence, en produisant des formes perfectives. Dans le cas des déterminés, il modifie les relations entre les actants, c'est-à-dire le sens du verbe, ce qui conduit à modifier le schéma temporel (Vendler 1967) qui lui est associé. Dans le cas

des préfixes *PRZY-* et *DO-*, la fonction sémantique semble primer. Selon qu'il s'agit de verbes déterminés ou indéterminés, ils contribuent à créer une configuration nouvelle en ce qui concerne les rapports entre les actants. Le nouveau contenu sémantique du verbe peut conduire, comme dans le cas des déterminés, à un aspect lexical différent de celui de la base. Les dérivés en *PRZY-* sont des achèvements, ceux en *DO-* relèvent des accomplissements. De par la structure situationnelle, le point de vue aspectuel ne peut-être que perfectif. En revanche, la perfectivisation n'est pas systématique dans le cas de ces préfixes, comme le montre l'exemple des verbes indéterminés. Le tableau 1 réunit nos observations.

| | modification de relations entre les actants | modification d'aspect lexical | changement de point de vue aspectuel (imperfectif → perfectif) |
|---------------------------------------|---|----------------------------------|---|
| <i>PO + V_{déterminé}</i> | + | + | + |
| <i>PO + V_{indéterminé}</i> | - | - | + |
| <i>PRZY + V_{déterminé}</i> | + | + | + |
| <i>PRZY + V_{indéterminé}</i> | + | - | - |
| <i>DO + V_{déterminé}</i> | + | + | + |
| <i>DO + V_{indéterminé}</i> | + | - | - |

Tableau 1 : Portée de préfixes polonais (exemples de *PO-*, *PRZY-* et *DO-*)

Il nous semble possible dès lors d'avancer l'hypothèse selon laquelle les difficultés ressenties chez les apprenants tiennent aux types de modifications effectuées par les préfixes dans les deux langues. En comparant la complexité liée à leur action, on constate que la description la plus courte d'un dérivé français en *RE-* et celle des formes polonaises en *PO-*, *PRZY-* et *DO-* comprend des informations de nature différente. De plus, en polonais, le nombre de propriétés de la base modifiées dans le dérivé varie selon la classe verbale à laquelle le préfixe est appliqué. Par conséquent, calculer la valeur d'une forme *PRÉF + V* nécessite la prise en compte de nombreux paramètres et le coût cognitif de leur traitement s'avère élevé. Les étudiants choisissent donc les verbes polonais non préfixés, plus « sécurisants », car réduisant le risque d'erreur dans ce calcul, osons-nous dire, complexe.

Bibliographie

- Amiot, D. 2002. *RE-*, préfixe aspectual. *Cahiers Chronos*, 10, 1-20.
- Andersen, H. 2010. Complexité à l'oral et à l'écrit. Communication au colloque international *La complexité et son acquisition en langue*, Paris, 5-7 juillet 2010, Université Paris Descartes.
- Apothéloz, D. 2002. *La construction du lexique français*. Paris : Ophrys.
- Apothéloz, D. 2005. *RE-* et les différentes manifestations de l'itérativité. *Pratiques*, 125/126, 48-69.
- Apothéloz, D. 2007. La préfixation en *RE-*, l'antonymie directionnelle et les phénomènes de polarité aspectuelle. *Journal of French Language Studies*, 17, 143-158.
- Asher, N., Sablayrolles P. 1995. A Typology and Discourse Semantics for Motion Verbs and Spatial PPs in French. *Journal of Semantics*, 12, 163-209.
- Asnès, M., Kupferman, L. 2008. Événements, prédicats, arguments : quelques points de repère. *Langages*, 169, 7-33.
- Aunargue, M. 2008. *Qu'est-ce qu'un verbe de déplacement ? Critères spatiaux pour une classification des verbes de déplacement intransitifs du français*. Paris : Institut de Linguistique Française, CMLF'08.
- Bach, E. 1981. On Time, Tense and Aspect : An Essay on English Metaphysics. In P. Cole (éd.). *Radical Pragmatics*. New York : Academic Press, 62-81.
- Bach, E. 1986. The Algebra of Events. *Linguistics and Philosophy*, 9, 5-16.
- Boons, J.-P. 1985. Préliminaires à la classification des verbes locatifs : les compléments de lieu, leurs critères, leurs valeurs aspectuelles. *Linguisticae Investigationes* IX, 2, 195-267.
- Boons, J.-P. 1987. La notion sémantique de déplacement dans une classification syntaxique des verbes locatifs. *Langue française*, 76, 5-40.
- Borillo, A. 1987. Notions de « massif » et de « comptable » dans la mesure temporelle. In J. David, G. Kleiber (éds). *Termes massifs et termes comptables*. Paris : Klincksieck, 215-238.
- Borillo, A. 1998. *L'espace et son expression en français*. Paris : Éditions Ophrys.
- Caudal, P., Veters, C. 2007. Passé composé et passé simple : sémantique diachronique et formelle. *Cahiers Chronos*, 16, 121-151.
- Dahl, Ö. 2004. *The growth and maintenance of linguistic complexity*, *Studies in language companion series* / v. 71. Philadelphia : John Benjamins.
- Depraetere, I. 1995. On the Necessity of Distinguishing between (Un)Boundedness and (A)Telicity. *Linguistics and Philosophy*, 18, 1-19.
- Dowty, D.R. 1979, *Word Meaning and Montague Grammar*. Dordrecht : Reidel.
- Giermak-Zielińska, T. 1979. *Polskie czasowniki przedrostkowe o znaczeniu przestrzennym i ich odpowiedniki w języku francuskim*. Wrocław, Warszawa, Kraków, Gdańsk : Zakład Narodowy im. Ossolińskich, Wydawnictwo Polskiej Akademii Nauk.
- Gross, M. 1975. *Méthodes en syntaxe : régime des constructions complétives*. Paris : Harmattan.
- Gross, M. 1995. La notion de lieu argument du verbe. In H. Bat-Zeev Shyldkrot, L. Kupferman L. (éds). *Tendances récentes en linguistique française et générale. Linguisticae Investigationes, Supplementa* 20. Amsterdam/Philadelphia : John Benjamins Publishing Company.

- Grzegorzczkova, R. 1998. Nowe spojrzenie na kategorię aspektu z perspektywie semantyki kognitywnej. In R. Grzegorzczkova et Z. Zaron (éds). *Semantyczna struktura słownictwa i wypowiedzi*. Warszawa : Wydawnictwo UW, 25-38.
- Guillet, A., Leclère Ch. 1992. *La structure des phrases simples en français. Constructions transitives locatives*. Genève–Paris : Droz.
- Herslund, M. 1990. Les verbes inaccusatifs comme problème lexicographique. *Cahiers de lexicologie*, 56-57, 35-44.
- Kenny, A. 1963. *Action, Emotion and Will*. London : Routledge and Kegan Paul.
- Kleiber, G. 1987. *Du côté de la référence verbale. Les phrases habituelles*. Berne, Francfort/M., New York, Paris : Peter Lang.
- Klemensiewicz, Z. 1951. Próba charakterystyki dwu naczelných funkcji przedrostka w polskim czasowniku. *Sprawozdania Polskiej Akademii Umiejętności*, 52, 651-654.
- Kopecka, A. 2004. Etude typologique de l'expression de l'espace : localisation et déplacement en français et en polonais. Thèse de doctorat, Université Lumière Lyon 2.
- Kopecka, A. 2009. L'expression du déplacement en français : l'interaction des facteurs sémantiques, aspectuels et pragmatiques dans la construction du sens spatial. *Langages*, 173, 54-73.
- Laskowski, R. 1998. Kategorie morfologiczne języka polskiego – charakterystyka funkcjonalna. In R. Grzegorzczkova. *Gramatyka współczesnego języka polskiego. Morfologia*, 151-203.
- Mascherin, L. 2007. Analyse morphosémantique de l'aspectuo-temporalité en français. Le cas du préfixe *RE-*. Thèse de doctorat, Université Nancy 2/ATILF-CNRS (UMR 7118).
- Mourelatos, A. 1978. Events, Processes and Status. *Linguistics and Philosophy* 2, 3, 415-434.
- Mourelatos, A. 1981. Events, Processes and States. In P. Tedeschi, A. Zaenen (éds). *Syntax and Semantics. Tense and Aspect*, 14. New York: Academic Press, 191-212.
- Paduceva, E. V. 1996. *Semanticeskije issledovanija. Semantika vremeni i vida v ruskom jazyke. Semantika narrativa*. Moscou : Jazyki kul'tury.
- Przybylska, R. 2006. *Schematy wyobrażeniowe a semantika polskich prefiksów czasownikowych do-, od-, prze-, roz-, u-*. Kraków : Universitas.
- Sikora, D. 2009. *Les Verbes de manière de mouvement en polonais et en français. Eléments pour une étude comparée des propriétés structurelles des prédicats*. Thèse de doctorat, Nancy, Université Nancy 2 – ATILF – CNRS (UMR 7118).
- Smith, C. 1986. A Speaker-based Approach to Aspect. *Linguistics and Philosophy*, 9, 97-115.
- Smith, C. 1991. *The Parameter of Aspect*. Dordrecht, Boston, London : Kluwer.
- Stosic, D. 2009. La notion de « manière » dans la sémantique de l'espace. *Langages*, 175, 103-122.
- Vendler, Z. 1967. Verbs and Times. *Linguistics in Philosophy*, 97-121.
- Vet, C. 1980. *Temps, aspect et adverbess de temps en français contemporain*. Genève : Droz.
- Wilmet, M. 2007. *Grammaire critique du français*. Bruxelles, Paris : De Boeck Duculot.

Ressources informatisées :

Base textuelle Frantexte www.frantext.fr

Centre National de Ressources Textuelles et Lexicales www.cnrtl.fr

Narodowy Korpus Języka Polskiego www.nkjp.pl

Trésor de la Langue Française informatisé <http://atilf.atilf.fr/tlf.htm>

Is the Polish verb *ić* an auxiliary to be?

A corpus-based study of the construction *ić* + infinitive*

Dorota Sikora

ATILF-CNRS (UMR 7118), Nancy, France

The aim of the present paper is to investigate the degree of grammaticalization displayed by the Polish verb *ić* 'to go' in its infinitive constructions. A significant number of such VP show features usually associated with ongoing auxiliariation (ambiguity in context, desemantization, relaxed selection constraints and loss of grammatical properties of the verb). However, semantic bleaching is far from being complete as long as antonyms of *ić* are banned from infinitive complement position. Considering available corpus data, we assume that the grammaticalization of *ić* in its infinitive constructions has presently attained the stage of bridging contexts (Heine 2002).

Keywords: grammaticalization; auxiliariation; Polish motion verbs; semantic change

1. Introduction

The present paper deals with the infinitive constructions (henceforth IC) of the Polish verb *ić* 'to go', focusing on those that display clear signs of ongoing auxiliariation. In its basic sense, *ić* denotes the motion event of walking, performed under agentive control of the walker. However, as illustrated in the example (1), in some ICs, the lexical meaning seems replaced by another:

- (1) X *idzie więzienie*
X go:3sg be-in prison:acc
'X is going to be sent to prison' (Internet)

* I am very grateful to anonymous reviewers for the fruitful remarks to a previous version of this paper.

In the example (1), *iść* is by no means motion verb. As shown in Sikora (forthcoming), such an IC cannot be reduced to a simple goal expression.¹ It turns to assume a new role, roughly speaking that of future marker of *siadzieć* 'to go to prison': by uttering (1), the speaker indicates that the person X will be in prison before long. That is to say that the spatial meaning of *iść* has been replaced with a functional one. The question arising here is to what extent *iść* behaves as a grammatical marker and, possibly, an auxiliary.

1.1 Polish future tenses

Polish grammatical system offers two ways of expressing future temporal reference. The first one relies on a complex structure composed of the future form of the auxiliary *być* 'to be' followed (or preceded) by the infinitive or the past of the main verb. Such a future tense characterizes imperfective verbs as, for example, *czekać* 'to wait'. In order to say that one will be waiting, both – male and female Polish speakers – are free to state: *będę czekał* (BE:FUT.1SG+INF:INF), as well as they have possibility to mark gender using past form of the main verb: *będę czekał* (BE:FUT.1SG + INF:PST.M) or *będę czekała* (BE:FUT.1SG + INF:PST.F).

Perfective verbs, as *počkać* 'to wait (PF)', express future reference by using simple forms. Although from purely synchronic point of view, they could be considered present exponents, semantically they only convey future meaning. The non-actual present reading, currently expressed by such simple forms in some South Slavic languages (Bulgarian, Serbo-Croatian, Macedonian), is no more relevant for speakers of contemporary Polish, although it still may be identified in a fine-grained semantic analysis.²

1.2 Auxiliarization – a case of grammaticalization

Scholars generally agree to define grammaticalization as linguistic change consisting in a transfer of a lexical unit from lexicon to grammatical system of the language (cf. Heine et al. 1991). In a broader sense, grammaticalization designates the study of the variational process mentioned above.

1. As a matter of fact, unlike the subordinate clause *Idę do biblioteki, żeby się poszczyć* 'I am going to the library to study a bit', the IC under analysis does not allow separate temporal and spatial localization for the event described in the main and in the subordinate clause.

2. For an exhaustive discussion concerning Polish future forms, see Vanders & Sychotaka 1998.

In the present paper, the term *grammaticalization* is used in the first, restricted meaning. Consequently, I assume that auxiliaryization is one of possible ways of grammaticalization relevant for verb classes: auxiliaryized verbal lexeme progressively develops into auxiliary acquiring a new, partially or fully, grammatical content. Following that definition, grammaticalization (i.e. auxiliaryization) is a gradual diachronic unidirectional³ process with uncertain results. As noticed in Beinton & Traugott (2005:26): "... textual evidence suggests that many changes involve periods of relative indeterminacy in which it is not clear whether the older or the newer usage is in evidence (...) Indeed, the first steps may never result in change, in the sense of acceptance by a community of speakers".

Following Heine (2002), the graduality of grammaticalization process materializes in four stages. During the *initial phase* the verb fully preserves its lexical content in all uses. Some ambiguities appear in the second stage, that of *bridging contexts*, in which new inferences, open the way for a new significance. A *switch contexts phase* may then occur, when the evolving item starts an independent life in contexts incompatible with the source meaning. The final stage of *conventionalization* is characterized by a double life of the word: on the one hand, it remains part of the lexicon as content verb, on the other, it has another existence in grammar, as it happened to *have* in English and to *avoir* 'to have' in French.

2. How the motion verb *ić* is special

"The moment a verb is given an infinitive complement, that verb starts down the road of auxiliaryity" - says Bolinger (1980: 287). However, although the semantic chain MOTION > ASPECT > TIME seems largely applicable across languages,⁴ any single motion verb does not necessarily become a future marker.

Before we focused on *ić*, let us get a closer look at the whole set of Polish motion verbs. Forsyth's (1970) distinction between determinate and indetermi-

3. As shown in Prévost 2003, although unidirectionality is a definitional property of grammaticalization, it does not preclude reversibility of the process in some cases. See also Croft (2000, chapter 8).

4. Heine & Kuteva 2002 quote fourteen languages across the world in which the source verb of the future auxiliary is a motion verb. See also Marchello-Nizia 2006 (especially Chapter 3), Croft 1990.

nate Russian motion verbs holds in Polish. The Table 1 portrays semantic relations between the two series of items forming corresponding classes.

Table 1. Polish determinate and indeterminate motion verbs

| | Determinate verbs | Indeterminate verbs | Translation |
|----|-------------------|---------------------|--------------------|
| 1. | <i>ić</i> | <i>chodzić</i> | 'to go, to walk' |
| 2. | <i>jechać</i> | <i>jeździć</i> | 'to ride' |
| 3. | <i>biec</i> | <i>biegać</i> | 'to run' |
| 4. | <i>lecieć</i> | <i>latać</i> | 'to fly' |
| 5. | <i>frunąć</i> | <i>frunąć</i> | 'to fly (birds)' |
| 6. | <i>plywać</i> | <i>plywać</i> | 'to swim, to sail' |
| 7. | <i>pełzać</i> | <i>pełzać</i> | 'to crawl' |
| 8. | <i>łazić</i> | <i>łazić</i> | 'to climb' |

In Polish, the so-called indeterminate verbs (see Table 1) allow infinitive complements, but their constructions convey iterative meaning as in (2):

- (2) *Ala ma już swoją ulubioną restaurację, do*
 but have:1SG already own:ACC favourite:ACC restaurant:ACC to
której chodzi jej.
 which:GEN go:1SG eat:1SG

'But I already do have my favorite restaurant where I usually go to eat.'

(Internet)

As Traugott (2010:98) puts it: 'there must be compatibility between the earlier ("source") and the later ("target") meaning'. The question to be answered is whether the semantic components of *ić* encourage the evolution of this verb towards grammatical uses?

The issue is interesting because *ić*, as well as all Polish motion verbs (except for some stylistically marked and somewhat constrained items as *przemierzać się* 'to move'), are manner verbs. They differ then at that point from the English verb *to go* or the French *aller* 'to go', considered path verbs, without any specification of manner the motion is carried out.

Nevertheless, recent studies in space semantics have shown that Talmy's (2000) typological bipolarity is rather an idealized conceptualization. In fact, different languages appear as much or less close to one of the two edges formed by the Path lexicalization pattern on one side, and the Manner lexicalization pattern on the other (for details, see Levin, Beavers & Shiao Wei 2009).

The semantic specificity of Polish determinate verbs has been described in terms of linearity (Laskowski 1999) or directionality (Włodarczyk 1980). Given available data, I suggest that determinate verbs, including *iść*, display a very particular lexicalization pattern conflating manner and path (Sikora 2009). Such a semantic model is crucial for aspectual properties of determinate Polish motion and transport verbs as long as they denote an event progressing along a spatially ordered Path, possibly towards a goal. For further discussion of the semantics of motion verbs, see Tenney 1995. The goal complement position may be fulfilled with an infinitive. Concomitantly, the manner lexicalized by *iść* relies only on physical performances of agentive subject. These semantic components – progressive motion and a particular manner – imply a third one, namely intentionality in a double sense. The motion denoted by *iść* is intentional, first, because the walker is supposed to control his/her motion, second, by reason of the goal-oriented path. The subject intention is then to reach the goal. The link between intention and future action is clear enough: by uttering something like the example (3):

- (3) *Idę dzwonić na policję.*
 go.1SG call.INF to police.ACC
 'I am going to call the police.' (NKJP)⁵

The speaker signals the intention of calling.

Let us observe however that *iść* does not really need infinitive complement to express subject's intention. In the example (4), *iść* is used as pragmatic marker of intentionality with reference to the action announced in the right – future – context.

- (4) *Idę napisać smsa, będę czekał*
 go.1SG write.PRS write.ACC, be.PRS wait.PRS
na odpowiedź.
 for answer.ACC
 'Here I go, I will write a text message and I will wait for an answer.'
 (Internet)

In sum, these three semantic features of *iść* provide a perfect ground for the pragmatic procedure that Traugott & König (1991), Traugott (2000:101) refer to as invited inferencing: "Say no more than you must, and mean more thereby".

5. *Narodowy Korpus Języka Polskiego* (Polish National Corpus).

3. Evidence of ongoing auxiliarization

The example (1) was the first insight of the change undergone by *iść*. It is time now to turn toward corpus data searching for linguistic evidence of ongoing auxiliarization. The present section focuses on pragmatic ambiguity in context (3.1) and semantic bleaching (3.2).

3.1 Pragmatic ambiguity in context

Corpus data show a number of examples with two or three possible readings. Let us recall that, according to Heine 2002, such ambiguity is typical of bridging contexts. A standard request to Polish National Corpus (Narodowy Korpus Języka Polskiego, NKJP) brings up 49% of ICs with future reading, while in 23.3% of occurrences, *iść* may be thought of as well as modal marker of intentionality and temporal marker of future. In some rare sentences (3%), all three readings – spatial, modal and temporal – remain available, as illustrated in the example (5) below:

- (5) *Panie, co pan robi, idę zawolać policjanta.*
 Mister:voc what Mister:nom do:3sg go:1sg call:1sg policeman:acc.
 'Hey, Mister, what are you doing, I am going to call a police officer.' (NKJP)

3.2 Semantic bleaching (desemantization)

Scholars still lack agreement as to the nature of desemantization. Should we define it as semantic bleaching, i.e. as loss of lexical meaning of the source verb, or rather in terms of its generalization as claimed in Bybee (2003)? However, beyond current discussions, there seems to be an agreement on how desemantization of a lexical unit may come to light.

3.2.1 Releasing selection constraints

Used as content verb, i.e. as full lexeme referring to a motion event, *iść* selects dynamic infinitives at complement position. But more the auxiliarization process advances, wider the range of selected complements can be. As a matter of fact, corpus data clearly show that *iść* selects a varied set of infinitives by no means dynamic. The most frequent among them denote emotions (*czekać* 'to long for', *rozpaczać* 'to despair', *zapomnieć* 'to forget', etc.), intellectual processes (*myśleć* 'to think', *analizować* 'to analyse', *zastanawiać się* 'to wonder', *wątpić* 'to doubt', etc.), inactivity (*odpoczywać* 'to rest', *leżać* 'to lie'), postures (*leżeć* 'to lie'), body processes (*spać* 'to sleep', *trawić* 'to digest'). The examples (6) and (7) provide

instances of ICs in which weakened selection constraints do not block verbs incompatible with spatial meaning of *ić*.

- (6) Tak więc (i) *idę zapomniać o* osłabieniu i złym
 So then (i) *go:1sg forget:1sv* about feeling-weak:LOC and bad:LOC
stanowieniu i (ii) *idę zrobić coś dla siebie...*
 feeling:LOC and (ii) *go:1sg do:1sv* something:ACC for myself:GEN
 'So, I am going to forget feeling weak and bad, and I am going to do
 something for myself...' (Internet)

- (7) *Wziąłem trzy kawały kurczaka i jogurt*
devoured:1sv three:ACC big-piece:ACC chicken:GEN and bit-intens:ADV
ogurkiem. Idę trawić i myśleć (pozytywnie): "Jestem
yogurt:INSTR go:1sg digest:1sv and think:1sv (positive:adv) *be:1sg*
na diecie."
on dukan-diet:LOC
 'I devoured three big pieces of chicken and a yogurt. Now I am going to
 digest it and to think positive: "I am following the Dukan diet." (Internet)

Let us observe that the verb *trawić* 'to digest' provides particularly strong evidence of desemantization of *ić*. Such a complement precludes spatial reading, given the fact that digestion cannot be separated from human body and, consequently, conceptualized as a goal to reach.

As put forward in Abraham (1990), a grammaticalized item assumes what he refers to as a twin role in the language. On the one hand, it continues being used as full lexeme, and on the other, it assumes a function of grammatical marker. That is to say, the item lives some kind of double life, independently in lexicon and in grammar. If so, then the auxiliary and the lexeme may happen to appear together in the same construction, as shown in (8) and in (9):

- (8) *On va aller à Montparnasse, il a décidé.* (Frantext)
 (9) *I'm going to go and pick them up now.* (BNC)

It would be exaggerated to claim that the auxiliaryizing *ić* is free to select its own source lexeme at complement position. However, the corpus provides some striking cases, as those in the examples (10) and (11), revealing that the Polish verb has possibly started the way towards such a double life. Let us investigate on how far it went.

- (10) *Idę iść. Będę wczoraj.*
go:1sg go:1sv *be:1sv* *evening:1sv*
 'I am going to walk. I will be back tonight.' (Internet)

- (11) ...*idę myśleć idąc. Działanie. Nic mi tak nie*
go:3sg think:inf walk:gera literally nothing:nom that as-well was
poprawia humoru jak szwandanie.
improve:3sg humorous as wandering:nom
 'I am going to think walking. Literally. Nothing makes me feel better than
 wandering around.' (Internet)

ICs with *ić* used as inflected auxiliary and, subsequently, as infinitive complement, remain infrequent: as for now, two instances could be isolated.⁶

Although in the example (11), the complement position is fulfilled with *myśleć* 'to think', the whole clause brings a strong evidence to support the claim of desemantization undergone by *ić*. Obviously, using the IC *idę myśleć*, the speaker seems perfectly aware of the fact the inflected form *idę* is no more a content verb. S/he found necessary to add the gerund *idąc* in order to specify the circumstances of the thinking process.

Basically, a grammaticalized verb is also free to introduce infinitives antonymous of the source lexeme, as illustrated in the examples (12) and (13) where *stay* and *rest* express the opposite of *to go* and *aller* (see Vetter & Lière 2009).

- (12) *He is going to stay here for a while.* (BNC)
 (13) *Tout va rester à sa place, comme il est.* (Frantext)

Comparatively, in Polish, an IC with *ić* followed by *zostać* (PF) or *zostać* (IPF) 'to stay, to remain' is infelicitous, as shown in (14a) and (14b).

- (14) a. **Idziesz zostać tutaj?*
go:2sg stay:pf.inf here?
 'Are you going to stay here?'
 b. **Wszystko idzie zostać na miejscu.*
everything:nom go:3sg remain:pf.inf in place:loc
 'Everything will remain at the same place'

Data discussed in the present section suggest that *ić* cannot be yet considered full-fledged auxiliary. The bleaching process, though incomplete, has attained the stage in which some selection constraints are released: stative verbs as well as, in some rare cases, *ić* at the complement position are allowed. It is reasonable to assume then that in some uses (examples (10) and (11)), the verb is able to play the Abraham's twin role of both, functional marker and main verb. In short, despite clear signs of ongoing semantic bleaching, in most of its uses, *ić* seems at least able

6. The other one, extracted from Andrzej Szaniuk novel "Mary Hebron" appears as stylistically marked and could hardly be analyzed as illustration of an ongoing change occurring in discourse. It should rather be looked at as a writer's literary distortion of every day language.

to assume twin role, without living yet the double life of an auxiliary. We are still dealing with a lexical unit with grammatical uses.

3.2.3 *Loss of grammatical properties*

Weakening lexical constraints are not the only mark of auxiliarization. Further evidence of ongoing bleaching appears in comparison of grammatical properties displayed by *iść* in its different VPs. Let us recall some examples discussed in previous sections and compare them from the angle of their ability to form prefixed forms and their reactions to the change of grammatical tense. In (15), we are clearly dealing with a motion verb. In the case of (6), the spatial reading cannot be definitively excluded, but such interpretation is strongly constrained as long as we need a special context switching it on. On the contrary, in (11), *iść* shifts to functional meaning of future marker.

- (15) *Idę na rynek.*
 go:1SG to market:ACC
 'I am going to the farmers market.'
- (6) *Tak więc (i) idę zapomnieć o ciążności i złym
 na then (i) go:1SG forget:INF about feeling-weak:ACC and bad:ART
 samopoczuciu i (ii) idę zrobić coś dla siebie...
 feeling:ART and (ii) go:1SG do:INF something:ACC for myself:GEN
 'So, I am going to forget feeling weak and bad, and I am going to do
 something for myself...' (Internet)*
- (11) ... *Idę myśleć idąc. Dołownie. Nic mi tak nie
 go:1SG think:INF walk:GER literally nothing:NOM I.DAT so-well not
 poprawia humoru jak wędlenie.
 improve:1SG humor:GEN as wandering:NOM
 'I am going to think walking. Literally. Nothing makes me feel better than
 wandering around.' (Internet)*

The sentence (15) is slightly ambiguous as to the temporal reference with a possible reading in terms of current or intended displacement. However, used as main verb, *iść* is available for a large set of prefixes (*po-*, *przy-*, *za-*, *prze-*, *pod-*, *do-*), that can be applied to derive perfective forms (15'). In (6), only *po-* may be selected (6'), whereas in (11) even that prefix becomes problematic (11').

- (15') *po- Pójdę na rynek.*
 go:PERF.1SG to market:ACC
 'I am about to go to the farmers market.'
- przy- Przyjdę na rynek.*
 come:PERF.1SG to market:ACC
 'I will come to the farmers market.'

- ze- Zaidę na rynek.
come.pr.1SG to market.ACC
'I will make a little trip to the farmers market.'
- prze- Przejdę na rynek.
move.pr.1SG to market.ACC
'I will move (from here) to the farmers market.'
- pod- Podojdę na rynek.
go.pr.1SG to market.ACC
'I will go to the farmers market (once on my way).'
- do- Dojdę na rynek.
join.pr.1SG to market.ACC
'I will join the farmers market.'

- (6') Tak więc (i) pójdę zapomnieć o osłabieniu i
so then (i) go.pr.1SG forget.INF about feeling-weak.ABS and
złym samopoczuciu i (ii) pójdę zrobić coś
bad.ABS feeling.ABS and (ii) go.pr.1SG do.INF something.ACC
dla siebie...
for myself.GEN
'So, I will forget feeling weak and bad, and I will do something for myself...'
- (11') Pójdę myśleć idąc. Dosłownie. Nic mi
go.pr.1SG think.INF walk.GEN literally nothing.SOME I.ABS
tak mi poprawia humoru jak szwędanie.
so-well I.ABS improve.INF humor.GEN as wandering.WOM
'I am going to think walking. Literally. Nothing makes me feel better than wandering around.'

To summarize, the above results, we can conclude that more the process has advanced, less prefixes remain available for derivation. We could even go farther, assuming that in the case of strongly grammaticalized use (see example (11')), such derivation is not possible at all, given the fact that (11') sounds oddly to Polish speakers.

Another way of checking whether a lexical item tends to auxiliarize is to use the verb in another grammatical tense than present. The reaction displayed by the verb is similar in all sentences under analysis: used at past tense, *ić* recovers its initial spatial meaning. The examples (15'), (6') and (11') refer to a motion event only.

- (15') Szedłem na rynek.
went.3PL.M to market.ACC
'I was walking to the farmers market.'

- (8⁷) Tak więc (i) *szedłem zapomniałem o osłabieniu*
 so then (i) went:IPF.M forgot:INF about feeling-weak:ABL
 i *złym samopoczuciu i (ii) szedłem*
 and bad:ABL feeling:ABL and (ii) went:IPF.M
zrobić coś dla siebie...
 do:PT:INF something:ACC for myself:GEN
 'So, I was walking to forget feeling weak and bad, and I was walking to do
 something for myself'
- (11⁷) '*...szedłem myśleć idąc. Dosłownie.*
 went:IPF.M think:INF walk:GER literally
 'I was going to think walking. Literally.'

It is worthy to notice that the change of grammatical tense in (11⁷) has just transferred the event into the past, while in (8⁷) and in (11⁷), the informational content of the sentences has been attained. In other words, (8) and (11) on the one hand, (8⁷) and (11⁷) on the other provide different kind of information.

To sum up, the verb *ić* displays different grammatical behavior when used with functional sense or with lexical meaning.

4. Semantics of the auxiliarized *ić*

In the present state of Polish corpora, a reliable statistic study about all aspects of *ić* in ICs is hard to realize. As a matter of fact, although users have access to collocations, extracting instances of infinitive complements only is not possible yet. Otherwise, only collocations with minimally five occurrences in the corpus are provided in answer to such a query.

Nevertheless, random outputs obtained in such a corpus study bring to light some regularity and encourage few explanatory generalizations. First, let us notice that corpora offer several examples of grammaticalized ICs inflected at different persons, as illustrated in (16):

- (16) *Nowi klienci, którzy teraz idą wznawiać*
 new:NOM.PL clients:NOM who:NOM.M now go:3SG renew:IPF:INF
ładź nową lub ubezpiecznić przynajmniej raz
 or get:PT:INF insurance:PL:ACC bring:3SG with REFL:INF
świeżą gotówką.
 fresh:ACC money:ACC
 'New clients, who at present are going to renew their insurance or get a new
 one, are bringing in fresh money.' (NKJF)

However, the majority of the auxiliarized occurrences are inflected at first person singular. That very fact matches with the invited inferencing discussed in the Section 3. As noticed by De Mulder (p. c.), the first person form is the most appropriate for expression of intentions. No one but the doer can better talk about what s/he wants to, and probably will, do.

Second, it is interesting to get a quick look on the contexts in which the IC under analysis appears. Some of them, as the one in the example (17), provide striking evidence of the functional role that the auxiliarized *ić* assumes in today's Polish.

- (17) *Pozdrawiam i idę rozmrozić kawał mięsa, potem*
 greet.1SG and go.1SG defrost.FUT.1SG big-piece.ACC meat.GEN then
je przysmę.
 it.ACC season.FUT.1SG
 'I am sending my regards, and I am going to defrost a big piece of meat,
 then I will season it.' (Internet)

Obviously, the IC *idę rozmrozić* conveys a future meaning. However, as shown in (17'), the clause cannot be replaced with perfective future of the infinitive complement *rozmrozić* 'to defrost', unless the sentence loses coherence: the present of *pozdrawiam* '(I) (am) greet(ing)', which is a performative act of greeting the addressee, and the future cooking activities would appear then as logically disconnected.

- (17') **Pozdrawiam i idę rozmrozić kawał mięsa, potem*
 greet.1SG and go.1SG defrost.FUT.1SG big-piece.ACC meat.GEN then
je przysmę.
 it.ACC season.FUT.1SG

The analysis of the example (17) points the function assumed by auxiliarized *ić* in ICs: it provides the speaker with a mean to tie up the present state and future actions. That is to say that the auxiliarized *ić* is currently relevant in its role of future marker as long as the expressed intension of doing something is the first condition – partly satisfied at the moment of speech – of the act referred to by the infinitive complement.

5. Conclusions

It is time now to sum up the above considerations and to evaluate the degree of auxiliarization displayed by *ić* in its ICs.

Given available data, the auxiliarization of the verb *ić* in ICs appears as still ongoing process, far from being completed. The future meaning remains often contextually constrained and the selectional restrictions due to the source meaning are not fully released. The pragmatic ambiguity of the large number of examples under analysis indicates that the auxiliarization process has reached the bridging context phase (Heine 2002): the one in which such ambiguity is not resolved, given the fact that none of the three possible readings (spatial, modal and temporal) is defeasible in context. However, some examples point towards the third stage, that of switch contexts, where *ić* appears in environments incompatible with its original spatial meaning. That is to say, in such contexts, only the "new" grammatical meaning (i.e. that of future marker) is involved. As shown in recent works on grammaticalization, such evolution may continue as well as it may never result in the next stage.

The path, or pattern, of grammaticalization and, to be precise, of auxiliarization followed by *ić* is, of course, by no means language-specific. We know – since Bybee, Perkins & Pagliuca 1994, Heine & Kuteva 2002 among others – that this is one of the universal conceptual channels of grammaticalization. The main aim of this paper was to provide evidence of similar conceptual device at work in the case of the IC under analysis. Unfortunately, given the state of development of Polish corpora, it is impossible to predict whether *ić* will continue the way of auxiliarization. Hardly 10% of texts gathered in the IPT Corpus of Polish are labeled "older prose", but these texts come from nineteenth century novels. Consequently, it is impossible to build diachronic sub-corpora representative for different periods of Polish language history. In other words, we lack data to run statistics and identify the tendency in the auxiliarization undergone by *ić*.

However, some insights are provided from the development of perfect present form in today's Polish. Sentences like (18), considered inappropriate thirty years ago, are currently used and understood as conveying particular aspectual and temporal information, different from that provided by the corresponding past and present forms.

- (18) *Dom ma zrobione zewnętrzne ocieplenie.*
 house:nom have:3sg done:part external:acc insulation:acc
 'The house has external insulation.' (Internet)

As stated (among others) in Paprocka-Piotrowska (2008), Polish tenses are poor in modal and aspectual information. Neither of the tenses described in handbooks of grammar can refer to a result state as the "present perfect" in (18) does. I suggest then that the auxiliarized verb *ić* with its strong modal, aspectual and temporal charge enables speakers to fill one of such gaps in Polish temporal system, saying no more than necessary, and meaning more thereby.

Abbreviations

| | | | | | |
|------|---|--------------|-----|---|------------|
| ACC | = | accusative | NON | = | nominative |
| ADV | = | adverb | PT | = | past |
| FUT | = | future | PR | = | perfective |
| GEN | = | genitive | PRS | = | present |
| GRD | = | ground | PL | = | plural |
| INS | = | instrumental | SG | = | singular |
| INST | = | instrumental | VOC | = | vocative |
| IMP | = | imperfective | | | |

References

- Abraham, Werner. 1990. Die Grammatikalisierung von Auxiliär- und Modalverben. *Beiträge zur Geschichte Sprache und Literatur* 112: 200–208.
- Bellinger, Dwight. 1980. *Wanna* and the gradience of auxiliaries. In *Ways to Universalienforschung: Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Sailer*, Gunter Brutschneider & Christian Lehmann (eds), 282–298. Tübingen: Gunter Narr.
- Brinton, Laurel J. & Traugott, Elizabeth C. 2005. *Lexicalisation and Language Change*. Cambridge: CUP.
- Bybee, Joan L. 2003. Mechanisms of change in grammaticalization: The role of frequency. In *The Handbook of Historical Linguistics*, Brian D. Joseph & Richard D. Janda, 602–623. Oxford: Blackwell.
- Bybee, Joan L., Perkins, Revere D. & Pagliuca, William. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago IL: Chicago University Press.
- Croft, William. 1990. *Typology and Universals*. Cambridge: CUP.
- Croft, William. 2000. *Explaining Linguistic Change: An Evolutionary Approach*. London: Longman.
- Foersth, James. 1970. *A Grammar of Aspect: Usage and Meaning in the Russian Verb*. Cambridge: CUP.
- Helme, Bernd. 2002. On the role of context in grammaticalization. In *New Reflections on Grammaticalization* [Typological Studies in Language 49], Elie Wücher & Gabriela Diewald (eds), 83–104. Amsterdam: John Benjamins.
- Helme, Bernd, Claud, Ulrike & Hünemeyer, Friederike. 1991. *Grammaticalization: A Conceptual Framework*. Chicago IL: Chicago University Press.
- Helme, Bernd & Kuteva, Tania. 2002. *World Lexicon of Grammaticalization*. Cambridge: CUP.
- Łaskowski, Roman. 1998. Czasowniki ruchu liniowego (czasowniki przemieszczania) w perspektywie kontrastywnej. *Semantyka a konfrontacja języków* 2: 49–57.
- Levin, Beth, Beavers, John & Shiao Wei, Tham. 2009. Manner of motion roots across languages: Same or different? In *Roots: Word formation from the perspective of "low lexical elements"* workshop, University of Stuttgart, 10–12 June.
- Maschello-Niina, Christiane. 2006. *Grammaticalization et changement linguistique*. Bruxelles: De Boeck.

- Paprocka-Potrowska, Urszula. 2008. *Contre au risque de tout changer. Complexité conceptuelle et référence aux procès dans l'acquisition du français L2 et du polonais L2*. Lublin: Towarzystwo Naukowe Katolickiego Lubelskiego.
- Prévost, Sophie. 2003. La grammaticalisation: Unidirectionnalité et statut. *Le Français moderne* LXXXI(2): 144-166.
- Sikora, Dorota. 2009. Les verbes de manière de mouvement en français et en polonais. Éléments pour une étude comparée des propriétés structurales des prédicats. Thèse de doctorat sous la direction de D. Apothélos, ATELJ-CNRS (UMR 7118)-Nancy-Université.
- Sikora, Dorota. Forthcoming. Sur le chemin d'auxiliarisation? Verbe polonais *tu* dans ses constructions infinitives – étude sur corpus. *Recoviaki Humanistyczne*, Université Catholique de Lublin.
- Talmy, Leonard. 2000. *Toward a Cognitive Semantics*, Vol. 1: Concept Structuring System, Vol. II: *Typology and Process in Concept Structuring*. Cambridge MA: The MIT Press.
- Toussy, Carol. 1995. How motion verbs are special. The interaction of semantic and pragmatic information in aspectual verb meaning. *Pragmatics and Cognition* 3(1): 31-73.
- Traugott, Elizabeth C. 2010. Grammaticalization. In *Historical Pragmatics*, Andreas H. Jucker & Irma Taavitsainen (eds), 97-126. Berlin: Mouton de Gruyter.
- Traugott, Elizabeth C. & König, Ekkehard. 1991. The semantics/pragmatics of grammaticalization revisited. In *Approaches to Grammaticalization*, Vol. 1 [Typological Studies in Language 19], Elizabeth C. Traugott & Bernd Heine (eds), 189-218. Amsterdam: John Benjamins.
- Vetters, Carl & Litze, Audrey. 2009. Quand une périphrase devient temps verbal: Le cas d'*aller* + infinitif. *Faits de langues* 35: 27-36.
- Vetters, Carl & Skibińska, Elżbieta. 1998. Le futur: Une question de temps ou de mode? Remarques générales du 'présent-futur' perfectif polonais. In *Cahiers Chronos 2*, André Bocillo, Carl Vetters & Marcell Vuillaume (eds), 247-266. Amsterdam: Rodopi.
- Włodarczyk, Hilaria. 1988. Czasowniki ruchu a kategorie aspektu w języku polskim. *Polonica* 6: 49-56.

On-line corpora:

Frantext (www.frantext.fr)

British National Corpus (www.natcorp.ox.ac.uk)

Korpus Języka Polskiego Informacji Polskiej Akademii Nauk (IPI PAN) (<http://korpus.pl>)

NKIP: Narodowy Korpus Języka Polskiego (Polish National Corpus) (www.alkp.pl)

**Quand le parfait est une affaire de prédication
seconde : analyse de la construction polonaise
*N₁ mieć participe passé N₂***

Dorota SIKORA

Université du Littoral-Côte d'Opale

Abstract

The purpose of the present study is to investigate the linguistic status of the construction *N₁ + mieć (have) + past participle + N₂* in contemporary Polish. Until recently, such a construction was mainly found in oral discourse. The central position of the verb *mieć* and the apparently perfect meaning marked by the construction as a whole prompts a comparison with the French *passé composé* tense. Indeed, various authors (e.g. Abraham, 2013) suggest an ongoing auxiliarisation of *mieć* and an interpretation of the construction as an emerging grammatical tense. However a closer analysis of its syntactic properties as well as the semantics of the verb do not make it possible to consider *mieć* as a perfect auxiliary. Rather than a Polish version of a *passé composé*, the suggestion is made that it is a particular form of secondary predication projecting the beneficiary argument into subject-theme position.

Keywords: Polish, perfect, auxiliarisation, secondary predication.

1. Présentation : le parfait polonais en questions¹

Il a été souvent dit (Paprocka-Piotrowska, 2008, Apothéoz et Nowakowska, 2010) que le système temporel du polonais est relativement pauvre en ce qui concerne aussi bien les possibilités de localiser les situations sur l'axe du temps que pour l'expression des catégories aspectuelles. En effet, les quatre

¹ Je tiens à remercier chaleureusement Marc Fryd et Pierre-Don Giancarli, ainsi que les relecteurs anonymes, ainsi que pour leurs commentaires. Leurs suggestions m'ont permis d'apporter de nombreuses améliorations à une première version de ce texte.

tiroirs (le passé, le présent, le présent-futur et le futur²) permettent simplement de situer un fait avant, après ou simultanément à un moment *t* de référence, le plus souvent le moment d'énonciation. Par ailleurs, les temps grammaticaux apportent peu, voire pas du tout, d'informations susceptibles de compter dans le calcul de l'aspect phrastique.

Dans ce contexte, il est d'autant plus intéressant d'observer l'émergence d'une construction à valeur de parfait (désormais CVP), construite sur le verbe *mieć* (avoir). L'exemple (1) en fournit l'illustration³ :

- (1) Sasiadka ma przygotowany
 sasiadka.Nom.sg mieć.Prés.3sg przygotowanyPART.PS.Acc.sg masc.
 obiad.
 obiad.Acc.sg masc.
 voisine-a-préparé déjeuner (*Le déjeuner de la voisine est prêt*)

Longtemps pourchassée comme un emprunt sauvage aux langues germaniques que l'on devait remplacer par le verbe qui a servi à construire le participe passé, elle a fini par s'imposer ces dernières années non seulement dans le registre oral, mais également – de plus en plus souvent – à l'écrit⁴.

Les travaux qui s'intéressent à cette construction (Muryn, 2009, Apothéloz et Nowakowska, 2010) mettent en avant une forte ressemblance de notre CVP d'une part avec le passé composé français à une certaine phase de son développement historique, et de l'autre avec des phrases telles que *Paul a les lunettes cassées*.

² Si la valeur temporelle du futur analytique des verbes imperfectifs, construit avec le l'auxiliaire *być* (être), n'est pas mise en doute, celle de la forme synthétique des perfectifs est parfois interprétée comme véhiculant un présent non actuel (i.e. n'englobant pas le moment d'énonciation, cf. Laskowski, 1984). Vetter et Skibińska (1998) montrent qu'il s'agit bien d'un futur, parfois de valeur modale. La distinction entre le futur (analytique) des imperfectifs et le présent-futur (synthétique) des perfectifs se fait ainsi sur des bases à la fois morphologique et sémantique.

³ La spécificité des CVP n'est pas sans poser des problèmes de traductions. L'emploi du passé composé en français ne garantit pas systématiquement la préservation de la valeur du parfait, alors qu'une construction purement résultative ne permet pas de montrer la place qu'y occupe le verbe *mieć* (avoir). Les traductions que nous proposons à la suite des gloses sont ainsi un compromis visant à rendre les aspects structurel et sémantique : la première séquence, avec les tirets, est destinée à caractériser la forme décrite en détail dans la glose, alors qu'entre parenthèses nous proposons une traduction moins fidèle, mais acceptable. Nous remercions les relecteurs anonymes pour cette suggestion.

⁴ Selon J. Bralczyk (c.p.), cette construction a aujourd'hui pris toute sa place dans la langue et ne saurait être critiquée comme impropre. Si elle suscite l'intérêt des linguistes, la plupart des grammaires du polonais continuent à l'ignorer.

L'objectif de cette étude est de déterminer la nature et le statut des CVP. Nous tenons en effet à réunir les arguments qui permettent de décider si elles relèvent de la flexion verbale ou de la syntaxe. Pour ce faire, nous explorons le degré de ressemblance entre la CVP polonaise et les deux structures du français. Nous espérons faire apparaître, à travers des analyses contrastives, les particularités de la construction polonaise. Nous chercherons ainsi à répondre aux trois questions suivantes :

1. Existe-t-il une ressemblance réelle entre une CVP et une phrase correspondante au PC en français, notamment en ce qui concerne le degré d'auxiliarisation du verbe *mieć* et la fonction qu'il assume ?

2. Quel est le degré de ressemblance entre une CVP et les structures bipropositionnelles françaises telles que *Paul a les lunettes cassées*, qui ont été décrites par Furukawa (1996) comme cas de prédication seconde, sous le nom de phrases à double thème ?

3. Quel est le statut grammatical des CVP en polonais actuel ?

Nous procéderons en quatre étapes. Dans la section 2, nous présentons la définition du parfait sur laquelle nous appuyons notre réflexion, ainsi que quelques faits de langue qui nous aideront à situer nos CVP dans un contexte plus large en polonais actuel. La section 3 est consacrée à l'analyse des structures syntaxiques des CVP, des phrases correspondantes au PC en français et des constructions bipropositionnelles. La section 4 s'intéresse aux éventuelles ressemblances et aux réelles différences entre les CVP en polonais et le PC français, notamment en ce qui concerne l'auxiliarisation des verbes *mieć* et *avoir*. La partie suivante compare les CVP et les phrases bipropositionnelles françaises. Les éléments ainsi dégagés devraient nous permettre, dans la section 6, de conclure nos investigations : les CVP polonaises présentent plusieurs caractéristiques des phrases bi-thématiques françaises. Il convient de les traiter donc comme un cas de double prédication, plutôt que comme un temps grammatical ou comme une périphrase aspectuelle.

Les exemples qui illustrent nos propos sont inspirés des énoncés attestés dans *Narodowy Korpus Języka Polskiego*⁵ (Corpus national du polonais, désormais *nkjp*) et sur le Web. Il s'agit en fait de versions simplifiées, expurgées d'éléments susceptibles d'alourdir les gloses. Compte tenu de notre objectif dans cet article, nous nous servirons de CVP construites sur le présent du verbe *mieć* (*avoir*), ne faisant qu'invoquer, dans la section 4.3.1, la référence temporelle des CVP contenant ses formes passées et futures.

⁵ www.nkjp.pl

2. Définition du parfait et quelques faits de langue

2.1 Définition du parfait

Comrie (1976 : 52) remarque le caractère spécifique du parfait par rapport aux « autres aspects » : « ...it tells us nothing directly about the situation itself, but rather relates some state to a preceding situation » en ce que « it expresses a relation between two time-points, on the one hand the time of the state resulting from a prior situation, and on the other, the time of that prior situation ».

Le caractère spécifique du parfait réside donc dans le fait que, contrairement aux autres aspects, il ne s'agit pas de la structuration interne d'une situation, mais d'une relation entre deux temporalités : une situation antérieure et un état résultant qui la suit sur l'axe du temps. Décrit ainsi, le parfait apparaît comme une catégorie sémantique.

Les définitions du parfait en tant que catégorie sémantique sont plutôt consensuelles dans la communauté des linguistes aspectologues. Elles mettent systématiquement en jeu deux éléments communs aux différentes valeurs du parfait : un état résultant *r* et une situation *e* qui lui est antérieure. La question des relations temporelles entre *r* et *e* est plus délicate en revanche. De Swart (1998) considère que la contiguïté temporelle, sur laquelle repose la métonymie entre *e* et *r*, impose une contrainte de perfectivité : une borne droite de *e* est indispensable pour marquer le point de passage entre les deux situations *e* et *r*. C'est d'ailleurs cette caractéristique qui a permis, dans plusieurs langues (cf. Bybee, Perkins et Pagliuca, 1994, Heine et Kuteva, 2002), la transformation du parfait en perfectif.

Cependant, les résultats des analyses des données multilingues, discutés dans Dahl (2000), montrent que si le parfait est largement utilisé dans de nombreuses langues, il s'agit d'un phénomène plutôt instable compte tenu de la diversité de ses manifestations d'une part, et de l'autre, de sa tendance à évoluer dans d'autres formes. Dans certaines langues, le parfait possède une structure formelle par laquelle il s'exprime, dans d'autres, cette valeur aspectuelle est véhiculée par d'autres moyens, par exemple par des inférences pragmatiques.

De même, la contrainte de perfectivité peut se révéler d'une pertinence variable. Caudal et Veters (2007 : 122) précisent que la situation *e* doit être « au moins en partie temporellement antérieure » à l'état résultant. Ainsi, une vue globale de *e* n'est pas une condition *sine qua non* pour qu'il y ait une valeur de parfait, notamment lorsqu'il s'agit d'un parfait d'expérience (cf. entre autres, Apothéloz et Nowakowska, 2010).

2.2 Construction à valeur de parfait en polonais actuel

Les exemples que nous présentons ci-dessous ont tous la même forme

syntactique N₁ mieć (avoir) PartPs N₂, sans pour autant véhiculer à chaque fois une valeur de parfait.

- (1) Sąsiadka ma przygotowany
 Sąsiadka.Nom.sg mieć.Prés.3sg przygotowany.PartPs.Acc.sg masc.
 obiad.
 obiad.Acc.sg masc.
 voisine-a-préparé déjeuner (*Le déjeuner de la voisine est prêt*)
- (2) Pies ma zrobione
 Pies.Nom.sg mieć.Prés.3sg zrobiony.PartPs.Acc.pl neutre
 szczepienia.
 szczepienie.Acc.pl neutre
 chien-a-faites vaccinations (*Le chien a toutes les vaccinations*)
- (3) Sąsiadka ma zrobioną.
 Sąsiadka.Nom.sg mieć.Prés.3sg zrobiony.PartPs.Acc.sg fém.
 trwałą
 trwała.Acc.sg.fém.
 voisine-a-fait permanente. (*La voisine a une permanente*)
- (4) Pacjent ma robioną
 Pacjent.Nom.sg mieć.Prés.3sg robiony.PartPs.Acc.sg fém
 tomografię.
 tomografia.Acc.sg fém.
 patient-a-fait scanner (*Le patient est en train de passer un scanner*)
- (5) Pacjent ma robioną
 Pacjent.Nom.sg mieć.Prés.3sg robiony.PartPs.Acc.sg fém
 tomografię raz na miesiąc.
 tomografia.Acc.sg fém raz na miesiąc.CIRCONSTANT
 patient-a-fait scanner une fois par mois (*Le patient passe un scanner une fois par mois*)

Les exemples (1) à (3) ont une valeur de parfait. Leurs PartPs sont formés sur des verbes perfectifs *przygotować* (*préparer*) dans (1) et *zrobić* (*faire*) dans (2) et (3). Dans les phrases (4) et (5), le PartPs est dérivé du verbe imperfectif *robić* (*faire*)⁶, et le sens de parfait est absent. La phrase (4) signifie que le patient est en train de subir un scanner, alors que l'adjonction d'un circonstant fréquentatif *raz na miesiąc* (*une fois par mois*) conduit à une lecture itérative de l'exemple (5).

⁶ Les verbes *robić* et *zrobić* signifient tous les deux *faire*, mais ils lexicalisent un point de vue aspectuel différent : imperfectif dans le cas de *robić*, perfectif dans le cas de *zrobić*. Le point de vue imperfectif n'empêche pas *robić* de former un prédicat télélique, comme le montrent nos exemples (4) et (5), qui sont des accomplissements. L'effet de l'aspect imperfectif est alors comparable à celui qu'exerce l'imparfait sur des prédicats téléliques.

Remarquons cependant que la perfectivité du verbe source du PartPs n'est pas une condition suffisante pour qu'une valeur de parfait apparaisse, ni même pour que notre construction soit acceptable. En effet, il est impossible de transformer l'exemple (6), construit sur le verbe perfectif *poczytał* (*lire pendant une courte période*) en une CVP : la phrase (7) est simplement agrammaticale.

- (6) Uczeń poczytał książkę.
 Uczeń.Nom.sg poczytał.Passé perf. masc książka.Acc.sg fém.
 L'élève a lu le livre (pendant un moment).
- (7) *Uczeń ma poczytany
 *Uczeń.Nom.sg mieć.Prés.3sg poczytany.PartPs.Acc.sg fém
 książkę.
 książka.Acc.sg fém.
 *élève-a lu-livre

Si (7) n'est pas possible, c'est parce que le préfixe perfectivisant *po-* opérant sur le verbe imperfectif *czytał* (*lire*) assume une fonction de délimitation temporelle. Il introduit des bornes temporelles, initiale et finale, sans que cette dernière soit télélique⁷. Il en ressort que la capacité qu'a le verbe-base du PartPs de former un prédicat télélique est décisive pour obtenir une CVP⁸.

Il est possible dès lors d'établir un état de faits et de situer les CVP dans un contexte plus large :

1. Il existe, en polonais contemporain, une construction syntaxique dont la forme peut être résumée par la formule : N_{Nom} mieć (*avoir*) PartPs_{Acc} N_{Acc} .

2. Du point de vue sémantique, cette construction peut véhiculer une valeur de parfait, lorsque le PartPs est formé sur un prédicat télélique exprimé par un verbe perfectif.

Cela signifie que les CVP sont des réalisations particulières d'un patron syntaxique plus large. Leur valeur de parfait tient aux propriétés

⁷ En suivant Depraetere (1995), nous distinguons la délimitation et la télélicité, en admettant que toute borne n'est pas télélique. En effet, la délimitation d'une séquence situationnelle peut s'opérer par bornage, sans qu'il y ait télélicité, c'est-à-dire sans que cette séquence se déroule de manière téléonomique et sans viser un changement quelconque susceptible d'affecter l'un des arguments. Pour une discussion plus détaillée de cette distinction appliquée à certains phénomènes linguistiques en français et en polonais, voir respectivement Sikora (2012a) et Sikora (2012b).

⁸ La valeur résultative susceptible d'apparaître auprès d'un verbe télélique n'est guère étonnante : le lien entre la télélicité et la résultativité a souvent été mis en avant (cf. à titre d'exemple, Creissels, 2000). Apothéloz et Nowakowska (2010) le soulignent également dans le contexte de la construction qui nous occupe ici.

aspectuelles ([+ perfectivité], [+ télicité]) du verbe dont le PartPs figure dans la construction.

3. Réseaux de dépendances syntaxiques dans une perspective comparée

Dans la présente section, nous analysons la structure syntaxique des CVP. Cette étape est nécessaire pour identifier la fonction qu'y assume le verbe *mieć* (avoir). Elle nous fournira des éléments de réponse à la question de savoir s'il peut être qualifié d'auxiliaire comme l'est *avoir* au passé composé en français.

3.1 Structure syntaxique d'une CVP polonaise

Reprenons l'exemple (1) présenté dans la section précédente :

- (1) Sąsiadka ma przygotowany obiad.
voisine-a-préparé déjeuner (Le déjeuner de la voisine est prêt)

Le réseau de dépendances syntaxiques dans une phrase telle que (1) se présente de la manière suivante :

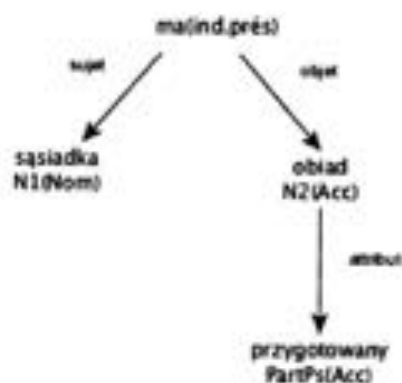


Figure 1 : Réseau de dépendances syntaxique dans une CVP

Dans une CVP telle que (1), le PartPs s'accorde en genre et en nombre avec N_2 , qui est son gouverneur syntaxique. Il y a donc entre ces deux éléments une dépendance de nature attributive et ils forment un syntagme nominal SN_{N_2Acc} , qui assume la fonction syntaxique de complément d'objet direct du verbe *mieć* (avoir). Il est dès lors possible de représenter notre construction sous une forme contractée SN_1 *mieć* (avoir) SN_2 . Comparons à

présent ce schéma de dépendances avec celui des phrases au passé composé en français.

3.2 Structure syntaxique des phrases au passé composé en français

Dans l'exemple (8), le passé composé du verbe *préparer* appelle une lecture en termes de parfait.

- (8) Je me lève, Agnese a préparé le café comme d'habitude, je le chauffe et le bois en regardant le puits blanc sur la place, en bas, bain, journaux sur les Zattere, Nuova strage in Parigi... (Frantext, Sollers Ph., *Le Cour absolu*, 1987, p. 410)

Pour plus de transparence, nous extrayons de l'exemple (8) la séquence *Agnese a préparé le café* pour observer le réseau de dépendances qui se construit à l'intérieur d'une phrase au passé composé où *avoir* a clairement un statut d'auxiliaire.

L'analyse syntaxique d'une phrase à auxiliaire n'est pas sans poser quelques problèmes de méthode. En effet, la relation entre celui-ci et son verbe principal peut être envisagée selon trois points de vue. Une première approche (Crystal, 1980, Huddleston, 1984), consiste à admettre que le verbe principal, c'est-à-dire *préparer* dans notre exemple (8), est le gouverneur syntaxique de son auxiliaire, puisque c'est lui qui est responsable de sa sélection. L'arbre de dépendances syntaxiques de la phrase est alors celui que nous proposons dans la Figure 2. Pour Crystal (1980), le statut de dépendant serait même définitoire pour les auxiliaires. Une deuxième position (cf. Akmajian et al., 1979) consiste à accorder aux deux éléments verbaux, l'auxiliaire et son verbe principal, un statut syntaxique égal dans une relation horizontale. Cette configuration est représentée dans la Figure 3 ci-dessous. Enfin, selon une troisième perspective (Hudson, 1987, Kahane, 2001), c'est l'auxiliaire, c'est-à-dire *avoir* dans le cas de (8), qui apparaît en position de gouverneur syntaxique, alors que le verbe principal *préparer* est son dépendant. Le réseau de dépendances entre les composantes de la phrase est alors celui de la Figure 4.

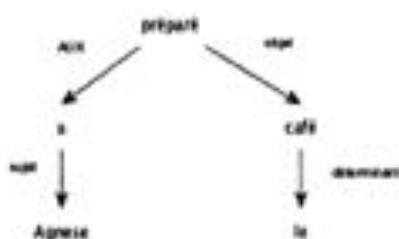


Figure 2 : Réseau de dépendances syntaxiques (hypothèse 1)

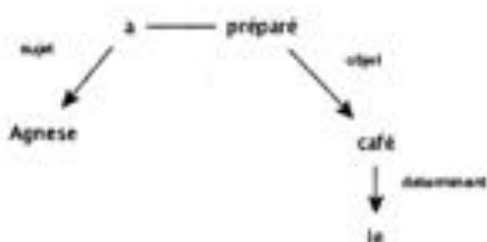


Figure 3 : Réseau de dépendances syntaxiques (hypothèse 2)

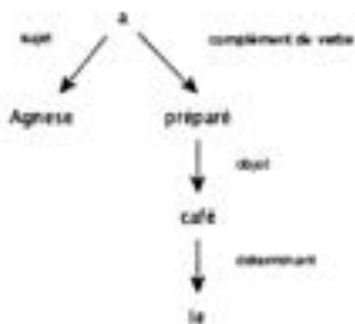


Figure 4 : Réseau de dépendances syntaxiques (hypothèse 3)

Il n'est pas de notre ressort de discuter, dans le présent article, les arguments qui plaident en faveur de chacune de ces trois hypothèses. Ce qui est frappant en revanche, c'est de s'apercevoir que quelle que soit l'option retenue, l'analyse met en avant deux points importants pour notre

comparaison entre les CVP polonaises et le passé composé construit avec l'auxiliaire *avoir*.

On remarque tout d'abord que l'identification des dépendances syntaxiques dans la phrase polonaise ne pose aucun problème, alors qu'en français, la relation entre l'auxiliaire *avoir* et son PartPs, nettement plus complexe, est soumise à trois interprétations différentes. Ce premier fait s'explique à notre avis par le second constat que nous sommes amenée à faire au sujet de la place du PartPs dans les réseaux de dépendances des phrases polonaise et française. Dans (1), le PartPs *przygotowany* n'est pas un dépendant syntaxique de *mieć* (*avoir*), mais il est régi par N_2 . Cela signifie que dans une CVP, le verbe *mieć* (*avoir*) et le PartPs n'ont pas le même rapport formel qu'entretiennent l'auxiliaire *avoir* et le PartPs du verbe principal au passé composé en français.

3.3 Structure syntaxique d'une phrase bipropositionnelle dite à double thème

Il a été souvent souligné que les CVP polonaises présentent une forte ressemblance avec des phrases bipropositionnelles dites à double thème (Furukawa, 1996), telles que (9) :

(9) Pierre a les lunettes cassées.

La phrase (9) a effectivement la même forme syntaxique SN_1 *avoir* SN_2 qui caractérise l'exemple (1), avec comme seule différence la postposition du PartPs à son gouverneur nominal N_2 . Il est intéressant dès lors de vérifier comment cette ressemblance de formule linéaire se traduit en arbre de dépendances syntaxiques, et quelle y est la relation entre *avoir* et le PartPs. L'arbre syntaxique de la phrase (9) est représenté ci-dessous, dans la Figure 5.

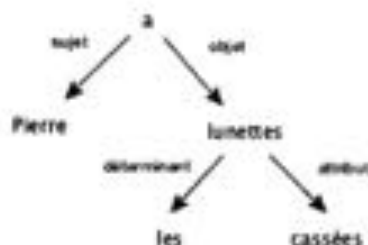


Figure 5 : Réseau de dépendances syntaxiques dans une phrase bipropositionnelle à double thème.

Si les schémas syntaxiques des phrases (1) et (9) ne sont pas identiques – ils diffèrent notamment par la présence du déterminant en français, ils font ressortir une caractéristique structurelle commune. Dans l'exemple polonais (1), comme dans la phrase française (9), les PartPs, respectivement *przygotowany* et *cassé*, sont des dépendants syntaxiques de N_2 , avec lequel ils forment des SN_2 . Cela signifie que le réseau de dépendances et la place du PartPs dans une CVP sont plus proches des phrases à double thème que du passé composé.

3.4 Bilan de la section 3

La comparaison des schémas de dépendances syntaxiques dans les exemples (1), (8) et (9) n'a pas permis de déceler de propriétés formelles communes à la construction polonaise et au passé composé en français. En revanche, elle montre que la CVP dans (1) et la phrase bipropositionnelle en (9) partagent trois caractéristiques :

1. Leurs PartPs sont syntaxiquement dépendants des N_2 , avec lesquels ils forment des SN_2 objets.
2. Il n'y a pas de lien de dépendance directe entre le participe passé et le verbe *mieć* d'une part, *avoir* de l'autre, qui, sauf l'hypothèse 2 (cf. Figure 3)⁹, apparaissent dans les deux cas comme gouverneurs syntaxiques de la phrase.
3. Aussi bien dans (1) que dans (9), le SN_2 est un complément d'objet direct respectivement de *mieć* et de *avoir*. Cela signifie que ces deux verbes n'ont pas perdu, dans ces emplois, la capacité de sélectionner des arguments sémantiques et syntaxiques.

Le point 3 nous amène à quitter le domaine syntaxique pour aborder les caractéristiques que les CVP manifestent à l'interface de la syntaxe et de la sémantique.

⁹ Rappelons que selon l'hypothèse 2, *avoir* ne peut pas être considéré à lui tout seul comme gouverneur syntaxique dans la mesure où il est placé au même niveau de l'arborescence que le PartPs. Sans se trouver invalidé, ce mode d'analyse se révèle peu pertinent pour notre propos.

4. La CVP en polonais et le passé composé en français

En analysant les réseaux de dépendances syntaxiques à l'intérieur des CVP polonaises, nous avons pu constater que le verbe *mieć* ne s'y comporte pas en auxiliaire. Il continue à sélectionner des COD et, à la différence du passé composé français, n'a pas de lien de dépendance direct avec le PartPs. Nous allons à présent vérifier si, du point de vue sémantique, le verbe *mieć* des CVP telles que (1) peut être considéré comme un auxiliaire du parfait.

4.1 *Mieć*, *avoir* et leurs équivalents dans une chaîne de grammaticalisation

Les travaux menés sur des données multilingues, et tout particulièrement les études qui s'inscrivent dans le vaste courant cognitiviste, ont permis, depuis les années 1990, de cerner certaines caractéristiques des auxiliaires du point de vue de leur évolution en diachronie, lors de processus de grammaticalisation (cf. à titre d'exemple, Heine, 1993, Bybee, Perkins et Pagliuca, 1994, Dahl, 2000, Heine et Kuteva, 2002). Si les auxiliaires présentent des particularités propres au système linguistique auquel ils appartiennent, il est désormais possible de les reconnaître grâce à l'identification des chaînes de grammaticalisation (Craig, 1991, Heine et al., 1991) et des critères permettant de les situer sur un continuum allant des verbes sémantiquement pleins aux morphèmes grammaticaux.

Ces travaux ont également permis d'identifier les verbes qui, dans leurs systèmes linguistiques respectifs, sont les meilleurs candidats pour subir une auxiliarisation. Ainsi, les verbes de déplacement tels que *to go*, *aller* ont tendance à évoluer en marqueurs d'intentionnalité, puis en auxiliaires de futur¹⁰.

Le verbe *avoir* et ceux que l'on peut considérer comme ses équivalents sont connus pour avoir évolué, dans plusieurs langues européennes, en marqueurs de parfait. Il ne serait pas étonnant dès lors qu'en suivant le même chemin d'auxiliarisation, le verbe polonais *mieć* devienne lui aussi un auxiliaire de parfait. En allant bien au-delà de cette valeur aspectuelle, Abraham (2013 : 190) constate que « *mieć* is on its way from lexical to auxiliary verb ». L'hypothèse mérite d'autant plus d'être examinée que la

¹⁰ Heine et Kuteva (2002) citent 14 langues dans lesquelles un verbe de déplacement a évolué en un marqueur du futur. Le degré d'auxiliarisation varie néanmoins selon les langues. En français, la forme *aller* + *infinitif* est la plus grammaticalisée des périphrases verbales (Vetters et Lière, 2009). Pour le polonais, nous avons pu montrer (Sikora, 2013) que même si le verbe *ić* ne peut aucunement être considéré comme un auxiliaire, il autorise des emplois caractéristiques de ce que Heine et Kuteva (2002) décrivent comme la phase de *bridging contexts*.

construction que nous analysons ici s'est répandue d'abord à l'oral, avant de pénétrer dans le registre écrit. Or, c'est ainsi que commence tout processus de grammaticalisation.

Pour vérifier cette hypothèse, il nous faut d'abord définir ce qu'est un auxiliaire et repérer des propriétés définitoires valides dans une perspective comparée. C'est seulement après que nous serons en mesure d'interpréter les données et de répondre à la question de savoir si *miéc'* est bien un auxiliaire.

4.2 Définir un auxiliaire

Dans une optique typologique, Heine (1993) propose de définir les auxiliaires de la manière suivante :

« An auxiliary is a linguistic item covering some range of uses along the Verb-to-TAM chain¹¹. » (Heine, 1993 : 70)

Dans les deux langues qui nous intéressent ici, les auxiliaires sont des items verbaux. On pourra donc restreindre cette définition, en admettant qu'un auxiliaire est un verbe renvoyant à des catégories sémantiques qui se situent sur différents points de la chaîne conceptuelle menant d'un sens plein vers l'expression de temps, d'aspect et/ou de modalité.

Les travaux typologiques discutés dans Heine (1993) montrent clairement qu'il est extrêmement difficile de déterminer un ensemble de propriétés définitoires que doit présenter un verbe pour être considéré comme un auxiliaire. Seule la désémantisation (blanchiment sémantique) semble commune aux auxiliaires tels qu'ils apparaissent dans les données multilingues : leur sens lexical s'efface, voire disparaît complètement, pour être remplacé par une fonction grammaticale. Pour un auxiliaire, ce sens fonctionnel consiste à qualifier la situation dénotée par le verbe en ce qui concerne sa localisation dans le temps, ses contours temporels et/ou le type de réalité dont elle relève selon le locuteur, c'est-à-dire à exprimer des concepts grammaticaux liés au temps, à l'aspect et à la modalité.

La définition proposée par Heine (1993) et le degré de désémantisation qui caractérise un verbe permettent de distinguer les auxiliaires des autres verbes assumant une fonction plus ou moins grammaticale :

1. Les verbes supports tels qu'ils sont définis dans des approches de type lexique-grammaire. Un verbe support est sémantiquement (plus ou moins) vide. Le sens du syntagme qu'il forme avec son complément, par exemple *faire de la moto*, est dû au sémantisme du lexème nominal. Dans la mesure où le lexique français ne dispose pas de lexème verbal dont le sens

¹¹ Il s'agit du chemin de grammaticalisation qui conduit un verbe depuis un sens plein (Verb) vers l'expression de temps (T), d'aspect (A) et/ou de modalité (M). Cf. Heine (1993, chapitre 2).

serait « pratiquer l'activité qui consiste à se déplacer en moto », le nom est verbalisé à l'aide d'un verbe support.

Les verbes supports présentent un fort degré de désémantisation et ils assument une fonction bien précise : ils permettent d'exprimer le contenu sémantique d'un lexème nominal sous une forme verbale. Cependant, conformément à notre définition, il ne faut pas les compter parmi les auxiliaires, puisqu'ils ne couvrent aucun segment de la chaîne TAM.

2. Les verbes tels que *commencer à*, *être en train de*, *finir de*, etc., qui permettent de former des périphrases aspectuelles. Ils expriment différentes phases de la situation dénotée par le verbe. Cependant, tout en véhiculant un contenu aspectuel, ils restent des verbes pleins, avec un sens lexical bien précis, ce qui interdit de les inclure dans la classe des auxiliaires.

4.3 *Mieć*, est-il un auxiliaire de parfait ?

En français, le verbe *avoir* qui permet de former les temps analytiques, dont le passé composé, est clairement un auxiliaire. Peut-on en dire autant de *mieć* dans les CVP polonaises ? Les dépendances syntaxiques que nous avons examinées dans la section 3 semblent indiquer que tel n'est pas le cas. Voyons à présent ce que l'on peut en dire d'un point de vue sémantique.

Dans la présente section, nous chercherons à répondre aux deux questions suivantes :

1. Quelle est la catégorie temporelle et/ou aspectuelle qui pourrait lui être associée ? En d'autres termes, il s'agit de savoir comment *mieć* localise la situation dénotée par le verbe sur l'axe du temps (sous-section 4.3.1) et/ou comment il façonne son contour temporel (sous-section 4.3.2).

2. Peut-on considérer que le verbe *mieć* des CVP a subi une désémantisation comparable à celle de l'auxiliaire *avoir* dans les temps composés français (sous-section 4.3.3) ?

4.3.1 *Mieć* et la référence temporelle des CVP en polonais

Rappelons le fragment de l'exemple (8) que nous avons analysé du point de vue syntaxique dans la section 3.

(8) Agnese a préparé le café.

La localisation temporelle de la situation désignée par cette phrase ne pose aucun problème. Elle est située dans le passé, c'est-à-dire à un moment t antérieur au moment d'énonciation t_0 . Le verbe *avoir* apparaît, certes, au présent, mais c'est seulement lorsqu'il est pris avec le PartPs du verbe principal que nous sommes en mesure de déterminer la référence temporelle de la phrase (8). Autrement dit, la valeur T associée à (8) n'est pas celle qu'indique la flexion de l'auxiliaire *avoir*.

En polonais, la référence temporelle des CVP est liée directement au tiroir temporel du verbe *mieć*. Ainsi, les exemples (1), (10) et (11) ci-dessous sont respectivement au présent, au passé et au futur, dont le verbe porte la marque flexionnelle.

- (1) Sąsiadka ma przygotowany
 Sąsiadka.Nom.sg mieć.Prés.3sg przygotowany.PartPs.Acc.sg masc.
 obiad.
 obiad.Acc.sg masc.
 voisine-a-préparé déjeuner (*Le déjeuner de la voisine est prêt*)
- (10) Sąsiadka miała przygotowany
 Sąsiadka.Nom.sg mieć.Passé imperf.3sg fém. przygotowany.PartPs.Acc.sg
 obiad.
 obiad.Acc.sg masc
 voisine-avait-préparé déjeuner (*Le déjeuner de la voisine était prêt*)
- (11) Sąsiadka będzie miała przygotowany
 Sąsiadka.Nom.sg mieć.Fut.3sg fém przygotowany.PartPs.Acc.sg masc
 obiad.
 obiad.Acc.sg masc
 voisine-aura-préparé déjeuner (*Le déjeuner de la voisine sera prêt*)

Le tiroir temporel du verbe *mieć* est bien celui qui place la situation désignée sur l'axe du temps. Ce résultat concorde avec celui de nos analyses syntaxiques, qui n'ont permis de repérer aucun lien de dépendance directe entre *mieć* et le PartPs.

4.3.2 *Mieć* et la valeur de parfait des CVP polonaises

Envisageons à présent l'hypothèse aspectuelle selon laquelle *mieć* serait un auxiliaire véhiculant une valeur de parfait. Autrement dit, plutôt que de couvrir une catégorie T de la chaîne TAM, on lui associerait une signification aspectuelle A. On pourrait ainsi considérer que dans les CVP, *mieć* est un auxiliaire de parfait.

Or, cette hypothèse ne résiste pas aux faits de langue que nous avons rapportés dans la section 1. En y présentant le patron syntaxique N_{Nom} mieć (avoir) PartPs_{Acc} N_{2Acc} commun aux phrases (1)-(5), nous avons vu que seulement certains de ces exemples exprimaient une valeur de parfait, celle-ci étant liée au PartPs dérivé d'un verbe perfectif à sens téléique. Cela signifie qu'on ne saurait imputer à *mieć* la valeur de parfait des exemples (1) à (3).

4.3.3 *Mieć*, est-t-il désémantisé dans les CVP ?

Une fois entré dans le système grammatical, un auxiliaire garde, dans la plupart des cas, un « jumeau » lexical : un même signifiant recouvre à la fois

un ou plusieurs lexèmes pleins et un ou plusieurs sens grammaticaux¹². Les différences sémantiques qui permettent de les distinguer se manifestent notamment dans leurs propriétés distributionnelles. Dans l'exemple (12), *avoir* – auxiliaire et *avoir* – verbe plein sélectionnent des compléments différents.

(12) Il a acheté une nouvelle voiture. Maintenant, il a un gros 4 x 4.

Les deux verbes *avoir* n'ayant pas de contenu sémantique commun, il est impossible de factoriser l'information qu'ils apportent sur le plan phrastique. Par conséquent, il est impossible de coordonner les éléments qui apparaissent en position de leurs compléments syntaxiques (le participe passé *acheté* et le SN *un gros 4 x 4*).

Le verbe *mieć* des CVP se distingue clairement sur ce plan de l'auxiliaire *avoir* en français. L'énoncé (13) en est une illustration :

(13) Mam *posprząta* *mieszkanie*
 avoir.Prés. 1sg *posprzątany*.PartPs.Acc.sg neutre *mieszkanie*.Acc neutre
 i *jeszcze* *czas* *dla*
 i.CONJ. *jeszcze*.ADV. *czas*.Acc. sg masc *dla*.PRÉP.
 siebie. (Internet)
 siebie.PRON.Gén. sg
J'ai rangé-appartement-et-encore-temps-pour-moi (J'ai l'appartement qui est rangé et en plus du temps pour moi)

La possibilité qu'offre le polonais de construire une phrase telle que (13) montre on ne peut plus clairement que le verbe *mieć* des CVP n'est ni complètement désémantisé¹³, ni auxiliarisé. Il conserve un sens lexical qui permet de factoriser le contenu sémantique du verbe *mieć* de la CVP *mam posprząta mieszkanie* et du syntagme verbal *mam jeszcze czas dla siebie* (*j'ai encore du temps pour moi*).

4.4 Bilan de la section 4

Les réflexions exposées ci-dessus conduisent à répondre négativement aux deux questions formulées au début de la section 4.3 au sujet de l'éventuelle auxiliarité de *mieć* dans les CVP. D'une part, étant donné qu'aucune

¹² Abraham (1990) évoque un double rôle des éléments grammaticalisés. Compte tenu des phénomènes de polysémie, il serait plus exact de parler de rôles multiples. En effet, un verbe comme *avoir* possède à la fois plusieurs sens lexicaux et plus d'un sens fonctionnel.

¹³ Certes, il serait difficile, dans cet emploi, de paraphraser *mieć* par *posiadać* (*posséder*). Nous ne nions pas l'évolution sémantique qui l'a éloigné de ce sens de base, mais aucun indice ne permet de conclure à sa désémantisation.

catégorie spécifique de la chaîne TAM ne peut lui être associée, il ne répond pas à la définition des auxiliaires que nous avons retenue. D'autre part, il apparaît également qu'il n'a pas subi une désémantisation comparable à celle de l'auxiliaire *avoir* en français. L'hypothèse d'auxiliarisation de *mieć* formulée par Abraham (2013) semble donc trop forte.

Des arguments sémantiques viennent ainsi étayer les éléments dégagés dans nos investigations syntaxiques. Le verbe *mieć* des CVP n'est pas un auxiliaire.

5. La CVP en polonais et les constructions bithématiques en français

Dans la section 3.3, nous avons observé que les CVP polonaises présentent le même type de dépendances syntaxiques que les constructions bipropositionnelles françaises telles que (9).

(9) Pierre a les lunettes cassées.

Selon l'analyse proposée par Furukawa (1996), il s'agit d'un cas de prédication seconde, plus précisément d'une construction à double thème avec un « petit thème », *lunettes*, en position de complément syntaxique, dont on prédique qu'elles sont cassées, et un « grand thème », *Pierre*, sujet syntaxique, qui fournit ce que Furukawa appelle le cadre thématique de la phrase ou le « grand thème ». Autrement dit, le locuteur choisit comme premier thème *Pierre* pour préciser ensuite le second thème, celui dont il sera véritablement question, en l'occurrence ses lunettes. Si la distinction entre le petit thème (le N₂ *lunettes*) dont le PartPs (*cassées*) est le rhème et le grand thème (*Pierre* dont le SN₂ *les lunettes cassées* constitue le rhème) se dessine clairement dans la structure de la phrase, les choix terminologiques restent discutables en ce qui concerne la structure communicative, c'est-à-dire l'élément localement prééminent (Kahane, 2001). En effet, on imagine facilement des situations où le petit thème et son rhème deviennent grands : par exemple, lorsque l'on utilise la phrase (9) pour expliquer à une vieille connaissance de Pierre pourquoi celui-ci ne l'a pas saluée dans la rue. L'information communicativement prééminente est alors celle des lunettes cassées, et elle vise l'inférence pragmatique 'Pierre n'a pas pu te voir'.

Dans la présente section, notre objectif est de vérifier si les ressemblances structurelles entre les CVP et les phrases bipropositionnelles telles que (9) trouvent une confirmation sur le plan sémantique.

5.1 Contraintes sémantiques dans les phrases bipropositionnelles et dans les CVP

Furukawa (1996) observe que les phrases bipropositionnelles à double thème, telles que *Sylvie a les yeux bleus* ou *Pierre a les lunettes cassées*, sont

soumises, en français, à une contrainte sémantique forte « qui exige que tout SN₂ figurant dans ce type de phrases soit partie intégrante du tout constitué par le SN₁ » (Furukawa, 1996 : 92). En fait, le N₂ des constructions à double thème dénote soit une partie (du corps) de N₁ (*Sylvie a les yeux bleus*¹⁴), soit un individu qui reste lié à N₁ (*Elle a son mari malade*¹⁵), soit un artefact dont N₁ désigne l'utilisateur (*Paul a les lunettes cassées*). Il semble légitime dès lors de nuancer la contrainte formulée par Furukawa, dans la mesure où un artefact utilisé par Paul ne peut être considéré comme sa partie intégrante.

Nous allons à présent vérifier si une contrainte sémantique similaire pèse sur les CVP polonaises. La question est importante, car elle nous amène à l'interface de la sémantique et de la syntaxe. En nous interrogeant sur le rapport entre les entités dénotées respectivement par N₁ et par N₂, nous n'échapperons pas au problème du PartPs, du verbe dont il est dérivé et de son sujet : N₁ est-il ou n'est-il pas le sujet grammatical du verbe perfectif et télique dont le PartPs figure dans une CVP ? C'est après avoir éclairci ces deux points que nous serons en mesure de déterminer le rôle qu'assument N₁ et le verbe *mieć* dans une CVP, et de comprendre la nature de cette construction.

5.2 Relâchement de la contrainte sémantique dans les CVP

Rappelons les exemples (1), (2), (3) discutés dans la section 2.2.

- (1) Sąsiadka ma przygotowany
 Sąsiadka.Nom.sg mieć.Prés.3sg przygotowany.PartPs.Acc.sg masc
 obiad.
 obiad.Acc.sg masc
 voisine-a-préparé déjeuner (*Le déjeuner de la voisine est prêt*)
- (2) Pies ma zrobione
 Pies.Nom.sg mieć.Prés.3sg zrobiony.PartPs.Acc.pl neutre
 szczepienia.
 szczepienie.Acc.pl neutre.
 chien-a-faites vaccinations (*Le chien a toutes les vaccinations*)
- (3) Sąsiadka ma zrobioną
 Sąsiadka.Nom.sg mieć.Prés.3sg zrobiony.PartPs.Acc.sg fém.
 trwałą.
 trwałość.Acc.sg fém.
 voisine-a-fait permanente (*La voisine a une permanente*)

La contrainte qui nous intéresse semble avoir un poids inégal dans ces trois CVP. Elle se confirme pleinement dans (3) puisqu'une permanente

¹⁴ Exemple emprunté à Furukawa (1996 : 87)

¹⁵ Exemple emprunté à Furukawa (1996 : 88)

($N_2 = trwała$) fait incontestablement partie intégrante de la voisine ($N_1 = sąsiadka$). On ne peut pas en dire autant d'un repas ($N_2 = obiad$) dans (1), ni des vaccinations ($N_2 = szczepienia$) dans (2). Forcé est de constater alors que la contrainte sémantique formulée pour le français ne tient pas dans le cas des CVP polonaises même dans sa version allégée, puisque N_2 ne dénote pas nécessairement un artefact dont N_1 serait l'utilisateur.

5.3 SN_2 dérivé d'un syntagme verbal

Le PartPs du SN_2 est formé sur un verbe dont N_2 pourrait être le complément direct. Sémantiquement, il s'agit de verbes de réalisation, équivalents de *faire, préparer, etc.* Autrement dit, le SN_2 est formé du PartPs d'un verbe perfectif dénotant une situation télique et d'un N_2 qui pourrait être son COD. Dans nos exemples (1) à (3), les SN_{2Acc} sont donc issus des syntagmes verbaux de forme $V_{\text{Part}} + \text{Telic} [N_2] N_{2Acc}$ suivants :

- (1) $SN_{2Acc} = \text{przygotowany obiad}$ (*déjeuner préparé*) – $VN_{2Acc} = \text{przygotować obiad}$ (*préparer le déjeuner*)
- (2) $SN_{2Acc} = \text{zrobione szczepienia}$ (*vaccinations faites*) – $VN_{2Acc} = \text{zrobić szczepienia}$ (*faire les vaccinations*)
- (3) $SN_{2Acc} = \text{zrobioną trwałą}$ (*permanente faite*) – $VN_{2Acc} = \text{zrobić trwałą}$ (*faire une permanente*)

La formule $V_{\text{Part}} + \text{Telic} [N_2] N_{2Acc}$ est pourtant incomplète, car il lui manque l'élément nominal qui est son sujet syntaxique. Peut-on attribuer ce rôle à N_1 , respectivement *sąsiadka* (*voisine*) et *pies* (*chien*) dans nos exemples ci-dessus ? La réponse à cette question est le dernier élément qui nous permettra d'expliquer la nature des CVP.

5.4 Y a-t-il un agent dans la CVP ?

Compte tenu du fait que le SN_{2Acc} ($\text{PartPs}_{Acc} N_{2Acc}$) des CVP est formé sur le syntagme verbal $V_{\text{Part}} + \text{Telic} [N_2] N_{2Acc}$, nous nous en servirons à présent pour construire les paraphrases de nos exemples (1), (2) et (3). Notre objectif est de vérifier si N_{1Nom} des CVP est le sujet de $SV_{\text{Part}} + \text{Telic} [N_2]$ et si tel n'est pas le cas, quelle en est la fonction.

Les $V_{\text{Part}} + \text{Telic} [N_2]$ en question sélectionnent deux arguments dont le second est réalisé par N_{2Acc} . La question de l'argument sujet est plus délicate. Peut-on admettre que cette fonction est assurée par N_1 ? S'agit-il d'un agent ? Sans être sûre de pouvoir y répondre par l'affirmative, nous désignerons l'argument sujet par N_X .

Reprenons à présent nos exemples (1), (2) et (3) sous (1a), (2a), (3a), alors qu'en (b) et (c), nous présentons leurs interprétations possibles.

- (1) a. Sąsiadka ($N_{1(\text{Dm})}$) ma przygotowany obiad ($N_{2(\text{Akc})}$).
 b. $N_{X(\text{Dm})}[N_X = N_1]$ przygotował(a) obiad ($N_{2(\text{Akc})}$) [sąsiadce ($N_{1(\text{Dm})}$)].
 c. $N_{X(\text{Dm})}[N_X \neq N_1]$ przygotował(a) obiad ($N_{2(\text{Akc})}$) sąsiadce ($N_{1(\text{Dm})}$).
 (2) a. Pies ($N_{1(\text{Dm})}$) ma zrobione szczepienia ($N_{2(\text{Akc})}$).
 b. $N_{X(\text{Dm})}$ zrobił(a) szczepienia ($N_{2(\text{Akc})}$) psu ($N_{1(\text{Dm})}$).
 (3) a. Sąsiadka ($N_{1(\text{Dm})}$) ma zrobioną trwałą ($N_{2(\text{Akc})}$).
 b. $N_{X(\text{Dm})}$ zrobił(a) trwałą ($N_{2(\text{Akc})}$) sąsiadce ($N_{1(\text{Dm})}$).

Il convient de souligner tout d'abord que les versions (b) et (c) de nos exemples ne sont pas des paraphrases exactes des phrases (a) construites sur les CVP. Comme nous l'avons montré dans la section 4.3.1, la référence temporelle des CVP est indiquée par le tiroir temporel du verbe *mieć*. Les versions (a) de nos exemples désignent donc un état résultant *r* présent, alors que (b) et (c) désignent la situation passée *e* qui a donné lieu à *r*. Sur le plan discursif, il est possible de construire une relation de type Résultat (Asher et Lascarides, 2003) entre la situation *e* désignée par une phrase (b), par exemple $N_X[N_X = N_1]$ przygotował(a)¹⁶ obiad (N_2) [sąsiadce (N_1)] ou (c) – $N_X[N_X \neq N_1]$ przygotował(a) obiad (N_2) sąsiadce (N_1) et une situation *r* associée à l'énoncé (a) Sąsiadka (N_1) ma przygotowany obiad (N_2).

Dans la paraphrase (1b), l'identité de N_X et de N_1 est hypothétique, avec néanmoins un degré de probabilité élevé. L'interprétation par défaut consiste à penser que la voisine a elle-même préparé le déjeuner dont elle est également bénéficiaire. Une lecture (c) reste néanmoins possible : elle présente la voisine dans ce second rôle de bénéficiaire, et non pas dans celui d'agent. Ce second rôle est indiqué par le complément circonstantiel datif entre crochets.

Dans (3), il n'est pas exclu que la voisine (N_1) ait réalisé sa permanente toute seule, mais l'expérience quotidienne montre que de tels exploits sont rares, ne serait-ce que pour des raisons pratiques. L'interprétation par défaut consiste alors à différencier l'agent N_X et le bénéficiaire N_1 exprimé au datif. Remarquons que tout comme dans (2b) analysé ci-dessous, le complément datif n'est pas un circonstantiel, mais un argument du verbe : il exprime l'un de ses actants.

Dans les exemples (2), N_1 (pies – chien) n'est pas soupçonné de s'être vacciné tout seul. N_X ne peut être que différent de N_1 . Le datif dont N_1 porte la marque flexionnelle dans (2b) indique clairement que l'entité dénotée est présentée comme patient ou bénéficiaire de l'action désignée par le syntagme verbal.

¹⁶ Au passé, le verbe polonais s'accorde en genre et en nombre avec son sujet. Le morphème *-a*, qui marque le féminin, est donc proposé entre parenthèses pour les cas où N_X serait au féminin.

Le rôle d'agent qu'il n'est pas exclu d'accorder, dans certaines CVP, à l'argument sujet N_1 , n'est lié ni à la structure syntaxique, ni aux indices sémantiques. Il est dû uniquement aux facteurs extralinguistiques et pragmatiques. D'un point de vue linguistique, l'agent N_X , qui apparaît en position de sujet dans les phrases (1b), (1c), (2b) et (2c), se trouve effacé dans les CVP (1a), (2a) et (3a).

Cet effacement de l'agent n'est pas indifférent pour le lien entre l'état résultant r et la situation e dénotée par les CVP. Malgré un rapport sémantique fort entre le PartPs des CVP et le $V_{\text{PART+TENSUM}}$ qui dénote e , l'absence d'agent ne permet pas d'instancier e . Dans Sikora (2009), nous avons pu constater qu'un agent offre un ancrage spatio-temporel, en constituant e en une instance situationnelle. Son absence dans les CVP fait que le lien entre r et e qui lui est antérieur ne peut se construire que par inférence pragmatique. On comprend aussi pourquoi il s'agit d'une inférence qu'Apothéloz et Nowakowska (2010 : 3) qualifient d'illative :

« Il s'agit du cas où la désignation de la situation qui vise à informer, non pas de la situation proprement dite, mais d'une conséquence de celle-ci. Autrement dit, de l'inférence situation \Rightarrow résultat, on ne retient pratiquement ici que le résultat. Ce dernier est valide et pertinent dans le contexte d'énonciation. »

6. Pour conclure : qu'est-ce qu'une CVP en polonais ?

Les analyses menées dans ces pages devaient permettre d'identifier le statut linguistique de la construction polonaise N_{INSTR} *mieć* (avoir) PartPs_{AUX} N_{DACC} à valeur de parfait (CVP). Elles se sont déroulées selon deux axes. Dans un premier temps, nous avons voulu vérifier le degré de ressemblance, souvent mis en avant dans la littérature, entre les CVP et les phrases au passé composé en français d'une part, et les phrases bipropositionnelles à double thème de l'autre. Deuxièmement, nous nous sommes intéressée aux caractéristiques des éléments qui composent une CVP. Notre objectif était de comprendre si les CVP relèvent de la flexion verbale ou de la syntaxe.

Nos analyses ont permis d'établir les points suivants :

1. Le réseau de dépendances syntaxiques dans une CVP situe clairement les structures de ce type du côté des phrases bipropositionnelles à double thème, et non pas parmi les phrases au passé composé. Le verbe *mieć* des CVP ne présente pas de caractéristiques qui permettraient de lui accorder un statut d'auxiliaire.

2. Deux éléments de nature fort différente sont à l'origine d'une temporalité plutôt originale des CVP, notamment lorsqu'on la compare à celle des parfaits en français et en anglais. D'une part, *mieć* n'étant pas un auxiliaire, c'est le tiroir temporel dont il porte la marque qui est déterminant pour situer dans le temps l'état résultant qu'il désigne. D'autre part, la

structure sémantique et syntaxique des CVP ne permet pas d'identifier l'agent impliqué dans la situation antérieure, qui a donné lieu à l'état résultant.

Qu'est-ce qu'une CVP alors ? Compte tenu des caractéristiques qui se sont manifestées dans nos analyses, une CVP est un cas particulier de construction bipropositionnelle à double thème.

En effet, une CVP articule deux prédications que nous illustrons à l'aide de l'exemple (1).

- (1) Sąsiadka ma przygotowany obiad.
voisine-a-préparé déjeuner (*Le déjeuner de la voisine est prêt*)

La première prédication s'opère à l'intérieur du SN₂ et elle porte sur N₂ (*obiad – déjeuner*). Elle se fait grâce au PartPs (*przygotowany – préparé*). Dérivé d'un verbe télique, il peut entrer dans une construction résultative. Communicativement, N₂ est un premier thème, alors que son PartPs est le rhème 1.

La seconde prédication s'effectue à l'aide du verbe *mieć* qui « raccroche » le SN₂ (*przygotowany obiad – déjeuner préparé/prêt*) à N₁ (*sąsiadka – voisine*), en lui assignant une certaine propriété, en l'occurrence un état de choses. C'est dire que sans être un auxiliaire, il assume une fonction dans ce type de constructions bipropositionnelles à double thème. D'un point de vue syntaxique, N₁ est le sujet de la phrase, mais nous avons pu constater qu'il s'agit en réalité du bénéficiaire de la situation. La construction N₁_{NOM} *mieć* (avoir) PartPs_{ACC} N₂_{ACC} permet ainsi de le thématiser, en le positionnant en tête de phrase, et de lui associer comme rhème le SN₂. La particularité d'une CVP réside dans le fait que son thème 2 s'exprime dans une construction résultative.

Abréviations

ADV – adverbe, Acc – accusatif, CONJ – conjonction, CVP – construction à valeur de parfait, Dat – datif, fém – féminin, Fut – futur, Gén – génitif, imperf – imperfectif, ind – indicatif, masc – masculin, N – nom, PartPs – participe passé, perf – perfectif, pl – pluriel, PRÉP – préposition, PRON – pronom, sg – singulier, SN – syntagme nominal, SV – syntagme verbal

Références

- Abraham, W. (1990). Die Grammatikalisierung von Auxiliar- und Modalverben, in: U. Demske; J. Haustein; S. Köbele; D. Nübling (eds), *Beiträge zur Geschichte Sprache und Literatur*, 112: 200-208.
Abraham, W. (2013). The developmental of the analytic past in German and in Polish, in: T. Lohndal (ed.), *In Search of Universal Grammar: From*

- Old Norse to Zoque*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 175-194.
- Akmajian, A.; Steel, S.; Wasow, T. (1979). The category AUX in Universal Grammar. *Linguistic Inquiry* 10: 1-64.
- Apothéloz, D.; Nowakowska, M. (2010). La résultativité et la valeur de parfait en français et en polonais. *Cahiers Chronos* 21 : 1-23.
- Asher, N.; Lascarides, A. (2003). *Logics of conversation*. Cambridge: Cambridge University Press.
- Bybee, J.; Perkins, R.; Pagliuca, W. (1994) *The Evolution of Grammar. Tense, Aspect, and Modality in the Languages of the World*. Chicago, London: The University of Chicago Press.
- Caudal, P.; Vetter, C. (2007). Passé composé et passé simple : Sémantique diachronique et formelle. *Cahiers Chronos* 16 : 121-151.
- Comrie, B. (1976). *Aspect: An Introduction to the Study of Verbal Aspect and related Problems*. Cambridge: Cambridge University Press.
- Craig, C., G. (1991). Ways to go in Rama: A case-study in polygrammaticalization, in: E. C. Traugott; B. Heine, (eds) *Approaches to grammaticalization*, vol. 2, Amsterdam/Philadelphia: John Benjamins, 455-492.
- Creissels, D. (2000). L'emploi résultatif de 'être + participe passé' en français, *Cahiers Chronos*, 6 : 133-142.
- Crystal, D. (1980). *A first dictionary of linguistics and phonetics*. London: Andre Deutsch.
- Dahl, Ö. (2000). The tense-aspect system of European languages in a typological perspective, in: Ö. Dahl (ed.), *Tense and Aspect in the Languages of Europe (Empirical Approaches to language typology 20-6)*, Berlin: de Gruyter, 385-401.
- Depraetere, I. (1995). On the Necessity of Distinguishing between (Un)Boundedness and (A)Telicity. *Linguistics and Philosophy*, 18, 1-19.
- Furukawa, N. (1996). *Grammaire de la prédication seconde. Forme, sens et contraintes*. Louvain-la-Neuve : Duculot.
- Heine, B. (1993). *Auxiliaries. Cognitive Forces and Grammaticalization*, New York, Oxford: Oxford University Press.
- Heine, B.; Claudi, U.; Hünnemeyer, F. (1991). *Grammaticalization: A conceptual Framework*, Chicago: Chicago University Press.
- Heine, B.; Kuteva, T. (2002) *World Lexicon of Grammaticalization*, Cambridge: Cambridge University Press.
- Huddleston, R., D. (1984). *Introduction to the Grammar of English*, Cambridge: Cambridge University Press.
- Hudson, R. (1987). Zwicky on heads. *Journal of Linguistics*, 23: 109-132.

- Kahane, S. (2001). Grammaires de dépendances formelles et théorie Sens-Texte. <http://www.kahane.fr/textespedagogiques.html> [consulté le 25 mai 2014].
- Muryn, T. (2009). Le passé composé en fonction de parfait non-testimonial. Une analyse contrastive français-polonais, in: J. Kortas ; J. Jereczek-Lipińska ; G. Quentel (éds), *La globalisation communicationnelle : Enrichissement et menace pour les langues*, Gdańsk : Fundacja Rozwoju Uniwersytetu Gdańskiego, 289-295.
- Paprocka-Piotrowska, U. (2008). *Contier au risque de tout changer*. Towarzystwo Naukowe KUL, Katolicki Uniwersytet Lubelski Jana Pawła II.
- Pilecka, E. (2013). Le polonais aurait-il un passé composé ?, in: Marsac, F., Pellat, J.-Ch. (éds) *Le participe passé entre accords et désaccords*. Strasbourg : Presses universitaires de Strasbourg, 147-160.
- Sikora, D. (2009). *Les verbes de manière de mouvement en français et en polonais. Éléments pour une étude comparée des propriétés structurelles des prédicats*, Thèse de doctorat, Nancy-Université.
- Sikora, D. (2012a). Compléments aspectuels et verbes de manière de déplacement en français : entre bornage et télécité. *Études Cognitives/Cognitive Studies*, 12, Varsovie : Éditions de l'Académie Polonaise des Sciences, 57-78.
- Sikora, D. (2012b). Portée des préfixes et complexité en français et en polonais. Le cas de *RE-* versus *PO-*, *PRZY-*, *DO-*, in: Paprocka-Piotrowska, U., Martinot, C., Goolimich, S. (éds) *La Complexité en langue et son acquisition*, Lublin : Towarzystwo Naukowe KUL, 131-147.
- Sikora, D. (2013). Is the Polish verb *ić* an auxiliary to be? A corpus-based study of the construction *ić* + *infinitive*, in: Kor Chahine, I. (ed.). *Current studies in Slavic Linguistics*. Amsterdam/Philadelphia: John Benjamins, 123-137.
- Vetters, C. ; Lière, A. (2009). Quand une périphrase devient temps verbal : le cas d'*aller* + *infinitif*, *Faits de langues* 33: 27-36.
- Vetters, C. ; Skibińska, E. (1998). Le futur : une question de temps ou de mode ? Remarques générales du « présent-futur » perfectif polonais. *Cahiers Chronos*, 2, Amsterdam – Atlanta, GA : Rodopi, 247-266.

DOROTA SIKORA

AJEF – CNRS (UMR 7118)

dorota.sikora@ajef.fr

COMPLÉMENTS ASPECTUELS ET VERBES DE MANIÈRE DE DÉPLACEMENT EN FRANÇAIS : ENTRE BORNAGE ET TÉLICITÉ¹

Résumé

Considérées comme équivalentes par certains, soigneusement distinguées par d'autres, les notions de télécité et de bornage ont une importance cruciale pour les études consacrées à l'aspect. Cet article propose une réflexion sur le rapport entre ces propriétés aspectuelles et certains compléments verbaux. À travers des analyses détaillées des syntagmes verbaux que forment *courir* et *nager* (*courir cent mètres*, *nager le deux-cents-mètres*), il est montré que malgré une similarité formelle apparente, les compléments tels que *cent mètres* et *le deux-cents-mètres* assument des fonctions différentes auprès du prédicat : dans le premier cas, ils opèrent un bornage, alors que dans le second, ils sont porteurs du trait [+ télécité]. L'étude d'un problème spécifique permet ainsi de prendre part dans une discussion plus générale, en dégagant des arguments en faveur de la distinction entre télécité et bornage.

Mots-clés : compléments aspectuels, télécité, bornage

Introduction

En 1967, Vendler écrivait :

« It appears, then, that running and its kind go on in time in a homogeneous way ; any part of the process is of the same nature as the whole. Not so with running a mile or writing a letter; they also go on in time, but they proceed toward a terminus which is logically necessary to their being what they are. »

Vendler (1967: 101)

Ainsi, les verbes *to run* et *to write* dénotent des activités, alors que *to run a mile* et *to write a letter* désignent des accomplissements². Les premiers sont atéliques,

¹Je tiens à remercier Denis Apothélos pour sa relecture d'une première version de cet article et pour ses remarques qui m'ont permis d'améliorer sensiblement ce texte.

²Tout au long de ce texte, nous nous reporterons à la typologie de Vendler (1967). Bien que révisée, perfectionnée et réorganisée à plusieurs reprises, elle continue à constituer un classement de référence qui ne nécessite pas de présentation supplémentaire.

alors que dans le second cas, la télécité impose aux situations un caractère téléonomique : elles progressent à travers le temps vers un terminus spécifié respectivement par les compléments *a mile* et *a letter*.

Dans le présent article, nous nous proposons d'examiner des énoncés français qui ressemblent fortement à celui de *run a mile* discuté par Vendler. Les exemples (1) à (4) en apportent l'illustration :

- (1) J'ai nagé un kilomètre et espère m'être musclée un petit peu... en tous cas, je me sens bien fatiguée! (forum.aufeminin.com)
- (2) J'ai couru 100 mètres avec trois types qui étaient sortis d'un bar et m'ont suivi pour me demander ce que je fais. (www.facebook.com)
- (3) Cathy a nagé le 200 mètres papillon. (blog.crdp-versailles.fr)
- (4) On a commencé par un peu d'échauffement, puis on a couru le 1000 mètres. (new.ecolemartigny.ch)

La problématique de dérivation aspectuelle a été beaucoup explorée depuis Vendler. Nous avons affaire à la dérivation aspectuelle le processus qui s'opère sur le plan de la prédication et qui conduit à modifier l'aspect lexical du verbe au moyen de sa complémentation ou bien par des temps grammaticaux.

De nombreuses recherches ont contribué à identifier et à formaliser le rôle des compléments pour la structure temporelle des situations (voir, à titre d'exemple, Krifka 1998)³, mais en dépit des avancées considérables dans ce domaine, quelque confusion semble se maintenir sur un point important lié à notre propos. En effet, les travaux que nous avons consultés semblent admettre tacitement que *to run a mile* cité par Vendler ainsi que *nager un kilomètre* de notre exemple (1) relèvent du même cas que *to write a letter* ou son équivalent français *écrire une lettre*, ipse *eo* dans les deux cas, nous aurions affaire à la dérivation aspectuelle transformant une activité en accomplissement par adjonction d'un complément approprié. Par voie de conséquence, rien n'empêche de conclure dès lors que *to run a mile*, *to write a letter* du côté anglais, ainsi un kilomètre, cent mètres, le deux-cents-mètres et le mille-mètres⁴ de nos exemples français (1) à (4) doivent recevoir un même traitement en tant que porteur du [+télécité].

Certes, dans tous les cas que nous venons d'évoquer, les compléments qui nous intéressent sont des syntagmes nominaux. Qui plus est, du point de vue sémantique, un vague air de famille semble jouer en faveur d'un rapprochement. Ainsi, Cappelle et Declercq (2005) indiquent que dans l'exemple (5), nous avons affaire à des compléments désignant la trajectoire de déplacement par spécification de sa longueur :

- (5) They ran {five kilometers / two laps / a marathon} (Cappelle et Declercq, 2005: 892)

³Nous reprenons ce fil de notre réflexion dans la section suivante.

⁴Nous tenons à signaler au lecteur le fait que, dans tous les exemples que nous avons identifiés, les compléments du type de ceux qu'on trouve dans les énoncés (3) et (4) sont notés en chiffres. La transcription en lettres est donc notre fait. La raison pour laquelle nous avons choisi de recourir aux traits d'union est exposée dans la section 2.1.1.

Remarquons que les compléments *five kilometers*, *two laps*, *a marathon* sont énumérés sous forme d'une liste de valeurs substituables. On peut donc raisonnablement admettre qu'aux yeux des auteurs, il s'agit d'expressions de même nature.

Le présent article se pose un double objectif. Nous nous efforcerons de montrer tout d'abord qu'en ce qui concerne le français, l'analyse linguistique fournit de bonnes raisons pour distinguer entre, d'une part, *nager un kilomètre* de (1) et *courir cent mètres* de (2), et de l'autre, *nager le deux-cents-mètres* et *courir le mille-mètres* des exemples (3) et (4). Nous serons ainsi amenés à suggérer, dans un second temps, que seuls les deux derniers relèvent du même cas que *écrire une lettre*, en permettant une dérivation aspectuelle.

En nous appuyant sur les travaux récents menés en sémantique de l'espace, nous commencerons, dans la première section, par vérifier s'il est légitime d'interpréter tous ces compléments en termes de trajectoire. Nous y exposerons également les articulations entre ce paramètre du déplacement et certaines propriétés aspectuelles des situations. Dans la deuxième section, nous passerons en revue les caractéristiques linguistiques de ces compléments en français. Comme nous nous efforcerons de le montrer dans la troisième section, les contrastes observés s'expliquent par un statut fonctionnel différent de ces deux types d'expressions, notamment en ce qui concerne leur rôle auprès des prédicats : si dans les exemples (3) et (4), les compléments sont des réalisations syntaxiques d'arguments internes, dans (1) et (2), ils assument des fonctions de modificateurs adverbiaux. Nous concluons, dans la quatrième et dernière section, par une discussion des implications que la distinction entre arguments et modificateurs a dans le domaine aspectuel, notamment pour ce qui est de la différence entre téléicité et bornage et de l'information qu'apporte en réalité la compatibilité d'une structure prédicative avec l'adverbiale en *X* temps.

1. Articulation de propriétés spatiales et aspectuelles à travers la notion de trajectoire

Depuis plus d'un quart de siècle néanmoins, la trajectoire (*path*) est devenue un lieu de rencontres privilégié de la sémantique de l'espace et de l'aspectologie⁵.

C'est à Talmy (2000)⁶ que nous devons le modèle largement utilisé aujourd'hui dans l'analyse de la conceptualisation de l'espace dans les langues. Le schéma de base de tout déplacement s'inscrit dans le cadre général décrit ainsi :

« The basic Motion event consist of one object (the Figure) moving or located

⁵Ce rapprochement n'est pas sans provoquer des réserves. Cf. Aunargue (2008, 2011) qui postule une ligne de démarcation claire et nette entre l'analyse des propriétés spatiales et aspectuelles. Tout particulièrement, le titre *How motion verbs are spatial* (Aunargue 2011) est une réponse polémique à Tenny (1995). Plus précisément, Aunargue postule la priorité de la description spatiale sur les analyses aspectuelles, cf. Aunargue (2011: 3) : « Relating spatial and aspectual properties of motion processes is not excluded but this task will have to be applied at some later stage, on the basis of a clear spatial characterization ». Nous souscrivons entièrement à cette vue en ce qui concerne le *modus operandi*, d'autant plus que, comme dans tous les autres groupes verbaux, le sens lexical est un élément déterminant pour les caractéristiques aspectuelles. Nous pensons néanmoins que dans son stade présent de développement, la sémantique de l'espace met à la disposition des chercheurs l'outillage conceptuel nécessaire pour entamer la description des articulations entre les propriétés spatiales et aspectuelles.

⁶La première version de ce texte fut publiée en 1987. Pour notre part, nous utilisons sa version revue dans Talmy (2000/1)

with respect to another (the reference object or Ground). It is analyzed as having four components. Besides Figure and Ground, there is Path and Motion. The Path is the path followed or site occupied by the Figure object with respect to the Ground object. »

Talmy (2000/1 : 25)

Considérée comme l'un des quatre composants conceptuels de base de tout déplacement (Talmy 2000), la trajectoire fait l'objet de nombreuses études qui se posent comme objectif de décrire et de systématiser la diversité de ses expressions linguistiques (formes lexicales, affixes, prépositions, etc.), aussi bien dans des perspectives intralinguistiques que comparées (cf. les travaux typologiques effectués dans le cadre du projet « Trajectoire »⁷, Pourcel et Kopecka 2005).

Parallèlement, ce paramètre spatial, initialement destiné à conceptualiser le déplacement, revêt une importance particulière dans l'étude de l'aspect dans la mesure où il permet, comme le montre Krifka (1992) à travers la thèse d'homomorphisme, de rendre compte de certaines propriétés aspectuelles des situations.

1.1 Trajectoire en sémantique de l'espace⁸

En parlant de *come* et *go*, Fillmore (1983 : 219) remarque que le mouvement linéaire qu'ils désignent commence à un moment dans un endroit et se termine plus tard, dans un autre lieu. Cette propriété forme la base de la distinction, couramment pratiquée aujourd'hui en sémantique de l'espace du français, entre les verbes de déplacement (*aller*, *sortir*, *partir*, etc.) et les verbes de manière de déplacement (*marcher*, *courir*, *nager*, etc.). Pour les premiers, on s'accorde pour parler de déplacement, puisqu'ils désignent un mouvement qui commence au moment t_0 dans un lieu l_0 pour s'arrêter à un moment t_n dans le lieu l_n . Pour les seconds, il convient d'admettre que si la trajectoire n'est pas lexicalisée dans ces verbes, il n'en est pas moins fréquent de les rencontrer dans des contextes qui réfèrent clairement à un déplacement, notamment quand le verbe est accompagné d'un complément interprété en termes de trajectoire, comme cela a lieu dans les exemples (6) et (7) ci-dessous.

- (6) Nous sommes restés une heure à la grotte, après quoi, nous avons couru à l'hôtel, j'ai étendu mon python sur le lit et j'ai attendu. (Frantext, Romain Gary, *Gros-Célin*, 1974)
- (7) Le soir enfin, après l'étude, Boris et moi avons couru vers les immeubles cernant une grande cour (Frantext, Schreiber, B. *Un silence d'environ une demi-heure*, 1996)

Il apparaît clairement dans (6), que l'agent nous exerce non seulement un mouvement consistant à courir, mais également – et peut-être avant tout – un déplacement, puisqu'au moment t_0 , il part d'un endroit l_0 pour se retrouver, au terme, c'est-à-dire à t_n , dans le lieu l_n , en l'occurrence à son hôtel. *À l'hôtel* désigne donc

⁷Cf. www.cdd-lyon.cnrs.fr/trajectoire/ProjetGF.html

⁸Les travaux consacrés à la trajectoire sont tellement nombreux qu'il nous est impossible d'en proposer ne serait-ce qu'un tour d'horizon dans le cadre de cet article. Par conséquent, la section 1.1 n'est pas pensée comme un exposé de l'état de l'art dans ce domaine. Son objectif consiste à mettre en avant les éléments qui ont permis d'employer la notion de trajectoire pour étudier les propriétés aspectuelles des situations.

un site l_n atteint au moment t_n qui marque le terme de la situation plutôt que la trajectoire, celle-ci restant néanmoins possible à inférer. Dans (7), au contraire, la borne $t_n l_n$ n'est pas précisée, puisque le complément *vers les immeubles* indique seulement que la course se déroule selon une trajectoire menant les garçons, au fil des moments t_1, t_2, t_3, \dots , par différents lieux l_1, l_2, l_3, \dots qui se suivent par rapport au site spécifié (*les immeubles*).

Il est donc possible, à travers la notion de trajectoire, de mettre en relation les propriétés spatiales et temporelles des situations. Les compléments qui expriment le terme du déplacement (*à l'hôtel*) et la trajectoire (*vers les immeubles*) ne sont pas indifférents pour les propriétés structurelles des situations qu'ils désignent. Il suffit de comparer notre exemple (6) à un cas comme (8)

- (8) Elle a couru plus vite que moi! (Prantext, D'Ormesson, J. *Le bonheur à Minato*, 1987),

pour remarquer que dans le premier, nous pouvons – grâce aux compléments *à l'hôtel* – identifier les paramètres spatio-temporels $t_n l_n$ qui marquent le terme de la situation, alors que celle associée au prédicat que forme *courir* dans (8) en est dépourvue. Plus précisément, dans l'exemple (8), la borne temporelle finale t_n est imposée par le temps grammatical, mais elle ne correspond à aucune borne spatiale. Dans la section 1.4, nous reprendrons la réflexion sur les différents types de bornes, en expliquant notre position quant à la distinction qu'il convient d'établir entre celles qui portent le trait de télicité et celles qui effectuent une opération de délimitation spatiale ou temporelle.

1.2 Trajectoire et propriétés des situations

Afin de formaliser les liens entre les propriétés spatiales et temporelles des situations via la trajectoire, les aspectologues ont réinterprété cette notion à la lumière des études développées par Dowty (1979) sur les arguments sémantiques assumant le rôle de thème incrémentaire. En effet, ce rôle thématique est attribué aux arguments internes du prédicat, lorsqu'il s'agit d'objets dont l'extension détermine la progression de la situation à travers le temps. Tel est le cas de la lettre dans *écrire une lettre*. Parvenu au bout, son auteur atteint le terme de la situation qui débouche sur un nouvel état des choses : *la lettre est écrite*.

En s'appuyant sur les travaux de Dowty (1979, 1991), ainsi que sur l'étude de Krifka (1992), Tenney (1994, 1995) propose de distinguer deux rôles aspectuels assumés par les arguments internes des prédicats. Un argument interne, syntaxiquement réalisé sous forme d'un complément direct ou indirect, est susceptible de jouer, sur le plan aspectuel, un rôle de terminus et/ou de mesure. Dans le premier cas, illustré ci-dessus par l'exemple (6), le lieu dénoté par le complément *à l'hôtel* marque le terme inhérent de la situation, c'est-à-dire le lieu l_n atteint au moment t_n . Un argument-mesure, quant à lui, fournit une échelle permettant d'évaluer précisément la progression temporelle de la situation associée à la structure argumentale du prédicat. Dans la mesure où il s'agit le plus souvent d'un objet doté d'une étendue physique, il permet en plus de fixer le terminus. L'exemple (9) emprunté à Tenney (1994 : 17) montre qu'un complément peut cumuler les deux rôles aspectuels : la linéarité de l'objet *Appalachian Trail* permet à la fois de mesurer le déplacement effectué par Sue et d'identifier les coordonnées spatio-temporelles $t_n l_n$ de son terme.

(9) Sue walked Appalachian Trail.

Krifka (1998) formalise la relation entre les parties temporelles d'une situation et la trajectoire selon laquelle celle-ci progresse. Cette relation de mouvement possède les quatre propriétés suivantes :

1. La co-extension consiste dans la correspondance entre d'une part les parties initiale et finale de la situation, et les segments initial et final de la trajectoire.
2. La contiguïté exige que les sous-parties temporellement adjacentes de la situation correspondent aux segments spatialement adjacents de la trajectoire.
3. La relation de mouvement est une application surjective, ce qui signifie que toute partie de la situation trouve une partie correspondante de la trajectoire et inversement.
4. Le rapport entre la situation et la trajectoire a un caractère minimal, c'est-à-dire que l'événement commence à la partie initiale de la trajectoire et il se termine avec sa partie finale.

1.3 Cent mètres de trajectoire ?

La relation de mouvement définie par Krifka (1998) a une importance particulière pour notre réflexion dans la mesure où elle fournit des critères permettant de décider si un complément appelle une interprétation en termes de trajectoire.

Avant de revenir à nos exemples, remarquons qu'en postulant, pour le complément *Appalachian Trail* (exemple 9), un statut d'argument-mesure, Tenny (1995) insiste sur le rôle de nos connaissances extralinguistiques : nous savons, en réalité, que traverser le Sentier d'Appalaches consiste à le parcourir en longueur plutôt qu'en largeur. Autrement dit, Tenny (1995) admet implicitement que les emplacements qui forment la trajectoire sont structurés selon un ordre tel que l'on pourrait les représenter sous forme d'une suite I_1, I_2, I_3, \dots . C'est justement cette contrainte de linéarité qui se trouve formalisée dans la relation de mouvement sous forme des quatre propriétés présentées ci-dessus.

Or, de toute évidence, nos exemples (1) à (4) échappent à la contrainte de linéarité. En effet, dans le cas de *j'ai nagé un kilomètre* (1) et dans celui de *on a couru le mille-mètres* (4), la relation entre les séquences temporelles qui forment la situation et l'étendue spatiale de l'objet n'a besoin, et très probablement n'est, ni co-extensive, ni minimale. Imaginons, en effet, un nageur faisant un kilomètre à la nage. Ni la signification de l'expression *un kilomètre*, ni notre expérience de la vie quotidienne ne permettent de s'attendre à ce que cette distance soit parcourue de manière linéaire, c'est-à-dire selon une trajectoire I_1, I_2, I_3, \dots . Bien au contraire, nous ne serons pas étonnés de voir le nageur faire plusieurs longueurs de piscine. Par conséquent, il n'y a pas de correspondance entre les parties temporelles successives de la situation et les parties adjacentes de cette prétendue trajectoire, puisque plusieurs séquences t correspondent à un même segment I de son parcours. La condition de co-extension n'est donc pas respectée. Remarquons que notre nageur peut ainsi terminer sa performance à l'endroit précis où il l'avait commencée, ce qui signifie que $t_0 \neq t_n$, mais $I_0 = I_n$ et exclut une relation minimale.

Le caractère subjectif de l'application ainsi que la contiguïté peuvent éventuellement être préservés à condition d'accepter l'idée qu'une même partie de la « trajectoire » correspond à plus d'une partie de la situation. Ces entorses nous semblent cependant coûteuses sur le plan conceptuel et surtout créées *ad hoc* pour sauver coûte que coûte une interprétation en termes de trajectoire. Or, pour saisir l'apport aspectuel des compléments qui nous intéressent, il est important de savoir s'il existe de bonnes raisons pour la préserver, en admettant par là même que ces expressions sont en mesure d'assumer les rôles aspectuels de mesure et de terminus de la manière exacte dont Tenny (1994, 1996) le suggère.

Si, comme nous le pensons, la relation de mouvement n'est pas préservée dans les exemples (1) à (4), il convient plutôt d'abandonner l'interprétation de nos compléments en termes de trajectoire sans toutefois nier leur importance sur le plan aspectuel. Débarrassés d'une notion susceptible de semer quelque confusion, nous avons de meilleures chances de saisir leurs spécificités.

1.4 Bornage et télécité

Avant de proposer une analyse de nos exemples (1) – (4) qui permettra de formuler une interprétation plus adéquate que celle qui ne se fait en termes de trajectoire, il convient de préciser notre position au sujet de différents types de bornes, c'est-à-dire d'explicitier la ligne de démarcation que nous dessinons entre le bornage et la télécité.

Il n'est pas rare parmi les aspectologues de traiter indifféremment le bornage, c'est-à-dire *gross modo* la délimitation quantitative, et la télécité. Ainsi, à titre d'exemple, Verkuyl (1993, 1999) et Verkuyl et al. (2004) préfèrent-ils parler de terminativité [\pm TERMINATIVITÉ], propriété complexe qui est le résultat d'un calcul opéré sur le plan de la prédication. On prend en compte le caractère dynamique ou non [\pm ADDTO] du verbe relevant de son aspect lexical et l'apport de syntagmes nominaux désignant des entités x dont la quantité peut être spécifiée [$+$ SQA]⁹ ou indéterminée [$-$ SQA]. En totalisant les informations aspectuelles apportées par les éléments de la prédication *Chantal* [NP + SQA] *tenir une lettre à la main* [VP | V – ADDTO] [NP + SQA] (Verkuyl et al. 2004 : 238-9), nous concluons à une situation non terminative.

En appliquant cet algorithme à nos exemples, dans tous les quatre cas on obtient la même formule [NP + SQA] + [VP | V + ADDTO] [NP + SQA] qui caractérise les situations terminatives. Il conviendrait alors d'admettre que les phrases (1) à (4) portent toutes le trait [$+$ TERMINATIVITÉ].

Dans certains travaux, la distinction entre la télécité et la délimitation est gommée sans être explicitement rejetée. En suivant Vendler (1967), de nombreux auteurs utilisent comme indice de télécité la compatibilité de la prédication avec l'adverbiale temporelle *en X temps*. Caudal (2006) établit qu'un énoncé « décrit une éventualité¹⁰ téléique s'il peut se combiner avec *en* [les italiques sont de notre fait] sans recevoir une interprétation inchoative ». La possibilité d'insérer cet adverbiale dans un énoncé sans l'effet interprétatif susmentionné est donc considérée comme caractéristique quasiment définitoire de télécité. Or, si l'on pense que toute situa-

⁹ La formule [$+$ SQA] se lit *specified quantity of x*.

¹⁰ Situation selon notre choix terminologique.

tion désignée par une phrase contenant l'adverbiale en *X* temps est télique, rien ne justifie de parler en plus de bornage, puisque celui-ci est nécessairement synonyme de télicité.

À l'opposé, Dahl (1981), Declerck (1979, 1980) et Depraetere (1995) adoptent la position consistant à penser que tout bornage n'est pas télique. En nous inspirant¹¹ de l'étude de Depraetere (1995), nous proposons de saisir les différences entre les deux notions de la manière suivante :

Définition 1 : Une situation est télique, lorsque sa structure interne possède une partie finale qui constitue le terme au-delà duquel elle ne peut plus continuer. Dans la mesure où un type de situations (ipse eo un type de structure situationnelle) est associé à la structure argumentale d'un prédicat, nous parlerons par extension de prédicat télique ou, au contraire, atélique.

Définition 2 : Une situation est bornée, lorsqu'elle couvre un intervalle fermé par une borne temporelle ou spatiale finale qui ne fait pas partie de sa structure interne.

Reprenons l'exemple (6) discuté dans la section 1.2 pour montrer qu'il illustre bien notre définition de télicité.

- (6) Nous sommes restés une heure à la grotte, après quoi, nous avons couru à l'hôtel, j'ai étendu mon python sur le lit et j'ai attendu. (Frantext, Romain Gary, *Gros-Célin*, 1974)

Dans (6), le verbe courir est employé comme un prédicat bi-argumental ($x = je$, y hôtel) dont l'argument interne fixe le terme du déplacement. Autrement dit, il assume un rôle aspectuel qui, conformément à Tenny (1994, 1995), incombe à un argument fermés. En plus de désigner le terme inhérent de la situation, il marque le passage vers un état résultant du déplacement qui consiste, pour x (je), à être à y (à l'hôtel).

Dans (10), au contraire, courir forme un prédicat monocargumental qui renvoie à une situation atélique.

- (10) J'ai couru 45 minutes aujourd'hui, mais j'ai eu une crampe atroce au ventre. (www.aujourd'hui.com)

Certes, l'occurrence particulière de cette situation décrite par (10) est présentée comme contenue dans un intervalle temporel de quarante cinq minutes. La borne finale qui correspond à la fin de la quarante cinquième minute, n'a pas un caractère télique pour autant, puisqu'elle ne fait pas partie de la structure situationnelle. La comparaison avec (6) est très instructive à cet égard. En effet, dans un cas comme

¹¹Depraetere (1995: 3) propose les définitions suivantes : « A clause is telic if the situation is described as having a natural (...) or an intended endpoint (...) which has to be reached for the situation as it is described in the sentence to be complete and beyond which it cannot continue ». Pour le bornage, nous lisons : « A sentence is bounded if it represents a situation as having reached a temporal boundary, irrespective of whether the situation has an intended or inherent endpoint or not. » Faut de place, nous ne pouvons développer la discussion de l'approche de Depraetere.

(10), il est possible d'étendre l'intervalle, en ajoutant des adverbiaux temporels sans que cette manipulation conduise à une lecture itérative¹² :

(11) J'ai couru 45 minutes, 50 minutes, une heure, mais j'ai eu une crampe atroce au ventre,

alors que la même manipulation sur un prédicat télique conduit à une référence plurielle :

(12) J'ai couru à l'hôtel, à la poste, à la gare.

Contrairement à Depraetere (1995) qui considère qu'aussi bien la télicité que le bornage sont des propriétés des phrases (sentences), nous pensons que les deux notions n'ont pas la même nature. Elles se manifestent effectivement sur le plan phrastique, mais relèvent de niveaux d'analyse différents. La télicité est une propriété structurelle qui caractérise certains types de situations. Certes, il se peut qu'une réalisation particulière s'arrête avant que cette partie finale soit atteinte, mais ce fait n'annule pas la présence du terminus dans la structure situationnelle. Une situation télique a donc un caractère téléonomique, c'est-à-dire qu'elle progresse vers son terme inhérent.

Le bornage (ou délimitation) tel que nous venons de le définir est une opération de prélevement quantitatif effectuée sur le plan de l'énoncé pour rendre les caractéristiques non structurelles d'une réalisation particulière de situation atélique. Nous verrons dans la section 2, que loin d'être uniquement un construit conceptuel, la différence entre la télicité et le bornage se traduit dans des faits linguistiques.

1.5 Bilan de la section

Cette première étape de notre réflexion a permis d'établir que les compléments des verbes *nager* et *courir* dans les exemples (1) à (4) ne s'interprètent pas en termes de trajectoire, contrairement à ce que suggère le cas de (5) avancé par Capelle et Declercq (2005). En écartant la lecture qu'ils en proposent, nous reprenons à notre compte l'idée de rôles aspectuels assumés par ces compléments. C'est afin de le circonscrire avec précision que nous nous sommes arrêtée, dans cette section, sur les notions aspectuelles dont nous nous servirons par la suite pour interpréter leurs propriétés grammaticales.

2. Propriétés grammaticales des compléments de distance

Dans la mesure où, comme nous l'avons indiqué dans la section 1.3, les compléments un kilomètre dans l'exemple (1), cent mètres dans (2), ainsi que le deux-cent mètres et le mille-mètres dans (3) et (4) n'expriment pas la trajectoire, nous proposons de les qualifier temporairement de compléments de distance. Ce terme, volontairement vague, est destiné à évoquer l'ensemble de ces expressions jusqu'au moment où nous disposerons d'éléments permettant d'en donner une caractérisation plus exhaustive.

Dans la présente section, nous nous concentrerons sur un certain nombre de faits linguistiques qui semblent converger pour confirmer la distinction qu'il convient

¹²L'hypothèse que nous avons avancée ailleurs pour expliquer ce phénomène consiste à penser que le bornage effectué sur des situations atéliques préserve leur structure homogène (pour un argumentaire plus détaillé, voir Sikora 2009).

d'établir entre les compléments de distance des exemples (1), (2) d'un côté, (3) et (4) de l'autre. Les propriétés que nous examinons se répartissent en deux catégories. Celles que nous discutons dans la section 2.1 s'accordent pour indiquer que, malgré les similarités de forme – dans les deux cas, il s'agit de syntagmes nominaux – les compléments en question possèdent des structures différentes, et par conséquent, n'assument pas les mêmes fonctions auprès du prédicat. Leurs propriétés transformationnelles discutées dans la section 2.2 montrent à leur tour que ces différences fonctionnelles sont décisives pour les rôles assumés par ces compléments de distance sur le plan aspectuel.

2.1 Ressemblance formelle et différences fonctionnelles

2.1.1 Structure de syntagmes nominaux

Force est d'admettre que les compléments de distance de nos exemples (1) à (4) présentent tous un (faux) air de famille qui joue certainement en faveur de leur rapprochement. En effet, dans les deux cas, nous avons affaire à des syntagmes nominaux composés d'un adjectif et d'un nom dénotant une unité de mesure. Un regard plus attentif permet néanmoins de déceler des différences frappantes.

De fait, dans les SN *un kilomètre dans* (1) et *cent mètres dans* (2), les adjectifs numériques quantifient l'unité de mesure dénotée par le nom. L'analyse de relations de dépendances syntaxiques permet de constater que ces SN s'analysent en termes de gouverneurs que constituent les noms *kilomètre* et *mètre* et d'adjectifs, en l'occurrence numériques, qu'ils régissent. Signalons cependant que dans le cas de *un kilomètre* une certaine ambiguïté brouille les pistes dans la mesure où un peu s'interpréter non seulement comme un numéral, mais également comme un article indéfini. Pour lever le doute, il suffit de procéder à un jeu de substitution en contexte : il est, en effet, possible de remplacer *un* par un autre adjectif numéral – mettons *deux*, alors qu'un article défini n'est pas permis¹³.

Si les SN *le deux-cents-mètres* et *le mille-mètres* comportent également un adjectif numéral et un nom, une analyse en termes de gouverneur nominal et de gouverné adjectival n'est pas adéquate. L'adjectif numéral et le substantif forment en réalité un nom composé forgé conformément aux principes dérivationnels du français qui ont produit, entre autres, *mille-pattes*, *deux-roues* et *trois-pièces*. Quant à son contenu sémantique, nous pensons qu'il lexicalise un sens différent de celui de *kilomètre* ou de *mètre*. Certes, une analyse sémantique profonde pourra faire ressortir le concept d'unité de mesure présent à la base¹⁴, mais interrogés sur la signification du *deux-cents-mètres* ou du *mille-mètres*, nous répondrons dans un cas comme dans l'autre, qu'il s'agit d'épreuve sportive réalisée sur une distance *d*.

2.1.2 Modification adjectivale

Les compléments de distance qui nous occupent ici présentent des divergences notables en ce qui concerne leur rapport aux adjectifs susceptibles de les qualifier.

L'exemple (13) repéré sur internet pointe vers la possibilité de modification adjectivale que l'on peut appliquer aux SN du type exemplifiés en (3) et (4), alors que ceux que nous trouvons en (1) et (2) s'y prêtent nettement moins bien.

¹³Nous ne discutons pas ici de la possibilité de l'article indéfini pluriel que l'on interprétera comme plusieurs.

¹⁴Je remercie Denis Apothélos d'avoir attiré mon attention sur cette composante que l'on pourrait appeler, selon sa proposition, *unité seconde*.

- (13) Bolt a couru le 200 mètres le plus rapide de l'année en s'imposant en 19⁰⁶ dimanche à Athènes. (<http://www.marathons.fr>)

Dans les versions c. de nos exemples, nous avons essayé de modifier le SN par un adjectif au superlatif. Le contraste entre (1) et (2) d'une part, (3) et (4) de l'autre est frappant.

- (1c) ? J'ai nagé un kilomètre le plus rapide de l'année et espère m'être musclée un petit peu... en tous cas, je me sens bien fatiguée!
- (2c) ? J'ai couru 100 mètres les plus rapides de ma vie avec trois types qui étaient sortis d'un bar et m'ont suivi pour me demander ce que je fais.
- (3c) Cathy a nagé le 200 mètres papillon le plus rapide de sa carrière.
- (4c) On a commencé par un peu d'échauffement, puis on a couru le 1000 mètres le plus rapide de l'année.

De toute évidence, les compléments de distance un kilomètre et cent mètres ne régissent pas à la modification adjectivale à la façon propre aux nominaux. Même si l'on admet, pour le premier, une lecture qui consiste à traiter un comme article indéfini susceptible d'être remplacé par un défini, *j'ai nagé le kilomètre le plus rapide de l'année*, bien qu'acceptable, reste moins bon que la phrase (3c). Les possibilités qu'offre la seconde des deux interprétations de un kilomètre indiquent que, dans le cas de nos exemples, la modification adjectivale nécessite une nominalisation fonctionnelle du SN. Elle permet également de constater a contrario que lorsque l'on interprète un dans (1), tout comme cent (2) comme adjectifs numériques, le comportement des compléments n'est pas pleinement nominal, puisqu'il exclut la modification par adjectifs.

Dans les exemples (3) et (4), au contraire, le deux-cents-mètres et le mille-mètres se prêtent sans résistance à l'adjonction du superlatif et à la modification adjectivale¹⁵ en général, en faisant preuve de leur statut pleinement nominal du point de vue non seulement formel, mais également fonctionnel.

2.1.3 Passivation

Si les compléments de distance le deux cents mètres dans (3) et le mille mètres dans (4) ont un statut pleinement nominal, les verbes nager et courir sont, dans ces deux phrases, transitifs. Rien d'étonnant dès lors qu'ils acceptent la passivation, aussi bien dans une forme périphrastique que pronominale. Les énoncés (3d) et (4d) sont des versions passives des exemples correspondants :

- (3d) Le 200 mètres papillon a été nagé par Cathy.

- (4d) Le 1000 mètres a été couru (par nous tous).

¹⁵Pensez, par exemple, à *J'ai nagé un deux-cents mètres époustouffant*. Une lecture pleinement nominale de un kilomètre autoriserait éventuellement *J'ai nagé un kilomètre époustouffant*, mais *J'ai couru cent mètres époustouffants* semble problématique.

Employés au passif périphrastique, *nager* et *courir* ont une lecture occurrence-tielle, puisqu'ils renvoient à une réalisation particulière de la situation par un ou plusieurs individus. L'on sait néanmoins que l'expression de l'agent est optionnelle dans le cas du passif périphrastique et qu'elle est proscrite dans la forme pronominale¹⁶. Néanmoins, en accordant à nos exemples une localisation temporelle et/ou spatiale, il est possible de former à partir de (3) et (4) un tel passif présentant à lecture occurrence-tielle même en l'absence d'un agent :

(3e) Le 200 mètres papillon s'est nagé hier à la piscine olympique.

(4e) Le 1000 mètres s'est couru hier au stade municipal.

Par ailleurs, il n'est pas rare de trouver, pour les emplois transitifs de *nager* et de *courir* avec des compléments de distance de type (3) – (4), des passifs pronominaux à lecture itérative de type habituel. (14) et (15) en sont une illustration :

(14a) En natation de compétition, une simple lecture des résultats obtenus lors des jeux olympiques depuis 1908 permet de constater que le 1500 mètres se nage aujourd'hui à une vitesse plus élevée que se nageait le 100 mètres en 1920. (<http://www.mont-devo-natation.org>)

(15a) Tout comme le 100 m haies et le 110 m haies, le 100 m se court en ligne droite. (<http://antony.athle.free.fr>)

Les énoncés (14) et (15) signifient que les épreuves de mille-cinq-cents-mètres et cent-mètres se déroulent habituellement selon certaines modalités qui leur sont propres (une vitesse particulière, structuration du parcours). Or, comme l'explique Carlier (2002 : 54), dans les phrases itératives « [l]es différentes occurrences du procès étant envisagées comme qualitativement identiques, il n'y a pas, d'une occurrence à l'autre, évolution. En cela, la valeur aspectuelle globale, obtenue au niveau de la série d'occurrences processives, se rapproche également de l'état. » Par conséquent, nos exemples (14) et (15) peuvent être paraphrasés, du moins dans les parties qui nous intéressent, de la manière suivante : *le mille-cinq-cents-mètres a la propriété d'être nagé aujourd'hui à une certaine vitesse, le cent-mètres a la propriété d'être couru en ligne droite*. On constate ainsi que le passif pronominal des exemples (14) et (15) nous amène à une lecture normative, voire générique, puisque les propriétés en question sont attribuées à toute entité qui relève de la même classe d'objets.

Il est important de remarquer que la lecture générique, ou du moins gnominique, de ces passifs pronominaux met en évidence le caractère pleinement nominal des

¹⁶C'est la raison pour laquelle Creissels (2006) considère que les formes *se-V* sont simplement proches du passif, « mais à la différence d'un véritable passif, ou bien elles suggèrent que le déclenchement du processus ne met en jeu aucune force identifiable (*La porte s'ouvre, Le tissu s'use*), ou bien elles minimisent l'intervention d'une force extérieure en insistant sur une prédisposition particulière du référent du sujet à subir le procès (*Ce tissu se lisse facilement*) » (Creissels, 2006 : 31). L'auteur reconnaît néanmoins que les phrases telles que *Le vin blanc se doit frais*, paraphrasable par *On boit le vin blanc frais*, sont caractéristiques des emplois de type passif. Dans la mesure où les exemples (14) et (15) semblent se ranger dans la même catégorie, nous avons choisi de les traiter en termes de passif pronominal.

compléments de distance de type (3) et (4). Bien plus qu'une forme de SN, ils possèdent également une caractéristique fonctionnelle propre aux nominaux : celle de dénoter des classes d'objets.

Le contraste avec un kilomètre et cent mètres, et plus généralement, avec les compléments de type (1) et (2), est frappant et, à notre avis, lourd en conséquences à tirer. Force est de commencer par un constat d'échec : nous n'avons trouvé aucune attestation de passif périphrastique ou pronominal formé à partir d'emplois de type (1) et (2). Il est par ailleurs difficile, voire impossible, d'en fabriquer pour les besoins de notre démonstration, comme le montrent les résultats de telles manipulations présentés sous (1d, e) et (2d, e) :

(1d) ? Un kilomètre a été nagé (par moi).

(1e) ? Un kilomètre se court en ligne droite.

(2d) ? 100 mètres ont été courus (par moi).

(2e) ?? 100 mètres se courent en ligne droite.

Si (1d) et (1e) paraissent légèrement moins bancals et, par conséquent, plus facilement acceptables, c'est grâce à l'ambiguïté due au déterminant un que nous avons signalé plus d'une fois depuis la section 2.1. En effet, si l'on interprète un comme article indéfini, le passif pronominal est admissible, tout comme il l'est pour les exemples de type (3) et (4). En témoignent les exemples (14b) et (15b) dans lesquels l'article un remplace le déterminant de des phrases originales (14a) et (15a) :

(14b) En natation de compétition, une simple lecture des résultats obtenus lors des jeux olympiques depuis 1908 permet de constater qu'un 1500 mètres se nage aujourd'hui à une vitesse plus élevée que se nageait un 100 mètres en 1920.

(15b) Tout comme un 100 m haies et un 110 m haies, un 100 m se court en ligne droite.

En revanche, là où – comme dans le cas de (2) – le bénéfice du doute ne peut pas être accordé, aucune forme de passif, qu'il soit périphrastique ou pronominal, n'est acceptable.

L'impossibilité de transformation passive des exemples (1) et (2), dans lesquels les verbes nager et courir ont pourtant une apparence de construction transitive, trouve une explication dans le fait que les compléments un kilomètre et cent mètres n'ont pas un statut pleinement nominal. Formellement, ce sont, certes des SN, mais ils semblent incapables d'assumer des fonctions propres aux nominaux, notamment celle de dénoter une entité susceptible de subir l'action exercée par un agent, comme cela a lieu dans le passif périphrastique, ou des classes d'entités dont on pourrait affirmer des propriétés génériques à l'aide, entre autres, de passif pronominal.

2.1.4 Bilan de la section 2.1

Tout au long de la section 2.1, nous avons tenté de réunir les faits linguistiques qui indiquent que les compléments de distance tels que un kilomètre dans (1) et cent mètres dans (2), malgré une forme de SN, n'ont pas un comportement nominal : ils

n'admettent pas la modification adjectivale, ils refusent la passivation. Quelle est donc leur véritable nature ?

Pour répondre à cette question, revenons à l'exemple (10) que nous avons évoqué pour illustrer le phénomène de boenage défini dans la section 1.4.

- (10) J'ai couru 45 minutes aujourd'hui, mais j'ai eu une crampe atroce au ventre.
(www.aujourd'hui.com)

Le complément *quarante cinq minutes* est un circonstant délimitant la séquence temporelle couverte par la course. En d'autres termes, c'est une expression quantifiante à caractère adverbial. On pourrait donc la rapprocher fonctionnellement des adverbes d'évaluation quantitative¹⁷ tels que *beaucoup*, *peu*, etc., à une différence près : le modifieur adverbial *quarante cinq minutes* impose des bornes temporelles à la séquence de course, ce qui n'est pas le cas de *beaucoup* ni de *peu*.

Notre hypothèse quant au statut des compléments de distance des exemples (1) et (2) consiste donc à considérer qu'il s'agit d'adverbiaux quantifiants qui délimitent une séquence de nage ou de course, en imposant des bornes à caractère spatial.

Le comportement nominal des compléments *le deux-cents-mètres* et *le mille-mètres* porte à croire, au contraire, que ce sont des réalisations syntaxiques des arguments internes des prédicats *nager* et *courir* dans les exemples (3) et (4).

Nous verrons, dans la section 2.2, que cette différence de statut est importante pour les structures des situations désignées dans nos exemples (1) et (2) d'un côté, (3) et (4) de l'autre.

2.2 Caractéristiques des structures situationnelles

Dans cette section, nous proposons de soumettre nos exemples (1) à (4) à des transformations qui mettent en lumière les propriétés des situations associées aux prédicats que forment *nager* et *courir* en présence, respectivement, de modifieurs adverbiaux et d'arguments internes.

2.2.1 Être + participe passé à valeur résultative

L'exemple (16) prélevé sur internet fournit un indice important quant aux structures situationnelles qui nous occupent. Il est à noter que la construction *est couru* dénote un état résultant sur lequel débouche la situation consistant à courir le quinté.

- (16) À mon avis, tu as dû bien fêter la victoire des bleus cette nuit. Tu postes ton message à 15h30, alors que le quinté est couru depuis 1h30 hihhi... (forum.lixium.fr)

En ce qui concerne nos exemples, seuls (3) et (4) acceptent la construction en *être + participe passé à valeur résultative*. Imaginons, en effet, qu'un interlocuteur peu au fait du calendrier sportif, pose les questions (17) et (18) :

- (17) Quand Cathy doit-elle nager le deux-cents-mètres papillon ?

- (18) Quand courra-t-on le mille-mètres ?

¹⁷Pour une discussion d'évaluation quantitative dans le domaine verbal, voir, entre autres, Boillat (1989).

Il serait alors tout à fait envisageable d'y répondre respectivement par (3f) et (4f) :

(3f) C'est fait, le deux-cents-mètres papillon est nagé depuis deux heures.

(4f) C'est fait, le mille-mètres est couru depuis deux heures.

Or, s'il est possible, dans des circonstances strictement similaires, d'émettre les mêmes interrogations relativement aux exemples (1) et (2), la réponse ne supportera pas de forme résultative :

(1f) ? Un kilomètre est couru depuis deux heures.

(2f) ? Deux-cents mètres sont courus depuis deux heures.

Il en ressort que les verbes *nager* et *courir* donnent lieu à une construction résultative, lorsqu'ils forment des prédicats bi-argumentaux, c'est-à-dire lorsqu'ils sélectionnent un argument interne. Or, la possibilité de forger une forme à valeur résultative, fortement contrainte comme l'indique Creissels (2000), est hautement révélatrice des propriétés aspectuelles liées à la structure situationnelle. Ainsi, « l'interprétation résultative de être + participe passé n'est possible qu'en liaison avec une interprétation téléque du sens lexical verbal » (Creissels, 2000 : 141)¹⁸. Précisons qu'en ce qui concerne les verbes *courir* et *nager*, il serait problématique de parler d'une interprétation téléque du sens lexical du verbe. C'est les prédicats qu'ils forment avec leurs arguments internes respectifs le deux-cents-mètres et le mille-mètres, ipse co leurs structures argumentales, qui se voient associer des situations téléques.

2.2.2 Les adverbiaux en entier et en partie

Il est courant, en aspectologie, de se servir d'adverbes complètement et partiellement pour vérifier si une situation possède une structure atomique, c'est-à-dire si elle forme un bloc impossible à scinder.

Pour notre part, il est clair que les situations associées aux prédicats *nager* et *courir* dans nos exemples n'ont pas le trait d'atomicité. Cependant, il est intéressant de remarquer qu'ils réagissent différemment à ce type d'adverbiaux, comme le montrent les énoncés (1g) et (2g) d'un côté, (3g) et (4g) de l'autre. Précisons que nous avons choisi de remplacer complètement et partiellement couramment utilisés (cf. à titre d'exemple, Caudal 2006), qui s'accordent mal avec les verbes et les compléments qui nous intéressent¹⁹, par les locutions en entier et en partie qui paraissent véhiculer un contenu sémantique semblable, tout en présentant l'avantage de mieux s'insérer dans les contextes incriminés. Regardons donc les effets que produit un tel ajout sur nos exemples.

¹⁸Précisons que cette contrainte n'est pas spécifique au français. Hyman et Dahl (1980) indiquent que le résultatif va de pair avec la téléité dans de nombreuses langues.

¹⁹De même, une situation non atomique désignée par *J'ai lu ce livre* accepte plutôt les adverbiaux en entier/en partie que complètement/partiellement. Il suffit de comparer *J'ai lu ce livre en entier/en partie* et *J'ai lu ce livre complètement/partiellement*.

- (1g) ? J'ai nagé un kilomètre en entier/en partie et espère m'être musclée un petit peu... en tous cas, je me sens bien fatiguée!
- (2g) ? J'ai couru 100 mètres en entier/en partie avec trois types qui étaient sortis d'un bar et m'ont suivi pour me demander ce que je fais.
- (3g) Cathy a nagé le 200 mètres papillon en partie/en entier.
- (4g) On a commencé par un peu d'échauffement, puis on a couru le 1000 mètres en entier/en partie.

Ces adverbiaux mettent en évidence la possibilité de distinguer des étapes ou des stades dans le déroulement d'une situation consistant, par exemple, à nager le deux-cents-mètres et à courir le mille-mètres. On conçoit facilement qu'un nageur peu entraîné ou un coureur en petite forme réalise l'épreuve en partie, sans la terminer, c'est-à-dire sans atteindre le terme de la situation. Cependant la présence de celui-ci dans la structure situationnelle n'en est pas remise en cause. En revanche, l'insertion problématique de *en entier* et de *en partie* dans les exemples (1) et (2) est révélatrice des fonctions assumées par un kilomètre et cent mètres. Quelqu'un qui aurait nagé un kilomètre ou couru cent mètres en partie, aurait tout simplement parcouru une distance plus courte que celle spécifiée par le complément²⁰. Il serait étrange de prétendre alors que x a couru un kilomètre en partie ou cent mètres en partie. Dire, au contraire, qu'on a parcouru un kilomètre ou cent mètres en entier serait pour le moins péonastique.

L'incompatibilité de (1) et (2) avec les adverbiaux *en partie* et *en entier* confirme notre hypothèse au sujet de la fonction qu'assument un kilomètre et cent mètres supéris des prédicats *nager* et *courir*. L'impossibilité de scinder les occurrences de situation désignées par (1) et (2) en stades réalisés en partie et la redondance de *en entier* indiquent que ces compléments sont des circonstants qui attribuent une propriété caractérisant non pas un type de situation, mais l'une de ses réalisations particulières qui se déroule dans un *hic et nunc*. Il s'agit donc de modifieurs adverbiaux dont le rôle consiste à délimiter l'étendue spatiale correspondant à une occurrence de situation.

2.2.3 L'adverbial temporel en X temps

La fiabilité du test aspectuel consistant à insérer, dans une structure prédicative, l'adverbial temporel en X temps est loin de faire l'unanimité des aspectologues. Si pour certains, comme nous l'avons indiqué dans la section 1, la compatibilité avec cette expression est une preuve de télicité, d'autres ne cachent pas leurs réticences, notamment en ce qui concerne son fonctionnement dans des systèmes linguistiques autres que le français ou l'anglais (cf. Verkuyl et al. 2004). Melendjik et al. (2004) considèrent, quant à eux, que loin d'être de simples tests, ces adverbiaux jouent un rôle actif dans la dérivation aspectuelle.

Revenons à présent à nos exemples (1) à (4) pour vérifier leurs réactions à l'adverbial en X temps.

²⁰A contrario, une fois que x a couru cent mètres, la séquence peut être rallongée. Il est possible de dire : x a couru cent mètres, deux cents mètres, un kilomètre avant de tomber de fatigue (cf. exemple 11 discuté dans la section 1.4).

- (1h) J'ai nagé un kilomètre en 45 minutes et espère m'être musclé un petit peu...
en tous cas, je me sens bien fatigué!
- (2h) J'ai couru 100 mètres en douze secondes avec trois types qui étaient sortis
d'un bar et m'ont suivi pour me demander ce que je fais.
- (3h) Cathy a nagé le 200 mètres papillon en 2 minutes 30 secondes (dans une eau
glaciale).
- (4h) On a commencé par un peu d'échauffement, puis on a couru le 1000 mètres
en 3 minutes 30 secondes.

Il ne fait pas de doute que tous ces énoncés sont bien formés, ce qui signifie qu'aussi bien les modificateurs que les arguments internes autorisent les adverbiaux en *X temps*. Pour les chercheurs qui confèrent à cette expression une valeur de test de télécité, (1h), (2h), (3h) et (4h) sont des descriptions de situations téliques. Or, nous avons constaté précédemment qu'une interprétation télique des exemples (1) et (2) devait être écartée notamment en vertu des propriétés grammaticales que nous avons étudiées. Si ce raisonnement est correct, nous sommes amenée à nous interroger sur la valeur du test lui-même. Autrement dit, la question à poser est celle de l'interprétation que l'on doit accorder aux résultats de son application.

Si tous les quatre énoncés sont compatibles avec *en X temps*, il convient de remarquer que l'adverbial ne produit pas toujours le même effet. Ainsi, dans le cas des exemples (1) et (2), on perçoit une différence de signification nette entre les versions a. et g. Loins d'un simple rajout d'information, l'insertion de l'adverbial remodèle la signification de ces phrases en créant un effet d'intensité : la séquence de nage dans (1h) ainsi que la course en (2h) apparaissent comme plus éprouvantes, plus difficiles à exécuter que celles des exemples (1a) et (2a). On en conclut que l'adverbial *en X temps* ajoute une circonstance qui demande un effort supplémentaire de la part de la nageuse ou du coureur. En nageant un kilomètre à son rythme, on peut espérer s'être musclé, mais une séquence de nage de la même longueur dans un temps limité est une performance encore plus favorable pour le développement musculaire, car elle décuple les efforts déployés. De même, dans (2h), une course de cent mètres avec trois individus suspects est encore plus éreintante, lorsqu'elle est exécutée dans un temps impliquant une vitesse très élevée.

En plus de spécifier l'intervalle inclusif couvert par une occurrence de situation, *en X temps* entre donc en interaction avec les autres circonstants de la phrase. Ensemble, ils remplissent une fonction propre aux adverbiaux qui, comme le remarque Bogusławski (2005), malgré leur diversité formelle, précisent le mode du déroulement de la situation désignée par le verbe, ses propriétés, autrement dit le comment de cet événement. C'est dire que, dans les exemples (1) et (2), un kilomètre, cent mètres et *en X temps* ont le même statut d'adverbiaux circonstants, ipse eo celui de modificateurs de descriptions événementielles : ils spécifient les propriétés de l'occurrence de situation plus que celles de la situation en tant que type²¹. On pourrait dès lors visualiser la structure de la prédication (1) et (2) de la manière suivante :

²¹Cf. Malendjk et al. (2006).



Figure 1 : Structure de la prédication dans les exemples (1h) et (2h)

Dans le cas de (3h) et de (4h), l'analyse semble plus simple. Dans la section précédente, nous avons montré que *le deux-cents-mètres* dans (3) et *le mille-mètres* dans (4) sont des arguments internes. C'est la raison pour laquelle dans l'exemple (3h), nous avons proposé, entre parenthèses, un circonstant (*dans une eau glaciale*) susceptible d'entrer en interaction avec *en X temps*.

Or, quelle que soit la configuration – avec ou sans circonstant – *en X temps* ne fait qu'ajouter une spécification temporelle de l'intervalle inclusif qui couvre l'occurrence situationnelle. En cas de circonstant supplémentaire en (3h), l'information semble s'ajouter, de manière séquentielle, comme nous avons tenté de le représenter dans la figure (2) :

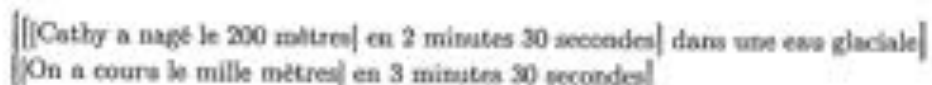


Figure 2 : Structure de la prédication dans les exemples (3h) et (4h).

2.2.4 Bilan de la section 2.2

Les faits linguistiques étudiés dans la section (2.2) conduisent à admettre qu'un modifieur (*un kilomètre* de 1, *cent mètres* de 2) et un argument interne (*le deux-cents-mètres* et *le mille-mètres* des exemples 3 et 4) n'ont pas le même poids pour la structure des situations associées au prédicat. Les premiers, c'est-à-dire les modifieurs adverbiaux, ne confèrent pas à la situation un caractère télique, mais ils opèrent un bornage. Les seconds, au contraire, sont porteurs du trait [+télicité]. Cela signifie qu'ils participent à la dérivation aspectuelle qui transforme une activité en accomplissement.

3. Compléments et/ou rôles aspectuels ?

L'étape suivante de notre réflexion consiste à poser la question des rôles aspectuels qu'assumeraient les compléments dans les deux types d'exemples étudiés. Cette question en comprend en fait deux autres qui nécessitent un traitement et des réponses séparés :

1. Peut-on parler, pour les cas qui nous intéressent dans cet article, de compléments aspectuels ?
2. Peut-on parler, pour les cas qui nous intéressent dans cet article, de rôles aspectuels ?

S'agissant de la première question, les faits de langue soumis à l'analyse dans les sections précédentes conduisent à admettre que les expressions *un kilomètre*, *cent mètres*, *le deux-cents-mètres* et *le mille-mètres* sont bel et bien des compléments aspectuels dans la mesure où elles confèrent certaines propriétés aspectuelles aux situations décrites dans les phrases (1) à (4). Cependant, en analysant leurs contributions respectives, on s'aperçoit qu'elles sont de natures fort différentes.

Les modificateurs adverbiaux *un kilomètre* et *cent mètres* dans (1) et (2) opèrent au niveau de la prédication, en fixant des bornes non téléiques. Autrement dit, il s'agit d'expressions qui servent à réaliser un prélèvement quantitatif par spécification de l'étendue qui caractérise non pas un type de situations, mais l'une de leurs réalisations particulières. On comprend dès lors pourquoi une telle opération ne débouche pas sur un état résultant (section 2.2.1) et ne peut pas être exécutée partiellement (section 2.2.2).

La réponse à la seconde question, celle de leur rôle aspectuel, est plus nuancée. Si ces expressions sont, comme nous le pensons, des modificateurs adverbiaux qui s'analysent syntaxiquement en termes de circonstants, il semble hautement problématique de leur accorder un rôle aspectuel. Rappelons, en effet, que ce concept est construit par Tenny (1994, 1995) sur la base de celui de rôles thématiques portés uniquement par les arguments sémantiques. Par conséquent, il serait plus judicieux de parler non pas de rôle, mais de fonction aspectuelle qu'assument les compléments tels que *un kilomètre* et *cent mètres* dans nos exemples (1) et (2).

Dans le cas du *deux-cents-mètres* et dans celui du *mille-mètres* (exemples 3 et 4), nous répondons à la première question par l'affirmative. En effet, il s'agit également de compléments aspectuels, mais à la différence des précédents, ces expressions sont des réalisations syntaxiques d'arguments internes du prédicat. Par conséquent, en répondant à la seconde question, il n'est pas impossible de caractériser, voire de modéliser, leur contribution aspectuelle en termes de rôles.

Le deux-cents-mètres dans (3) ainsi que *le mille-mètres* dans (4) assument en réalité un double rôle aspectuel. D'une part, en tant que porteurs du trait [+téléicité], ils fixent le terminus de la situation. De l'autre, ils permettent non pas de mesurer le déroulement de la situation, mais plutôt d'évaluer sa progression.

En effet, dans la section 1.3, nous avons été amenée à écarter l'interprétation de ces compléments en termes de trajectoire, en renonçant par là même à l'idée d'une échelle composée d'emplacements t_1, t_2, t_3, \dots sur laquelle on projetterait les valeurs successives t_1, t_2, t_3, \dots du temps interne de la situation. C'est la raison pour laquelle le rôle de mesure ne correspond pas à l'apport aspectuel de ces compléments-arguments.

4. Conclusions

En analysant les exemples (1) à (4), nous avons commencé par renoncer à l'interprétation de leurs compléments de distance en termes de trajectoire. Nous nous sommes ensuite appliquée à montrer que sous un air de famille formel, se cochaient des expressions aux propriétés fonctionnelles et aspectuelles différentes.

Le problème discuté dans cet article est, certes, spécifique à des verbes particuliers employés avec certains types de compléments. La distinction que nous avons dégagée entre les arguments et les modificateurs semble cependant avoir des implica-

tions plus générales, notamment en ce qui concerne deux thèmes d'importance en aspectologie. Premièrement, nous pensons que nos analyses fournissent des arguments en faveur de la distinction entre la télélicité et le bornage. En effet, les deux phénomènes relèvent bien de l'aspectualité, mais ils se situent sur des niveaux de calcul aspectuel distincts (structure situationnelle vs propriétés d'occurrences) et se manifestent selon des modes particuliers, comme nous avons essayé de le montrer dans la section 2.

Deuxièmement, les cas que nous avons étudiés confirment les objections formulées à l'encontre de l'adverbial en *X temps*, notamment en ce qui concerne son statut de test. Il doit être manipulé avec précaution, car si, comme les faits examinés semblent l'indiquer, il y a lieu de distinguer entre télélicité et bornage, en *X temps* est compatible avec les deux. De ce fait, il ne peut pas être considéré comme marque de télélicité, notamment parce que sa valeur inclusive s'accorde aussi bien avec un argument porteur de rôle aspectuel qu'avec un modifieur délimitatif. Il s'agit avant tout d'une expression de langue naturelle qui entre en interaction avec les autres éléments de la prédication pour créer une signification.

Références

- Aunargue, Michel (2008). *Qu'est-ce qu'un verbe de déplacement ? Critères spatiaux pour une classification des verbes de déplacement intransitifs du français*. Paris, Institut de Linguistique Française, CMLF'08.
- Aunargue, Michel (2011). *How motion verbs are spatial : The spatial foundations of intransitive motion verbs in French*. *Linguisticae Investigationes*, vol. 34, n°1, pages 1-34.
- Bogusławski, Andrzej (2005). *O operacjach przyimkowych*. In Grochowski, M. (Ed.) *Przyimki i przysłówki. Studia ze składni i semantyki języka polskiego*. Wydawnictwo Uniwersytetu Mikołaja Kopernika, Toruń, pages 15-44. (In Polish).
- Borillo, André (1989). *Notions de « massif » et de « comptable » dans la mesure temporelle*. In David, Jean & Kleiber, Georges. (Eds) *Termes massifs et termes comptables*, Klincksieck, Paris, pages 215-238.
- Borillo, André (1998). *L'espace et son expression en français*. Éditions Ophrys, Paris.
- Hybee, Joan L. & Dahl, Östen (1989). *The creation of tense and aspect systems in the languages of the world*, *Studies in Language* 13, pages 51-103.
- Cappelle, Bert & Declerck, Renaat (2005). *Spatial and temporal boundedness in English motion events*. *Journal of Pragmatics* 37, pages 889-917.
- Carlier, Anne (2002). *Les propriétés aspectuelles du passif*. *Cahiers Chronos* 10, pages 41-63.
- Caudal, Patrick (2000). *La polysémie aspectuelle - contraste français / anglais*. Thèse de doctorat, Université Paris VII - Denis Diderot.
- (2006). *Aspect*. In Godard, Danièle, Roussarie, Laurent & Corblin, Francis (Eds.), *Sémantico-pédagogie : dictionnaire de sémantique*, GDR Sémantique & Modélisation, CNRS, <http://www.semantique-gdr.net/dico/>
- Croisels, Denis (2000). *L'emploi résultatif de 'être + participe passé' en français*. *Cahiers Chronos* 6, pages 133-142.
- Croisels, Denis (2006). *Syntaxe générale. Une introduction typologique. Tome 2 : La phrase*. Hermes-Lavoisier, Paris.

- Dahl, Östen (1981). On the Telic/Atelic (Bounded/Nonbounded) Distinction. In Tedeschi, Philip J. & Zaenen, Annie (Eds.). *Tense and Aspect (Syntax and Semantics)*, vol. 14, pages 79-90.
- Declerck, Renaat (1979). *Aspect and the Bounded/Unbounded (telic/atelic) Distinction*. *Linguistics*, 17, pages 761-794.
- Declerck, Renaat (1989). *Boundedness and the Structure of Situations*. *Louvain Bijdragen*, 78, pages 275-308.
- Depraetere, Ine (1995). On the necessity of distinguishing between (un)boundedness and (a)telicity. *Linguistics and Philosophy*, vol. 18 n°1, pages 1-19.
- Dowty, David, R. (1979). *Word Meaning and Montague Grammar*. Dordrecht, Reidel.
- Dowty, David, R. (1991). Thematic Proto-Roles and Argument Selection. *Language*, 67, pages 547-619.
- Fillmore, Charles, J. (1983). How to know whether you're coming or going (Reprint). In Rauh, Gisa (Ed.), *Essays on Deixis*, Gunter Narr Verlag, Tübingen, pages 219-227.
- Krifka, Manfred (1992). Thematic Relations as Links between Nominal Reference and Temporal Constitution. In Sag, Ivan, A. & Scabolesi, Anne (Eds.). *Lexical Matters*, Stanford California, CSLI Publications, pages 29-54.
- (1998). The Origins of Telicity. In Rothstein, Susan (Ed.). *Events and Grammar*. Dordrecht, Kluwer, pages 197-235.
- Lagae, Véronique (2002). *Le passif pronominal : une forme complémentaire du passif périphrastique ?* *Cahier Chronos*, 10, pages 133-149.
- Lagae, Véronique (2005). *Les formes en être + participe passé à valeur résultative dans le système verbal français*. *Cahiers Chronos*, 12, pages 125-142.
- Molendijk, Aris, De Swart, Henriëtte, Vetter, Carl, Borillo, Andrés, Bras, Myriam, Le Draoulec, Anne, Vieu, Laure, Verkuyl, Henk, Vet. Co (2004). *Meaning and Use of Past Tenses in Discourse*. In *Handbook of French Semantics*. Corblin, Franck & De Swart, Henriëtte (Eds.). Stanford, CSLI, pages 271-308.
- Pourel, Stéphanie, Kopecka, Aneta. (2006). *Motion Events in French : Typological Intricacies*. www.sumex.ac.uk.
- Sikora, Dorota (2009). *Les Verbes de manière de mouvement en polonais et en français. Éléments pour une étude comparée des propriétés structurelles des prédicats*. Thèse de doctorat, Nancy, Université Nancy 2 - ATILF - CNRS (UMR 7118).
- Sikora, Dorota (& paraitre). *Le verbe courir et ses acceptions sportives*. *Études de linguistique appliquée*.
- Smith, Carlotta (1991). *The Parameter of Aspect*. Dordrecht, Boston, London, Kluwer.
- Talmy, Leonard (1987). *Lexicalization patterns: Typologies and universals*. *Berkeley CognitiveScience Report 47*.
- Talmy, Leonard (2000). *Toward a Cognitive Semantics. Volume I: Concept Structuring System. Volume II: Typology and Process in Concept Structuring*. MIT Press, Harvard.
- Tenny, Carol (1994). *Aspectual Roles and the Syntax-Semantics Interface*. Dordrecht, Boston, Londres, Kluwer Academic Publishers.
- Tenny, Carol (1995). *How Motion Verbs Are Special. The Interaction of Semantic and Pragmatic Information in Aspectual Verb Meaning*. *Pragmatics and Cognition*, 3, 1, pages 31-73.
- Vendler, Zeno (1967). *Verbs and Times*. *Linguistics in Philosophy*, pages 97-121.
- Verkuyl, Henk (1972). *On the Compositionnal Nature of Aspect*. Dordrecht, Reidel.
- Verkuyl, Henk (1993). *A Theory of Aspectuality. The Interaction between Temporal and Atemporal Structure*. *Cambridge Studies in Linguistics*, 64, Cambridge, Cambridge University Press.

Verkuyt, Henk (1990). *Aspectual Issues : Studies on Time and Quantity*. Stanford, CA, CSLI Publications.

Verkuyt, Henk, Vet, Co, Boello, Andr e, Bras, Myriam, Le Draoulec, Anne, Molendijk, Aris, De Swart, Henri tte, Vetter, Carl, Vies, Laure (2004). In Tense and Aspect in Sentences. Corblin, Francis & De Swart, Henri tte (Eds). *Handbook of French Semantics*, Stanford, CSLI Publications, pages 233-270.

Ressource  lectronique

www.frantext.fr

L'adverbial *en X temps*, est-il vraiment un test de télicité?*

Dorota Sikora

Université du Littoral — Côte d'Opale & HLLI (EA 3040)

Introduction

Il est courant en aspectologie de considérer que l'adverbial anglais *in X time* et son équivalent français *en X temps* sont des tests permettant de vérifier si un prédicat est télique ou non. En dépit d'objections récurrentes dans la littérature (cf. Verkuyl et *all.* 2004, Molendjik, 2004), ces adverbiaux bénéficient toujours de la confiance des auteurs, quand il s'agit de démontrer la télicité des prédicats.¹

L'objectif de cet article est de repérer les propriétés de *en X temps* pour essayer de répondre à la question de savoir si cet adverbial constitue réellement un test fiable. Pour ce faire, nous avons analysé quelques 450 énoncés extraits de la base textuelle Frantext, d'un corpus de presse régionale (L'Est Républicain, désormais ER) et ramenés par des requêtes envoyées *via* un moteur de recherche. Les caractéristiques de l'adverbial *en X temps* que nous avons pu observer dans cet échantillon permettront ensuite d'évaluer si cette expression mérite la confiance qu'on lui accorde dans son rôle de test de télicité. Nous adoptons donc le chemin inverse de celui suivi habituellement: plutôt que de partir du prédicat pour vérifier s'il s'accorde avec *en X temps*, nous nous intéressons avant tout à l'adverbial lui-même, tel qu'il apparaît dans son « élément linguistique ».

* Les commentaires des relecteurs anonymes m'ont grandement aidée à améliorer le texte. Je tiens à les en remercier chaleureusement

1. Souvent, les auteurs évoquent, dans leurs parties introductives, les difficultés que présente le maniement de *in X time* et de *en X temps*, tout en continuant de s'en servir dans l'argumentation qu'ils développent (cf. plusieurs articles réunis dans R. Binnick, 2012).

Dans la Section 1, nous rappelons brièvement son parcours depuis l'usage qui en a été fait chez Z. Vendler (1957/1967)², nous revenons sur la notion de télicité pour enfin nous interroger sur ce que devrait être un test. La Section 2 est entièrement consacrée à la discussion des exemples de corpus dont nous réunirons les conclusions dans la troisième et dernière partie.

1. Un bref historique d'une carrière étonnante

1.1 *In X time* et les schémas temporels des verbes anglais chez Z. Vendler (1967)

La paternité de ce test est attribuée à Z. Vendler (1967), qui utilise les adverbiaux temporels *in X time* et *for X time* pour caractériser les « schémas temporels » associés aux verbes anglais. Il remarque notamment que les accomplissements tels que *running a mile* et *drawing a circle* s'accordent avec le premier de ces deux adverbiaux. De même, dans le cas d'un achèvement tel que *to reach the top*, on pourra dire *He reached the top in three hours*, même si quelques réajustements interprétatifs sont alors nécessaires, puisque ce qui a pris trois heures, ce n'est pas l'action même d'atteindre le sommet, mais l'escalade qui l'avait précédée.

Deux points doivent être soulignés à ce propos. Premièrement, les accomplissements et les achèvements cités par Z. Vendler comme compatibles avec *in X time* relèvent des classes aspectuelles qui existent également en français. À cet égard, le parallèle avec *en X temps* est pleinement justifié, puisque *courir le deux-cents mètres* (accomplissement) et *atteindre le sommet* (achèvement) sont des prédicats téliques, parfaitement compatibles avec cet adverbial temporel.

Deuxièmement, il est néanmoins frappant de constater qu'à aucun moment, Z. Vendler (1967) ne qualifie *in X time* de preuve, ni même de test de télicité. Il s'agit, pour lui, d'une expression permettant de mettre en évidence des propriétés temporelles, et non pas d'un moyen de montrer que la situation dénotée par le verbe possède un terme inhérent. Nous reprendrons le fil de cette réflexion dans la Section 1.3.

1.2 Télicité *versus* bornage

Selon Z. Vendler (1967), un verbe est télique, lorsque le schéma temporel qui lui est associé possède un *climax*, un point terminal. La télicité est donc associée au

2. Tout au long de cet article, nous nous référons au texte de Z. Vendler reproduit en 1967 dans *Linguistics and Philosophy*.

fait de dénoter la fin d'une situation. Cette conception a été formalisée plus récemment par Krifka (1992) dans le cadre de son approche mérologique et de sa thèse d'incrémentalité. Ainsi, un objet spatialement borné (par exemple, la distance d'un marathon dans (1), ci-dessous), confère un sens télélique au prédicat *courir le marathon de Paris*. Si toute situation pourvue d'une borne finale est télélique, on sera amené à penser que *courir 17 kilomètres* dans l'exemple (2) l'est également.

- (1) Dans le civil, Christian G. est un grand sportif qui a couru le marathon de Paris en avril dernier. [ER]
- (2) Nous avons couru 17 km hier ... [ER]

Or, dans D. Sikora (2009, 2012), nous avons montré que malgré une similarité de surface, les deux prédicats ont des propriétés structurales différentes. Le premier est effectivement télélique, avec un caractère indivisible, alors que le second est au contraire atélique avec une référence homogène (D. Sikora, 2009). Nombre de caractéristiques grammaticales spécifiques à chacun de ces syntagmes indiquent en effet qu'il s'agit de cas différents. Faute de place, nous ne pouvons pas reprendre ici l'argumentaire de D. Sikora (2012). Rappelons simplement que la phrase (1) se prête bien à la passivation aussi bien périphrastique que pronominale (notamment à lecture gnomique) et autorise la construction *être + participe passé* à valeur résultative, alors que dans (2) de telles transformations sont impossibles³.

Ces faits de langue confirment le bien-fondé de la distinction entre la télélicité et le bornage proposée par R. Declerck (1979, 1989), Ö. Dahl (1981) et de I. Depraetere (1995). Les deux phénomènes sont en effet conceptuellement différents dans la mesure où la télélicité est liée à un certain type de changement dénoté par le prédicat : un changement qui débouche sur un nouvel état de choses. Ainsi, les situations dénotées par les prédicats tels que *mourir*, *boire un verre de vin* ou *courir le marathon de Paris* possèdent un terme inhérent où elles basculent vers un nouvel état de choses : *X est mort*, *le verre est bu*, *le marathon de Paris est couru*.

Cela signifie *a contrario* qu'une borne finale de nature spatiale (exemple 2) ou temporelle (exemple 3 ci-dessous) ne garantit pas la télélicité du prédicat dans la mesure où aucun état résultant ne peut leur être associé. Les bornes ne marquent pas un nouvel état de choses, elles ne garantissent pas le caractère téléonomique de la situation. Elles délimitent une séquence d'activité qui reste atélique, comme cela ce fait dans l'exemple (3).

3. À partir de l'exemple (1), il est en effet possible de former un passif périphrastique *Le marathon de Paris a été couru par Christian G. en avril dernier* et un passif pronominal *Le marathon de Paris s'est couru en avril dernier*. L'exemple (2) n'autorise pas ce type de modifications.

- (3) L'organisation par classe est la suivante : les CP-CE1 courent 30 minutes sur un circuit de 200 m, 2 élèves courent en même temps et afin d'assurer les relais à chaque tour, il est souhaitable de composer des équipes de 6 élèves.
[ER]

1.3 De l'utilité d'un test (linguistique)

Réfléchissons à présent sur ce que devrait être un test de télicité. Ou bien, pour poser la question d'un point de vue fonctionnel : à quoi sert un test ? En chimie, un indicateur de pH appliqué à une solution permet de savoir si celle-ci est un acide ou une base. Si l'on considère que les adverbiaux temporels *in X time* en anglais et *en X temps* en français sont des tests de télicité, on est en droit de s'attendre à ce que leur application à un prédicat ou à une prédication permette d'en détecter la télicité.

Notons que la confiance accordée à ces adverbiaux n'est pas générale. H. Verkuyl et *alii* (2004) remarquent qu'il n'est pas certain à quel niveau ils doivent s'appliquer : celui du verbe, du prédicat ou de la prédication. A. Molendjik et *alii* (2004) indiquent qu'il s'agit d'une expression qui, loin d'être neutre, participe à la dérivation aspectuelle.

Une autre difficulté surgit dans une perspective multilingue. La télicité est une catégorie sémantique et conceptuelle, alors que *in X time* et ses correspondants appartiennent à des langues particulières. Certes, on peut admettre sans risque d'erreur que *en X temps* est l'équivalent français de *in X time* en anglais et que, de plus, il cooccurre avec les prédicats téliques. Mais peut-on être sûr qu'une catégorie conceptuelle telle que *type de situations* maintienne la même relation avec les traductions de *in X temps* dans les différentes langues ?

Des difficultés d'application de *en X temps* en français n'auront pas échappé aux aspectologues. Un certain nombre de contraintes ont été formulées pour anticiper et contourner les difficultés susceptibles d'apparaître lors de son insertion dans la phrase. Ainsi, l'adverbial ne doit pas conduire à une lecture inchoative. On connaît également l'effet itératif produit par la conjonction de *en X temps* et d'un temps à valeur imperfective tel que l'imparfait dans le cas des achèvements. Il est frappant de constater par ailleurs que dans l'immense majorité des cas, le « test » est appliqué à des exemples construits pour les besoins des exposés scientifiques. Dans ce cas, la démarche consiste plutôt à illustrer la télicité d'un prédicat que de la vérifier.

Pour notre part, nous n'avons connaissance d'aucun travail qui justifie, sur le plan épistémologique, la promotion de *en X temps* au rang de tests de télicité. Le raisonnement qui y a conduit reste donc implicite. Il semblerait que celui-ci repose sur une généralisation accidentelle : *en X temps* étant compatible avec les verbes

qui forment des prédicats téliques, il a été admis qu'il permettait de détecter le trait [+télicité]. Or, les corpus fournissent plusieurs exemples dans lesquels notre adverbial apparaît avec un prédicat clairement atélique. C'est plus qu'il n'en faut pour invalider une généralisation accidentelle.

2. L'adverbial *en X temps* en corpus

Dans cette section, nous présentons le résultat de notre travail sur corpus. Observer *en X temps* dans des énoncés attestés pour en saisir les caractéristiques sémantiques et le fonctionnement permettra de déterminer le degré de confiance que l'on peut lui accorder dans son rôle de test. Faute de place, nous limitons la discussion aux cas particulièrement problématiques.

2.1 Adverbial de télicité et de bornage

Dans la Section 1.2, nous avons rappelé quelques faits de langue qui plaident en faveur d'une distinction claire entre la télicité et le bornage. Or, les exemples (4) et (5), illustrant respectivement le premier et le second cas, montrent une parfaite compatibilité avec l'adverbial *en X temps*.

(4) Au premier tour, il a couru le 200 mètres en 22,2. [Internet]

(5) Il a couru 100 mètres en 9 secondes et 58 centièmes. [Internet]

En admettant que la compatibilité avec *en X temps* met en évidence le caractère télique du prédicat, on serait obligé de conclure que *courir* est télique dans nos deux exemples. Or, comme nous l'avons indiqué dans Sikora (2012), il existe de bonnes raisons (linguistiques) de considérer qu'il s'agit de deux cas différents, dans la mesure où les syntagmes nominaux *le 200 mètres* dans (4) et *100 mètres* dans (5) n'ont pas la même fonction auprès du prédicat. Dans le premier cas, nous avons effectivement affaire à un prédicat télique, alors que dans le second, il s'agit de bornage. Force est de constater dès lors qu'aussi bien les prédicats téliques que les prédicats atéliques bornés peuvent s'accorder avec l'adverbial *en X temps*.

2.2 *En X temps* et les prédicats potentiellement téliques

Il existe des prédicats susceptibles d'appeler aussi bien une interprétation télique qu'atélique. À titre d'exemple, J. Jayez (1998: 138) cite le cas de *lire le livre en une heure/pendant une heure*. L'étude d'exemples attestés montre qu'il ne s'agit pas d'un cas isolé. Regardons le fonctionnement de *se coiffer* dans (6) et (7):

- (6) Je me suis coiffé pendant une heure ce matin! C'est con de me balancer sous la pluie! [Internet]
- (7) Nouveau record! Je me suis coiffé en 5 secondes. [Internet]

Les prédicats potentiellement téliques s'accordent aussi bien avec les adverbiaux traditionnellement considérés comme « tests » d'atélicité (*pendant X temps*) qu'avec ceux qui doivent mettre au jour leur télicité (*en X temps*). En réalité, ces adverbiaux font nettement plus que simplement s'accorder avec le prédicat. C'est eux qui désambigüisent les énoncés, en leur conférant respectivement une interprétation atélique et télique. C'est comme si un indicateur de pH introduisait subrepticement des ions hydrogènes dans la solution soumise à l'examen pour ensuite en faire ressortir le caractère acide.

2.3 *En X temps* et le calcul des relations discursives

Le corpus de travail que nous avons réuni recueille un nombre élevé de phrases dans lesquelles l'adverbial *en X temps* accompagne un prédicat d'état qui — par définition — dénote une situation atélique. Ce cas de figure impose une perspective d'analyse qui va au-delà de la compatibilité observée sur le plan phrastique. C'est la raison pour laquelle il nous faut, dans les sous-sections 2.3.1 et 2.3.2, examiner le fonctionnement de notre adverbial sur le plan discursif.

2.3.1 *En X temps* et les prédicats d'état

Comparons les exemples (8) et (9):

- (8) L'accord fut conclu en quelques minutes, à la satisfaction des deux parties: Alexis aurait un professeur, Mignon aurait un toit. [Frantext]
- (9) Avec son pinceau large, une seule couche suffit, et tout est sec en 2 minutes. [Internet]

Les prédicats téliques tels que *conclure un accord* dans (8), on le sait, autorisent une construction *être + participe passé* qui s'interprète en termes d'état résultant. Si l'on admet qu'il s'agit là d'une forme de leur passif appartenant au paradigme flexionnel⁴ du verbe télique *conclure*, on ne peut s'étonner de la compatibilité de celle-ci avec *en X temps*. En d'autres termes, le fait que notre adverbial prédique sur l'état résultant d'une situation télique ne suffit pas pour le disqualifier dans son rôle de test, car il existe un lien sémantique avec le processus qui le précède.

4. Une autre position consiste à penser que *être + participe passé* est une construction attributive. Les exemples (8) et (9) relèveraient alors du même cas de figure. Pour une discussion détaillée et argumentée des deux positions, voir V. Lagae (2005).

La situation est plus complexe dans le cas de *être sec* puisqu'il s'agit d'une construction attributive avec un adjectif, et non pas du participe passé d'un verbe téléique. Bien que compatible avec *en deux minutes* sur le plan phrastique, contrairement à *être conclu* dans l'exemple (8), *être sec* ne désigne pas la phase résultative sémantiquement associée à un procès téléique.

A. Borillo (1998) indique néanmoins que *sec* appartient à la classe d'adjectifs résultatifs. Ces adjectifs sont souvent liés à un verbe « qui exprime une action causative ou un déroulement de processus » (A. Borillo 1998: 184). Ils peuvent de ce fait s'interpréter en termes d'états vus comme « l'aboutissement de cette action ou de ce processus » (*ibidem*).

Il n'en est pas moins que sur le plan phrastique, l'adverbial *en deux minutes* s'accorde avec un prédicat dénotant un état, certes temporaire, mais atéléique par définition. Pour retrouver le lien entre notre adverbial et la téléicité, il ne suffit pas de limiter l'analyse de (9) à la sémantique verbale. Il est nécessaire de procéder à un calcul des relations qui se construisent sur le plan discursif. En effet, les situations e_1 (*poser une seule couche de vernis à ongles*) et e_2 *être sec* sont reliées par une relation discursive de type Narration (A. Lascarides et N. Asher 1993, N. Asher et A. Lascarides 2003): sans qu'il y ait de causation *sensu stricto*, e_2 (*être sec*) est bien une conséquence de e_1 (*application d'une couche de vernis*). Un élément implicite, à savoir le verbe *sécher* dénotant le processus e_p , est également à prendre en compte dans ce calcul: le vernis appliqué doit sécher. En réalité, *en deux minutes* ne porte pas sur *être sec*, mais il précise l'intervalle temporel dans lequel s'inscrit le processus de *sécher*, qui sans être exprimé, fait objet d'une inférence pragmatique construite sur la base des relations discursives. C'est parce que l'organisation du discours s'appuie sur ce calcul que *être sec* s'accorde avec l'adverbial *en deux minutes*, sans pour autant « prouver » sa téléicité.

2.3.2 *En X temps en position détachée*

Dans les sections précédentes, nous nous sommes concentrés sur les emplois intraprédicatifs de *en X temps*. Notre adverbial porte alors sur le verbe, en prédisquant l'intervalle temporel dans lequel s'inscrit la situation dénotée.

Avant de réunir nos observations en conclusions, il nous semble important de nous arrêter sur les caractéristiques que présente *en X temps* en emploi extraprédicatif. D'une part, ces emplois frappent par leur nombre dans tous les corpus que nous avons explorés, de l'autre, ils permettent de compléter le portrait de *en X temps*, car ils font ressortir l'une de ses caractéristiques essentielles.

Dans les exemples (10) et (11), *en X temps* est détaché en tête de phrase.

(10) En une heure, j'ai avancé de 10 mètres. [Internet]

- (11) En quelques minutes, elle avait renversé le cendrier, balancé son kir sur la moquette et allumé sa Winston par le filtre. [Frantext]

Placé en position extrapredicative, *en X temps* porte sur l'ensemble de la phrase: une proposition simple dans (10) ou une coordination dans (11). Notons cependant une différence dans le comportement des adverbiaux *en une heure* de (10) et *en quelques minutes* de (11). Le premier peut sans problème être déplacé en position post-verbale, sans que la signification de la phrase en soit touchée (exemple 12), alors que (11) résiste à une telle modification, puisque (13) paraît pour le moins bancal.

- (12) J'ai avancé de 10 mètres en une heure.

- (13) ?Elle avait renversé le cendrier, balancé son kir sur la moquette et allumé sa Winston par le filtre en quelques minutes.

En effet, le transfert en position post-verbale échoue en (11)–(13), car il enlève à notre adverbial la fonction de cadre temporel qui va de pair avec son antéposition. Tout comme dans ses emplois post-verbaux intrapredicatifs, *en X temps* détaché en tête de phrase exprime un intervalle temporel inclusif. Une différence de portée doit cependant être notée: dans cet intervalle s'inscrivent tous les événements dénotés par les prédicats successifs de (11) *renverser le cendrier, balancer son kir, allumer sa Winston*.

Ainsi, notre détour extrapredicatif met en évidence une caractéristique de *en X temps* commune à tous ses emplois. Indépendamment de sa position syntaxique et de sa portée (prédicat, prédication simple ou bien prédication multiple), cet adverbial reste invariablement endophrastique (C. Guimier, 1996). En d'autres termes, partout où ils apparaissent, les adverbiaux de ce type « affectent le contenu même de l'élément sur lequel ils portent et, ce faisant, ils participent à la construction du sens référentiel de la phrase. » (C. Guimier, 1996: 6). Les exemples que nous venons d'étudier vont bien dans ce sens: *en X temps* est de nature endophrastique en ce qu'il participe « à la construction du sens référentiel » des prédicats dont il est censé tester la télécité.

3. Conclusions

Les exemples que nous venons de discuter montrent qu'il existe des raisons fortes de se méfier de l'adverbial *en X temps*, quand il s'agit de vérifier la télécité, tant sur le plan du prédicat qu'au niveau phrastique. Dès que l'on dépasse le cadre rassurant des exemples construits pour les besoins d'une démonstration, le prétendu « test » montre son véritable triple visage:

1. Sémantiquement, *en X temps* ne permet pas de faire la distinction entre télicité et bornage. Cela n'a rien d'étonnant, puisqu'il s'agit d'une expression à caractère inclusif qui définit l'intervalle temporel dans lequel s'inscrit la situation décrite par le prédicat. Cela signifie que notre adverbial détecte simplement la présence d'une borne temporelle finale, mais ne permet pas de conclure qu'il s'agit d'une borne tétique.
2. *En X temps* est un adverbial endophrastique, aussi bien en position intraprédicative qu'extraprédicative. Il contribue activement à la construction de la signification et de ce fait, n'est pas neutre pour le calcul des propriétés aspectuelles. Nous avons vu, dans la Section 2.2, que dans de nombreux cas, il introduit dans la phrase le trait de télicité qu'il est censé tester.
3. Bien au-delà de la phrase, *en X temps* est souvent intégré dans un réseau de relations qui structurent le discours. C'est la raison pour laquelle il peut très bien s'accorder avec des prédicats atéliques, comme nous l'avons vu dans la Section 2.3.1.

L'utilisation d'un indicateur de pH est soumise à des contraintes, notamment en ce qui concerne certaines températures qui affectent son fonctionnement et faussent les résultats. Les difficultés et les précautions qu'il convient d'adopter en maniant *en X temps* vont au-delà des contraintes que l'on devrait prendre en compte en appliquant un test de télicité. On aurait donc tout à gagner à prendre *en X temps* pour ce qu'il est : un adverbial inclusif qui apporte du sens.

Œuvres citées

- Asher, N., & Alex, L. (2003). *Logics of conversation*. Cambridge: Cambridge University Press.
- Binnick, R. (Ed.). (2012). *The Oxford handbook of tense and aspect*. New York: Oxford University Press.
- Borillo, A. (1998). Les adjectifs et l'aspect en français. *Cahiers Chronos*, 2. Amsterdam & Atlanta: Rodopi.
- Dahl, Ö. (1981). On the definition of the Telic-Atelic (bounded-unbounded) distinction. In P. J. Tedeschi & A. Zaenen (Eds.), *Syntax and Semantics. Tense and Aspect*, vol. 14. New York: Academic Press.
- Declerck, R. (1979). Aspect and the bounded/unbounded distinction. *Linguistics*, 17. The Hague: Mouton Publishers.
- Declerck, R. (1989). Boundedness and the structure of situations. *Leuvense Bijdragen*, 78.
- Depraetere, I. (1995). On the necessity of distinguishing between (un)boundedness and (a)telicity. *Linguistics and Philosophy*, 18. Dordrecht, Boston: D. Reidel.
- Guimier, C. (1996). *Les adverbes du français: le cas des adverbes en -ment*. Paris: Ophrys.
- Jayez, J. (1998). DRT et imparfait. Un exemple de traitement formel du temps. In J. Moeschler (Ed.), *Le temps des événements. Pragmatique de la référence temporelle*. Paris: Éditions Kimé.

- Krifka, M. (1992). Thematic relations and links between nominal reference and temporal constitution. In I. Sag & A. Szabolcsi (Eds.), *Lexical Matters*. CSLI Publications, Chicago University Press.
- Lagae, V. (2005). Les formes en *être* + participe passé à valeur résultative dans le système verbal français. *Cahiers Chronos*, 12, Amsterdam/New York: Rodopi.
- Lascarides, A., & Nicolas, A. (1993). Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy*, 16, Dordrecht: Kluwer.
- Molendijk, A., Borillo, A., Bras, M., Le Draoulec, A., de Swart, H., Verkuyl, H., Vet, C., Veters, C., & Vieu, L. (2004). Meaning and use of past tenses in discourse. In F. Corblin & H. de Swart, (Eds.), *Handbook of French semantics*. Stanford: CSLI Publications.
- Sikora, D. (2009). *Les Verbes de manière de mouvement en polonais et en français. Éléments pour une étude comparée des propriétés structurelles des prédicats*. Thèse de doctorat, Nancy, Université Nancy 2 — ATILF-CNRS (UMR 7118).
- Sikora, D. (2012). Compléments aspectuels et verbes de manière de déplacement en français: entre bornage et télicité. *Cognitive Studies/Études cognitives*, 12. Varsovie: Slawistyczny Ośrodek Wydawniczy.
- Vendler, Z. (1957). Verbs and times. In *The philosophical review*, 66, reproduit dans Vendler, Z. 1967. *Linguistics in philosophy*. Ithaca: Cornell University Press, New York.
- Verkuyl, Henk, Borillo, A., Bras, M., Le Draoulec, A., Molendijk, A., de Swart, H., Vet, C., Veters, C., & Vieu, L. 2004. Tense and Aspect in Sentences. In: *Handbook of French Semantics*, F. Corblin, H. de Swart, (eds), Stanford: CSLI Publications.

Résumé

L'adverbial *en X temps* est souvent utilisé comme test: inséré dans une phrase, il est supposé mettre en évidence la télicité du prédicat. Le présent article propose une réflexion sur le degré de confiance que l'on peut accorder à cet adverbial dans son rôle de test en français. En examinant ses caractéristiques sémantiques et discursives, il convient d'admettre que *en X temps* n'est pas un moyen fiable de détecter le caractère télique d'un prédicat, voire d'une prédication.

Mots-clés: aspect, télicité, adverbiaux temporels

Summary

En X temps and the problem of the (supposed) telicity test in French

The French adverbial PP *en X temps* is often considered telicity test. The present paper address the question of its reliability: how can we be sure that a VP is telic as long as it is compatible with *en X temps*? Corpus data investigated in the article provide evidence that the supposed telicity test often appears in clearly atelic contexts. Further investigation of the linguistic features of the PP under analysis in such contexts leads to the conclusion that there is no reason to rise the adverbial clause *en X temps* to the rank of telicity test.

Keywords: aspect, telicity, temporal adverbials.

Adresse de l'auteur :

Dorota Sikora
Université du Littoral-Côte d'Opale
ISCID-CO
280, avenue de l'Université
BP 69
F-59942 Dunkerque cedex 02
Dorota.Sikora@univ-littoral.fr

Reçu en septembre 2013

D'achèvement à état : vers une modélisation de quelques structures de polysémie aspectuelle

Dorota Sikora*

1. INTRODUCTION : OBJET ET OBJECTIFS DE L'ÉTUDE¹

Le phénomène est parfois mentionné au détour de travaux consacrés à d'autres problèmes de sémantique verbale : la polysémie de certains verbes comprend des acceptions radicalement opposées sur le plan aspectuel. Apothéloz (2008) évoque le cas de *se taire* qui lexicalise deux sens : celui de 'cesser de parler' appartenant à la classe d'achèvements (*se taire*₁) et celui de 'rester sans parler', *se taire*₂, à classer parmi les états.

Cet article s'intéresse à ce type de polysémie aspectuelle. Nous examinons des verbes qui, dans leur sens intuitivement perçu comme premier, dénotent un achèvement, tout en comptant parmi leurs acceptions (i.e. parmi leurs sens copolysèmes) un état. C'est le cas du verbe *arriver* dans les exemples (1)² :

- (1) a. Après de multiples interventions téléphoniques, le train est arrivé à 10 heures à Saint-Mesmin.
b. La chaussée aux pavés anciens arrivait à la gare.

Si ce type de mouvement polysémique a attiré notre attention, c'est tout d'abord en raison du paradoxe apparent qu'il recouvre. D'une part, les achèvements s'opposent aux états par le plus grand nombre de traits définitoires : les premiers sont dépourvus de durée, dynamiques et téléliques, alors que les seconds recouvrent un intervalle temporel de manière homogène, sans changement et sans borne télélique. Une structure polysémique comme celle des verbes *se taire* et *arriver* comprend donc des acceptions aspectuellement opposées. D'autre part, l'abondante littérature consacrée à l'aspect met en avant

* Université du Littoral-Côte d'Opale, courriel : Dorota.Sikora@univ-littoral.fr

¹ Je tiens à remercier chaleureusement les relecteurs anonymes pour leurs remarques et commentaires, qui m'ont permis d'améliorer une première version de cet article.

² Nos exemples sont extraits de la Base textuelle *Frantext*, du corpus du Web français *FrWac* et de celui de presse locale (corpus du journal *L'Est Républicain*). Faute de place, et dans un souci de clarté, nous avons souvent choisi de les abrégés, en veillant à ne pas modifier le cotexte immédiat du verbe analysé.

plusieurs caractéristiques communes aux achèvements et aux états, qui se manifestent notamment sur le plan phrastique. Plusieurs auteurs observent en effet que certains verbes anglais (*to know, to remember, etc.*) peuvent désigner un achèvement dans des emplois tels que 2. L'exemple 3 montre qu'il en est de même avec le verbe *savoir* en français³.

- (2) And then suddenly I knew! (Vendler, 1967 : 111, Mourelatos, 1981 : 196)
 (3) J'ai su dans la seconde que mon neveu disait vrai, que tu étais quelque part vivante. (Frantext)

Lenci (1995) note un autre lien entre achèvements et états, lien qui se manifeste dans les phrases gnominiques et habituelles. Quand celles-ci se construisent sur des prédicats d'achèvements (*episodic predicates*, cf. Bertinetto et Lenci, 2012), elles peuvent désigner des habitudes, c'est-à-dire des propriétés (i.e. situations statives plus ou moins étendues dans le temps) que l'on assigne à un argument externe.

La présente étude aborde le problème de polysémie aspectuelle sur un plan purement lexical. Sans nier la nature compositionnelle de l'aspect, nous nous intéressons aux structures situationnelles telles qu'elles sont codées dans le sens des lexèmes verbaux. En effet, l'aspect lexical d'une acception particulière est la première donnée à prendre en compte dans le calcul des propriétés aspectuelles d'une prédication. Nous avons cherché à répondre à trois questions :

1. La structure polysémique achèvement-état, est-elle fréquente en français ou bien s'agit-il d'un phénomène marginal ?
2. Est-il possible de dégager un ou plusieurs modèles qui sous-tend(ent) ce mouvement polysémique ?
3. Peut-on dégager des facteurs sémantiques qui favorisent ou, au contraire, bloquent la perte de dynamicité qui fait apparaître un sens statif ?

Pour explorer ce phénomène, nous avons analysé 212 verbes prélevés dans deux bases de données lexicales : *Le Réseau Lexical du Français* (désormais *RL-fr*, cf. Lux-Pogodalla et Polguère, 2011, Polguère et Sikora, 2013) et *Les Verbes Français* (*LVF*, cf. Dubois et Dubois-Charlier, 1997). Dans la section 2, nous présentons les positions théoriques et les choix méthodologiques de notre travail. La section 3 est entièrement consacrée à la discussion de nos résultats. En 4, nous proposons un bilan succinct plutôt qu'une conclusion, dans la mesure où nos investigations sont toujours en cours.

2. QUESTIONS THÉORIQUES ET MÉTHODES D'ANALYSE DES DONNÉES

Les questions que nous avons choisi d'étudier nous amènent dans un périmètre où il est délicat d'avancer. D'une part, la polysémie fait objet de plusieurs

³ Mourelatos (1981 : 196) parle à ce propos de « semantic multivalence of State Verbs ». Par ailleurs, dans une approche constructionniste, Croft (2012 : 38) souligne également que ces prédicats statifs (*inceptive states*) peuvent se construire en états transitoires ou en achèvements.

débats, qui vont de la question de son existence même (Cadiot et Visetti, 2001, Kleiber, 2008, à titre d'exemple) jusqu'aux modes de sa conceptualisation et de ses représentations possibles (cf. François, 2010 pour une discussion récente). D'autre part, la classe d'achèvements n'est pas unanimement admise. Certains auteurs (Verkuyl, 1972, Parsons, 1985) contestent la distinction entre accomplissements et achèvements, alors que de nombreuses études (Mittwoch, 1991, Piñon, 1997, Martin, 2011) apportent, au contraire, des arguments en faveur du maintien de cette classe aspectuelle.

Le cadre de cet article ne permet pas, faute de place, de justifier la position que nous adoptons dans ces débats. Le titre et l'objet de cette étude indiquent néanmoins clairement que nous nous situons à la fois parmi celles et ceux qui admettent les phénomènes de polysémie en langue et accordent aux achèvements une place dans les typologies de situations. Les sous-sections 2.1 et 2.2 exposent notre approche, en insistant sur les outils adoptés dans le traitement des données.

2.1 Polysémie et discrimination des sens multiples

La polysémie, pluralité de sens sémantiquement liés, n'est pas traitée de la même façon dans les deux ressources exploitées pour notre recherche. La base *LVB* adopte une approche homonymisante (cf. la répartition proposée dans François, 2010) : elle décrit les acceptions d'un vocable polysémique, sans dégager de liens sémantiques entre elles.

Le *RL-fr* analyse au contraire ces relations dites de copolysémie pour situer chaque sens par rapport à ses acceptions copolysèmes, dont chacune est considérée comme lexème à part. Ainsi, l'exemple (1) ci-dessus présente deux lexèmes copolysèmes : *arriver*₁ en (1a) et *arriver*₂ en (1b).

L'approche mise en œuvre dans le *RF-fr* se situe dans une perspective synchronique dérivationnelle (cf. François, 2010), dans la mesure où elle repose sur l'hypothèse d'un sens de base. Il s'agit d'un lexème qui, dans un état donné de l'évolution de la langue, en l'occurrence en période contemporaine, est le plus saillant pour le locuteur (cf. à ce propos, Wilkins, 1996). Dans le cas du vocable *arriver*, c'est le verbe de déplacement illustré dans (1a), et non pas le premier sens attesté en diachronie 'toucher la rive', qui constitue l'acception de base pour l'ensemble des locuteurs contemporains. Il s'agit, certes, d'une hypothèse, mais celle-ci peut être vérifiée : si nous demandons à un francophone ce que signifie le verbe *arriver*, il y a fort à parier que sa réponse consistera à expliciter spontanément le sens de déplacement illustré dans (1a), plutôt que celui de *arriver*₂ de (1b). En résumé, c'est une optique dérivationnelle en ce qu'elle admet qu'à partir d'un sens de base, un locuteur contemporain perçoit et construit des liens sémantiques entre les copolysèmes.

C'est en adoptant ce point de vue, i. e. une approche synchronique dérivationnelle telle qu'elle est pratiquée dans le *RL-fr*, que nous avons procédé à une première série de vérifications des 212 verbes isolés dans les deux bases. L'objectif en était de nous assurer qu'il s'agit bien de pluralité des sens non-unifiables (Kleiber, 2008), et non pas d'effets contextuels.

Si les écarts théoriques dans le traitement de la polysémie sont très marqués, on observe – au-delà des terminologies disjointes – un certain accord, quant à la façon dont la non-unifiabilité des sens (i. e. l'impossibilité de réduire deux copolysèmes à un seul sens) se manifeste dans la langue⁴. Pour l'évaluer, nous avons eu recours aux trois paramètres régulièrement mentionnés dans la littérature :

1. (Im)possibilité de factorisation des contenus sémantiques. Sauf jeu de mots, il est impossible, en cas de copolysèmes, de coordonner les arguments propres à chacun d'entre eux – sujets et/ou compléments – pour les combiner avec une même occurrence verbale. Ainsi, l'énoncé (4) ci-dessous est difficilement acceptable dans la mesure où les deux lexèmes *arriver*₁ et *arriver*₂ résistent à toute tentative d'unification dans une même occurrence, susceptible de sélectionner comme sujet grammatical à la fois *train* et *chaussée*.

(4) ? Le train et la chaussée aux pavés anciens arrivent à la gare.

2. Autonomie relationnelle (Croft et Cruse, 2004), qui se manifeste dans des réseaux de relations sémantiques distincts. Deux types de liens de sens sont ici particulièrement significatifs. Premièrement, les liens de dérivation sémantique (Mel'čuk et al., 1995), notamment les noms déverbaux (noms d'action ou noms d'actants). Ainsi, le sens du verbe *arriver*₁ dans (1a) permet d'exprimer celui de *arrivée*₁ (fait, pour X, d'*arriver*₁ à Y). Il est également possible de désigner le premier actant X de ce verbe comme *arrivant* ou *arrivante*. *Arriver*₂ dans (1b) n'est pas connecté à ces lexèmes nominaux. Deuxièmement, l'autonomie relationnelle se manifeste dans des réseaux de quasi-synonymes et d'antonymes distincts : le verbe de déplacement *arriver*₁ dans (1a) est en relation de quasi-synonymie avec *venir*₁, antonyme de *partir*₁, *quitter*₁, alors que *arriver*₂ employé dans (1b) se laisse paraphraser par *aboutir*, *déboucher*, *donner (sur)*, voire *se terminer*, alors que pour exprimer son contraire, on utilisera le verbe *commencer*⁵.

3. L'autonomie compositionnelle (Croft et Cruse, 2004) des copolysèmes est un autre signe de leur non-unifiabilité sémantique. Elle se manifeste par des réseaux de cooccurrents distincts de chacune des acceptions. Ainsi, *arriver*₁ et *arriver*₂ sélectionnent des modificateurs adverbiaux différents. Dans (5a), *arriver*₁ accepte l'adverbial *à toute vitesse*, alors que *arriver*₂ dans (5b) ne peut pas se combiner avec cette locution.

(5) a. Le train arrive à la gare de Saint Mesmin à toute vitesse.

⁴ Cf. à titre d'exemple, les propositions Mel'čuk et al. (1995) qui s'inscrivent dans un courant néostructuraliste, la description de la polysémie par Croft et Cruse (2004) d'un point de vue cognitiviste et l'approche sémantico-référentielle de Kleiber (2008).

⁵ Notons que ce critère doit bien être renforcé par d'autres dans la mesure où certains synonymes et antonymes de l'acception de base développent une polysémie qui suit le même modèle. Ainsi, un achèvement tel que *partir*₁, antonyme de *arriver*₁, compte parmi ses copolysèmes un sens statif *partir*₂ (*Le chemin part de l'extrémité sud du parking*), antonyme du verbe d'état *arriver*₂ (*La chaussée aux pavés anciens arrivait à la gare*).

- b. ? La chaussée aux pavés anciens arrive à la gare à toute vitesse.

Selon ces critères, les deux emplois du vocable *arriver* dans (1a) et (1b) apparaissent comme non-unifiables. En revanche, dans (6a) et (6b), il s'agit d'une même acception du verbe *entrer* : les deux emplois acceptent la nominalisation par *entrée*, ils s'inscrivent dans le même réseau de relations paradigmatiques et ils se combinent de façon identique sur le plan syntagmatique.

- (6) a. Nous sommes entrés dans un bistrot à l'enseigne banale du *Rendez-vous des amis*. (Frantext)
 b. J'étais en train d'entrer dans la cathédrale quand j'ai entendu une détonation. (Métro, 21 mai 2013)

Dans l'exemple (6b), la locution prépositionnelle *en train de* confère une certaine durée à une situation qui en est initialement dépourvue. Nous avons affaire à la coercition comprise non pas comme changement de type aspectuel, mais comme résultat de procédés « whereby aspects and adverbial modifiers compositionally add layers of temporal predication such as preparation, initiation, iteration, and culmination, and the like » (Steedman, 2012 : 104).

2.2 Identification des types de situation

Une seconde vérification a eu pour objectif d'établir le diagnostic aspectuel pour chacun des copolysèmes. Il fallait nous assurer que, comme pour *arriver*₁ et *arriver*₂, le sens premier appartient réellement à la classe d'achèvements et que le second est bien un verbe d'état.

Pour ce faire, nous avons mis en relation un ensemble de traits considérés dans la littérature comme définitoires pour les achèvements, et les expressions linguistiques susceptibles de confirmer ou d'exclure leur présence.

1. Dans la mesure où les achèvements dénotent des changements instantanés, ils sont ontologiquement liés à d'autres événements qui les précèdent⁶ ou qui les suivent (Piñón, 1997). Lorsqu'ils réfèrent à une borne télique située à droite, i. e. clôturant une situation qui les précède, les achèvements véhiculent un présupposé. Ainsi, l'interprétation de l'énoncé (7a) présuppose que le train était déjà en route, attendu en gare. La lecture de (7b) n'en a nullement besoin.

- (7) a. Le train est arrivé à Saint Mesmin.
 b. La chaussée aux pavés anciens arrivait à la gare.

2. De nombreux achèvements acceptent le progressif qui, en français, s'exprime entre autres par la locution *en train de*. La prédication renvoie alors non pas à la

⁶ L'analyse de Piñón (1997) se situe dans une approche méréologique. Si l'on adopte celle que propose Smith (1991), on pourra rapprocher les situations situées à gauche et à droite de l'achèvement respectivement des phases préparatoire et résultative.

transition elle-même, mais à sa phase préparatoire⁷. Le progressif porte ainsi sur cette dernière, alors que la transition est présentée selon un mode prospectif. C'est la lecture que l'on accorde à (8a)⁸. (8b), qui selon notre hypothèse est un état, n'accepte pas l'ajout de *en train de*.

- (8) a. Le train est en train d'arriver à Saint Mesmin.
b. ? La chaussée aux pavés anciens est en train d'arriver à la gare.

3. Lorsqu'un verbe d'achèvement renvoie à la borne droite d'une situation (cf. *left achievements* in Martin, 2011), il se combine avec l'adverbial *en x temps*. Ce modifieur délimite alors la période qui précède immédiatement et qui conduit à la transition dénotée par le verbe d'achèvement. Dans l'exemple (9a), l'adverbial *en vingt minutes* précise en effet non pas une durée de l'arrivée du train, mais celle de la partie finale du trajet : il est par conséquent paraphrasable par *au bout de x temps*. Dans (9b), l'adjonction de *en x temps à arriver₂* produit un énoncé plus que problématique, du moins sans une opération discursive complexe visant à (re)construire un contexte particulier.

- (9) a. Le train est arrivé à Saint Mesmin en vingt minutes (= au bout de vingt minutes).
b. ? La chaussée aux pavés anciens arrivait à la gare en vingt minutes/au bout de vingt minutes.

4. De manière générale, les verbes d'achèvement résistent à l'adverbial *pendant x temps*, comme c'est le cas d'*arriver₁* en (10a). Notons cependant que certains achèvements droits acceptent une telle modification (cf. l'exemple (10b) emprunté à Apothéloz, 2008). Celle-ci porte alors non pas sur la transition dénotée par le verbe, mais sur la durée de son état résultant. Avec un verbe d'état en (10c), notamment employé avec un temps à point de vue imparfaitif, il est tout à fait possible d'utiliser un tel modifieur.

- (10) a. ? Le train est arrivé à Saint Mesmin pendant vingt minutes.
b. À cette époque, il part pendant plusieurs semaines vers l'Arizona, via la Virginie et le Texas. (Apothéloz, 2008 : 202)
c. Pendant vingt ans, la chaussée aux pavés anciens arrivait à la gare.

5. Le caractère instantané des achèvements s'accorde avec la localisation temporelle à *10 heures*, comme le montrait dès la section 1 l'exemple (1a). Dans (1b), un tel ajout est impossible, quel que soit le tiroir temporel du verbe d'état.

6. Plutôt que de critère, il est plus exact de parler de tendance générale : les verbes d'achèvement ont une nette préférence pour les temps accomplis.

⁷ Piñón (1997 : 279) observe qu'en anglais, le progressif appliqué aux achèvements introduit une modification sémantique. Ainsi, pour rendre la signification de *Rebecca was reaching the summit when it began to rain*, Piñón propose d'employer le verbe *to approach* : *Rebecca was approaching the summit when it began to rain*.

⁸ Notons qu'elle diffère de celle de l'exemple (6b) ci-dessus, où l'on a affaire à une coercition.

Employés à l'imparfait, ils appellent une lecture narrative ou habituelle illustrée en (11a). Leurs copolysèmes statifs apparaissent en général avec des temps inaccomplis (cf. 11b).

- (11) a. Le train arrivait à 10 heures à Saint-Mesmin.
b. ? La chaussée aux pavés anciens est arrivée à la gare.

L'incompatibilité avec les adverbiaux agentifs (*délibérément, volontairement*, etc.) n'a pas été considérée à ce stade, car elle caractérise aussi bien les achevements que les états. Ce critère nous sera cependant utile dans la section 3.2.

3. PRÉSENTATION ET DISCUSSION DES RÉSULTATS

Le travail de collecte et d'analyse de données s'est déroulé en deux étapes. Dans un premier temps, nous avons examiné 148 verbes prélevés dans la base *RL-fr*. Les résultats partiels, que nous ne discutons pas ici faute de place, ont ensuite été confirmés lors de la seconde étape, qui a élargi l'échantillon initial, en y intégrant des données supplémentaires en provenance de la base *LVF*.

Parmi les 212 vocables analysés, 112 (soit 56,6 % de l'ensemble examiné) possèdent une structure polysémique avec un premier sens appartenant à la classe d'achevements et un sens statif (section 3.1). 92 achevements, représentant 43,4 % des vocables de notre liste n'ont pas de copolysème d'état (section 3.2).

3.1 Vers une modélisation de polysémie achèvement-état

Les régularités repérées dans ce type de polysémie vont au-delà des classes sémantico-syntaxiques établies par exemple par Levin (1993) et Sénéchal et Willems (2007). Le cas le plus fréquent, discuté dans la sous-section 3.1.1, en est une illustration. Face à la diversité des classes verbales qui s'inscrivent dans ce patron, nous avons opté, dans l'état présent de notre recherche, pour une caractérisation métalexicaux du contenu sémantique des deux copolysèmes.

3.1.1 Modèle de phases relationnelles

Parmi les 120 achevements dotés de copolysème statif, 58 s'inscrivent dans un patron sémantique que nous avons qualifié de modèle de phases relationnelles. En effet, en tant qu'achevements (sens premier), ces verbes dénotent une entrée ou une mise en relation de deux participants. Leur sens statif décrit le fait, pour leur premier actant X, d'être dans une certaine relation avec l'actant Y. Regardons les exemples (12) à (14).

- (12) a. Nous franchissons la frontière hollandaise au milieu des forêts. (FrWac)
b. La route franchissait une petite colline. (FrWac)

- (13) a. Un grand nombre d'ouvriers adhèrent à la franc-maçonnerie et aussi à la « Ligue de l'enseignement ». (Frantext)
 b. Sa jupe, qui dans l'état normal descendait sous le genou, était tombée à la cheville ; la soie trempée pesait, comme alourdie de plomb, elle adhérait à la peau. (Frantext)
- (14) a. Lee m'a donné un bonbon anglais. (Frantext)
 b. Ma chambre donne sur la rue principale du faubourg. (Camus, A., *L'étranger*, 1942, Folio, p. 35)⁹

Les achèvements qui donnent lieu à ce mouvement polysémique se caractérisent par une grande diversité sémantique. On y trouve des verbes de déplacement tels que *franchir*₁ en (12a), des verbes de contact (par exemple *adhérer*₁ en 13a) et plus rarement des verbes de transfert (*donner*₁ en 14a). Avec deux ou trois actants, ils dénotent l'entrée ou la mise en relation interactionnelle entre X et Y (12a, 13a) ou entre Y et Z (14a). Selon la classe syntactico-sémantique dans laquelle se range leur sens de base, l'entrée ou la mise en relation peut consister dans un contact (*atteindre*₁, *rejoindre*₁, etc.), dans un changement de relation locative (*entrer*₁, *sortir*₁, etc.), sociale (*adhérer*₁, *épouser*₁, *se marier*₁) entre les premier et second actants, ou bien dans une nouvelle relation possessive (*donner*₁, *offrir*₁) qui relie Y (*un bonbon anglais*) à Z (*moi*).

Dans cette configuration sémantique, les verbes d'état *franchir*₂, *adhérer*₂ et *donner*₂ dans (12b), (13b) et (14b) dénotent une relation de positionnement ou de configuration spatiale de l'actant X (*la route*, *la soie de la robe*, *la chambre*) par rapport à un Y (*une petite colline*, *la peau*, *la rue principale du faubourg*). Ces verbes d'état sont systématiquement des prédicats bi-argumentaux : ils n'engagent que deux actants non humains, quelle que soit la structure actancielle du verbe d'achèvement de base.¹⁰

C'est dans ce patron de polysémie que pourraient s'inscrire une douzaine d'états inceptifs (cf. section 1), par exemple *se taire*, *se cacher*, *se rappeler*, *se souvenir*, *oublier*, etc. Cela reviendrait cependant à considérer que dans leurs sens premier, ils dénotent une transition, ce qui est une hypothèse difficilement acceptable.

3.1.2 Modèle de polysémie régulière de verbes de communication

30 achèvements forment un groupe qui, contrairement à celui de la section précédente, se démarque par une forte homogénéité sémantique. Il s'agit de verbes qui dénotent un acte de communication réalisé par un agent : *affirmer*₁, *appeler*₁, *autoriser*₁, *confirmer*₁, *dire*₁, *exprimer*₁, *publier*₁, etc. En tant que

⁹ Nous reviendrons sur le cas de *donner* dans la section 3.2, notamment en discutant les raisons pour lesquelles nous le comptons parmi les achèvements.

¹⁰ Pour certains verbes de ce groupe (*arriver*, *adhérer*, *franchir*, etc.), on peut se demander dans quelle mesure ce mouvement polysémique s'explique sur le plan psychologique par le concept de *mouvement abstrait* proposé par Langacker (1987). Je remercie le relecteur anonyme d'avoir attiré mon attention sur cette question.

verbes d'état *affirmer*₂, *appeler*₂, *autoriser*₂, *confirmer*₂, *dire*₂, *exprimer*₂, *publier*₂, etc., ils prédisent une propriété, à savoir le fait, pour un document, de contenir un certain type d'information. Les phrases (15a) et (15b) illustrent ce modèle de polysémie :

- (15) a. Le policier confirme le récit de la victime dans son rapport.
b. Le rapport du policier confirme le récit de la victime.

C'est un cas de polysémie régulière (Apresjan, 1974, Willems, 2005), que François et Sénéchal (2008) qualifient de syntaxique : la modification qui s'opère au passage de *confirmer*₁ à *confirmer*₂ se manifeste effectivement sur le plan syntaxique, puisque le SN du circonstant *dans son rapport* en (15a) se retrouve en position de sujet dans (15b).

Cependant, le lien syntaxique semble sémantiquement motivé. Le sens lexical de *confirmer*₁ décrit l'acte de communication qui engage trois participants : X (*le policier*), un certain type d'information(s) au sujet de l'objet Y (*le récit de la victime*) à l'intention du destinataire Z. Un transfert de contenu informationnel entre les interlocuteurs, que ce soit par écrit ou à l'oral, se fait *via* un support textuel, dont le type peut être spécifié par un circonstant tel que *dans son rapport*. Si son expression sur le plan phrastique reste optionnelle, le support de contenu informationnel est strictement impliqué par le sémantisme du verbe. C'est la raison pour laquelle il est difficile de considérer qu'il s'agit là d'un participant, i. e. d'un actant, susceptible d'être réalisé syntaxiquement soit comme sujet, soit comme complément. Il semble plus exact d'y voir une spécification d'une certaine composante du sens lexical des verbes de communication. Si notre raisonnement est exact, la différence sémantique entre (15a) et (15b) ne se réduit pas à une alternance de réalisation argumentale d'un seul et même sens verbal.

En tant que prédicat d'état dans l'exemple (15b), *confirmer*₂ sélectionne deux actants : un support textuel devient le premier actant X, alors que Y représente son thème. La phrase (15b) montre par ailleurs que l'agent X du verbe d'achèvement *confirmer*₁, en tant qu'auteur de l'énoncé, peut réapparaître sous forme de complément du nom-sujet sélectionné par le verbe d'état *confirmer*₂. Il apparaît ainsi que le lien de copolysémie entre l'achèvement *confirmer*₁ et l'état *confirmer*₂ est de nature métonymique, puisque le sens du premier inclut celui du second.

3.1.3 Modèle de métaphores situationnelles

En choisissant de regrouper 20 achèvements dans une même catégorie, nous avons pris en compte avant tout le mécanisme d'extension polysémique en œuvre. Il se peut que des analyses ultérieures conduisent à des distinctions plus fines.

Dans cette catégorie, la polysémie se construit autour d'un achèvement dénotant un événement (*crever*₁, *hériter*₁, *mourir*₁, *tomber*₁, etc) ou bien un acte volontaire réalisé par un agent (*abdiquer*₁, *anticiper*₁, *démissionner*₁, etc.). Un

lien métaphorique mène alors à un verbe statif qui décrit un état physique, psychique ou l'attitude adoptée par un premier actant humain (*crever₂*, *hériter₂*, *mourir₂*, *tomber₂*, *abdiquer₂*, *anticiper₂*, *démissionner₂*, etc.). Les exemples (17a) et (17b) d'une part, (18a) et (18b) de l'autre illustrent les configurations sémantiques achèvement-état caractéristiques de ces deux sous-ensembles :

- (17) a. La dame tombait alors à terre, sérieusement blessée (15 jours d'ITT). (L'Est Républicain)
 b. — Tu dois tomber de fatigue... (Frantext)
- (18) a. Après la manifestation qui rassembla, le 26 janvier, place de l'Opéra, environ quarante mille personnes, le cabinet démissionna ; Daladier forma un nouveau gouvernement et limogea Chiappe. (Frantext)
 b. Les pouvoirs politiques démissionnent face au pouvoir financier. (Internet)

Les achèvements tels que *tomber₁* dans (17a) sont non agentifs : le premier actant X est impliqué dans un événement dont il ne contrôle pas le déroulement, tout en en supportant les effets caractérisés par une certaine intensité. Le verbe d'état est systématiquement bi-argumental et il dénote le fait, pour X, d'éprouver de manière intense une certaine sensation physique ou bien un sentiment. Remarquons que le nom qui réalise le second actant Y (*fatigue* dans 17b) a un caractère prédicatif : il dénote un état, auprès duquel le verbe tel que *tomber₂* assume un double rôle de verbe-support et d'intensificateur. En tant que verbe support, il se laisse paraphraser par *être*, *éprouver* et *ressentir*, qui cooccurrent régulièrement avec des noms de sensations et de sentiments. Les propriétés aspectuelles semblent donc celles du prédicat nominal, alors que l'intensification du contenu sémantique de celui-ci peut être perçue comme un héritage sémantique de l'achèvement de base.

Bien que de nature métaphorique également, le mécanisme polysémique semble fonctionner différemment dans le cas des actes volontaires que dénotent les achèvements *abdiquer₁*, *anticiper₁*, *démissionner₁*, etc. Dans (18b), d'agent contrôlant son action *démissionner₁*, le premier actant du verbe d'état *démissionner₂* devient porteur d'une certaine attitude ou disposition d'esprit, qui se traduit par son inactivité face à un type de situation (l'activité du pouvoir financier dans 18b).

3.2 Absence de polysémie aspectuelle achèvement-état

92 achèvements, soit 43,4 % de tous ceux que nous avons analysés, n'ont pas de copolysème statif. Parfois, l'absence de celui-ci est facile à expliquer : les achèvements tels que *se fiancer* ou *décéder* sont monosémiques. Cependant, dans l'immense majorité des cas, ces achèvements ont développé une riche structure polysémique, sans pour autant aboutir à un copolysème statif.

L'analyse de nos données a permis de repérer des caractéristiques sémantiques communes à une forte majorité – 62 (soit 67,4 %) – de ces 92 achèvements dont la riche structure polysémique n'inclut aucun sens statif. Il s'agit de verbes syntaxiquement transitifs (*arrêter₁*, *annuler₁*, *arracher₁*, *blessé₁*, *briser₁*, *casser₁*, *déclencher₁*, *tuer₁*, etc.) ou ditransitifs (*adjoindre₁*, *affecter₁*, *arracher₁*,

*décerner*₁, *déposer*₁, *envoyer*₁, *expédier*₁, *mettre*₁, etc.). Ils forment des prédicats bi- ou tri-actanciels, qui dénotent une action exercée par l'agent X sur un patient Y, de manière à ce qu'un changement d'état de Y se produise (cf. *arracher*₁ dans 22a). L'état résultant associé à Y peut ensuite s'exprimer par une construction *être + participe passé* à valeur résultative, comme dans (22b).

- (22) a. Le facteur arracha le petit drapeau rouge fiché sur la boîte aux lettres, il le jeta à l'intérieur avec le courrier¹¹. (Frantext)
 b. Le petit drapeau rouge est arraché.

Avant d'émettre une hypothèse sur l'élément sémantique susceptible d'empêcher l'extension polysémique vers un verbe d'état, il faut se demander si le sens de base, généralement considéré comme causatif, est réellement un achèvement. Nous discutons cette question dans la sous-section 3.2.1 ci-dessous. Si, comme nous le pensons, on peut les compter parmi les achèvements, est-ce leur caractère causatif qui bloque l'apparition d'un copolysème statif? Nous essaierons de montrer, dans la sous-section 3.2.2, qu'une explication en termes de transitivité sémantique fait mieux apparaître la corrélation entre un certain type de sens lexicalisé par les verbes d'achèvements et l'absence d'acceptation stative dans la structure polysémique qu'ils développent.

3.2.1 Une structure causative en cause ?

Les 62 achèvements sans copolysèmes statifs sont des verbes que l'on qualifie souvent de causatifs. Développées dans la lignée de McCawley (1976) et de Dowty (1979), Grimshaw (1990), Levin et Rappaport-Hovav (2005), de nombreuses études recourent aux méta-prédicats tels que CAUSE et BECOME pour rendre le contenu conceptuel plutôt que lexical de ces verbes. Ainsi, McCawley (1976) décrit le sens de *to kill* par la formule CAUSE_X (BECOME(DO NOT ALIVE))_Y.

La question qui se pose dès lors est celle de savoir si ces 62 verbes sont réellement des achèvements. En effet, ces derniers sont habituellement associés avec le méta-prédicat BECOME : *to die*, comme *mourir*, se décompose en (BECOME(DO NOT ALIVE))_X.

Cependant, la causativité, comme d'ailleurs l'agentivité, est une propriété transversale des typologies de situations (cf. Levin et Rappaport Hovav, 2005, Croft, 2012) : on trouve des verbes causatifs dans toutes les classes aspectuelles, du moment où celles-ci regroupent des situations dynamiques.

Pour ce qui est des propriétés aspectuelles, *arracher*₁ de l'exemple (22) passe avec succès les tests que nous avons proposés dans la section 2.2. Beaucoup de ces verbes (*arrêter*₁, *blessé*₁, *briser*₁, *casser*₁, *déplacer*₁, *jeter*₁, *lancer*₁, *renverser*₁, *tuer*₁) véhiculent un présupposé et leur réaction aux adverbiaux *en train de*, *en x temps*, *pendant x temps* est celle des achèvements. De plus, ils

¹¹ Notons que cet exemple comporte également *jeter*₁, un autre achèvement qui relève du même cas que *arracher*₁.

acceptent la localisation temporelle et appellent une lecture itérative, quand ils sont employés à un temps imperfectif. Les critères convergent donc pour faire apparaître des caractéristiques propres aux achevements : il s'agit de transitions ponctuelles, donc dépourvues de durée, dotées d'une structure atomique.

Cependant, deux éléments vont à l'encontre de l'hypothèse selon laquelle la causativité des acceptions de base empêche une extension polysémique stative. Premièrement, il existe des verbes qui acceptent une paraphrase construite sur CAUSER (*coller*₁, *donner*₁, *réconcilier*₁, etc.) qui possèdent un verbe d'état parmi leurs acceptions copolysèmes. Deuxièmement, pour nombre d'achevements qui n'ont pas développé ce type de polysémie (*acheter*₁, *attraper*₁, *déposer*₁, *empoisonner*₁, *inscrire*₁, *vendre*₁, etc.), une paraphrase construite sur le méta-prédicat CAUSER paraît artificielle tant le niveau d'analyse qu'elle traduit s'éloigne du sens lexical du verbe, en gommant chemin faisant ses spécificités. L'hypothèse selon laquelle le caractère causatif du sens premier empêche l'apparition d'un verbe d'état dans la structure polysémique d'un vocable semble donc trop forte. Considérons à présent celle qui s'appuie sur la notion de transitivité sémantique.

3.2.2 Transitivité sémantique

Contrairement à la causativité, qui est un rapport entre deux faits (une cause et un effet), la transitivité sémantique caractérise un type d'interaction entre deux participants¹² de la situation. Hopper et Thompson (1980) insistent sur son caractère scalaire : la transitivité sémantique se décrit relativement à un ensemble de paramètres liés à la fois au sens lexical du verbe et à ses actants. Un prédicat peut donc être plus ou moins transitif sur le plan sémantique, selon le nombre de paramètres qu'il vérifie. Les caractéristiques situationnelles permettant de définir le degré de transitivité d'un sens verbal sont liées à la kinésis, à l'aspect téléique ou atélique, à la durée ou à son absence, à sa perfectivité ou imperfectivité, ainsi qu'au mode de référence à la situation (affirmatif ou négatif, factuel ou contrefactuel). Des traits liés aux actants sont également à prendre en compte. Ainsi, le degré d'agentivité se mesure selon le caractère intentionnel de l'action et l'engagement volitionnel de X. Le second actant Y, présenté de manière individuelle ou non, définie ou indéfinie, peut être affecté de différentes façons par l'action de X.

Sur le plan aspectuel, les achevements réunissent bien des caractéristiques qui tendent vers le pôle fort de transitivité (télicité, ponctualité, préférence pour les temps perfectifs). De plus, nos données montrent que l'extension de sens vers un copolysème statif ne se fait pas en présence de paramètres de forte transitivité liés aux actants. Comparons les exemples (12), (15) et (21) à l'énoncé (22).

¹² La tradition aristotélicienne (*Métaphysique*, V) reconnaît cependant des *causes efficientes*, qui incluent des individus. Ainsi, Croft (2012) admet des relations causales entre un participant et un fait.

- (12) Nous franchissons la frontière hollandaise au milieu des forêts. (FrWac, cf. la sous-section 3.1.1)
- (15) Le policier confirme le récit de la victime dans son rapport. (cf. la sous-section 3.1.2)
- (21) Et puis non. Coup de bol. Au dernier moment, la boule tombe dans un trou. (Frantext, cf. la sous-section 3.1.3)
- (22) Le facteur arracha le petit drapeau rouge fiché sur la boîte aux lettres, il le jeta à l'intérieur avec le courrier. (Frantext, cf. la sous-section 3.2)

Quel que soit leur modèle de polysémie (exemples 12, 15 et 21), les achèvements copolysèmes d'un verbe d'état possèdent des caractéristiques participatives qui tendent vers le pôle faible de transitivité sémantique. L'agentivité n'est pas un trait décisif de leur premier actant : certains achèvements relevant du modèle de phases relationnelles (*entrer₁*, *sortir₁*, etc.) sélectionnent un X agentif, alors que d'autres (*arriver₁*, *franchir₁*, *rencontrer₁*, etc.) restent neutres à cet égard. Les verbes de communication impliquent au contraire un agent qui contrôle l'acte de parole. Le troisième patron polysémique caractérise aussi bien des verbes agentifs (*abdiquer₁*, *démissionner₁*, etc.) que non-agentifs (*crever₁*, *tomber₁*, *mourir₁*). On remarque cependant que même lorsqu'ils sont agentifs, ces achèvements ne dénotent pas une action intentionnelle orientée vers un changement d'état du patient Y. En effet, ni *la frontière* dans (12), ni *le récit* dans (15) ne se trouvent affectés par l'action exercée par X.

La plupart des achèvements qui, comme *arracher₁*, n'ont pas de copolysème statif, se caractérisent au contraire par une très forte dimension participative. Le premier actant X est systématiquement agentif : engagé sur le plan volitionnel, il contrôle pleinement son action. Celle-ci est orientée vers le patient Y qui s'en trouve fortement affecté. Ainsi, les paramètres participatifs définis par Hopper et Thompson (1980) viennent renforcer les caractéristiques aspectuelles « naturelles » des achèvements, en convergeant vers un degré élevé de transitivité sémantique. Comparons à cet égard deux verbes quasi-synonymes tels que *donner₁* et *transmettre₁*. Le premier a une extension polysémique stative (exemples 14a et 14b rappelés ci-dessous), alors que le second (23) en est dépourvu.

- (14) a. Lee m'a donné un bonbon anglais. (Frantext)
b. Ma chambre donne sur la rue principale du faubourg. (Camus, A., *L'étranger*, 1942, Folio, p. 35)
- (23) Mais le Palais plaît à Khrouchtchev : on s'est borné à transmettre les lettres aux architectes. (Frantext)

Les verbes *donner₁* et *transmettre₁* dénotent tous les deux le transfert de l'objet Y effectué par X vers Z, ce qui détermine leur structure tri-argumentale. Cependant, les relations entre les participants telles qu'elles sont saisies dans le sens lexical diffèrent sensiblement, notamment en ce qui concerne le caractère plus ou moins transitif de l'action effectuée par X. Dans le cas de *donner₁*, le second actant Y (*un bonbon anglais*) est, certes, un objet de transfert, mais on ne

saurait le qualifier de patient. L'action de X (*Lee*) n'est pas à proprement parler orientée vers le bonbon et un éventuel changement d'état de la friandise, mais vers le bénéficiaire Z. C'est la raison pour laquelle le résultat de l'action de X est mieux décrit par la phrase (14c) que par (14d). L'action de *transmettre*₁ effectuée par X est plus nettement focalisée sur le second actant Y (*les lettres*), ce qui explique pourquoi on préférera rendre compte du transfert décrit dans (23) par la phrase résultative (23b) plutôt que par (23a).

- (14) c. J'ai un bonbon anglais.
 d. ? Un bonbon anglais m'est donné.
- (23) a. Les architectes ont les lettres de Khrouchtchev.
 b. Les lettres de Khrouchtchev sont transmises aux architectes.

Parmi les 92 achèvements sans copolysème statif, 62 se caractérisent par une forte transitivité sémantique. Cette propriété de leur sens lexical pourrait ainsi expliquer l'absence d'extension polysémique vers un verbe d'état.

4. POUR CONCLURE... AVANT DE POURSUIVRE

Les résultats de notre recherche permettent de proposer des réponses aux questions posées dans l'Introduction. Il s'avère que les relations de copolysémie partant d'un achèvement vers un verbe d'état sont attestées dans 56,6 % des cas examinés : 120 verbes d'achèvement donnent lieu à une extension polysémique stative. Dans le cas des 43,4 % des vocables analysés, nous n'avons pas repéré de verbe d'état dans la structure polysémique : 92 achèvements n'ont pas de copolysème statif.

On note ensuite que, pour les verbes analysés, la polysémie achèvement-état s'organise selon trois principaux patrons sémantiques. Le modèle de phases relationnelles est de loin le plus fréquent : l'achèvement de base dénote une entrée en relation, alors que son copolysème d'état renvoie au fait, pour X, de maintenir un certain type de rapport avec Y (cf. sous-section 3.1.1). Une autre configuration sémantique caractérise la polysémie régulière des verbes de communication (3.1.2). Enfin, il existe des verbes d'achèvement qui, par un mécanisme métaphorique, conduisent à des copolysèmes statifs désignant des états physiques ou psychiques. Notre étude a permis d'identifier ces trois modèles généraux, mais un travail ultérieur est nécessaire pour affiner cette répartition.

Si les achèvements donnant lieu à un sens statif se répartissent bien en (au moins) trois classes, leurs sens lexicaux ont néanmoins en commun une caractéristique importante. En effet, nous avons mis en avant une forte corrélation entre le faible degré de transitivité sémantique liée aux caractéristiques participatives de l'achèvement de base et la présence d'une acception stative dans la structure polysémique du vocable. *A contrario*, 67,4 % des achèvements sans copolysème statif sont fortement transitifs sur le plan sémantique.

Ces résultats autorisent à formuler l'hypothèse selon laquelle le degré de transitivité sémantique qui caractérise le sens lexical d'un verbe d'achèvement n'est pas indifférent pour l'apparition d'un verbe d'état parmi ses acceptions copolysèmes. La probabilité d'avoir une extension stativale est très faible, voire inexistante, lorsque l'achèvement dénote une action agentive, orientée vers un patient et susceptible d'opérer un changement d'état de celui-ci, c'est-à-dire lorsqu'il tend vers le pôle fort de transitivité sémantique.

Comme tout modèle, les patrons que nous avons dégagés sont une représentation simplifiée d'un phénomène autrement plus complexe. Si le présent article s'arrête là, il n'en est pas autant de notre recherche et de nos analyses. De nombreuses questions apparues au fil de ce travail méritent d'être explorées. Il sera notamment intéressant de proposer ensuite des analyses plus fines des modèles discutés, en y incluant une description plus approfondie des structures argumentales des prédicats.

BIBLIOGRAPHIE

- Apothéloz, D., 2008, *Entrer quelques instants vs arriver quelques instants : le problème de la spécification de durée de l'état résultant*, *Verbum*, XXX 2-3, p. 199-219.
- Apresjan, Ju., 1974, Regular polysemy, *Linguistics*, 142, p. 5-32.
- Bertinetto, P. M. & Lenci, A., 2012, Pluriactionality, habituality, and gnomic imperfectivity, in R. Binnick (ed), *Oxford Handbook of Tense and Aspect*, Oxford, Oxford University Press, p. 852-880.
- Cadiot, P. & Visetti, J.-Y., 2001, *Pour une théorie des formes sémantiques : motifs, profils, thèmes*, Paris, Presses universitaires de France.
- Croft, W., 2012, *Verbs. Aspect and Causal Structure*, Oxford, New York, Oxford University Press.
- Croft, W. & Cruse A. D., 2004, *Cognitive Linguistics*, Cambridge, Cambridge University Press.
- Dowty, D. R., 1979, *Word Meaning and Montague Grammar*, Dordrecht, Reidel.
- Dubois, J. & Dubois-Charlier, F., 1997, *Les Verbes français*, Paris, Larousse.
- François, J., 2010, L'étude de la polysémie verbale entre dérivation et invariance, in F. Neveu, V. Muni Toke, J. Durand, T. Klingler, L. Mondada & S. Prévost (ed), *Congrès Mondial de Linguistique Française – CMLF 2010*, Paris, Institut de Linguistique Française, p. 21-52.
- François, J., Sénéchal, M., 2008, Le sémantisme propre des cadres prédicatifs et la polysémie des verbes de production de parole, *Faits de langue*, 31/32, p. 325-335.
- Grimshaw, J., 1990, *Argument Structure*, Cambridge (Mass.), London, MIT Press.
- Hopper, P. J. & Thompson, S., A., 1980, Transitivity in Grammar and Discourse, *Language*, 56/2, p. 251-299.
- Kleiber, G., 2008, Petit essai pour montrer que la polysémie n'est pas un sens interdit, in J. Durand, B. Habert & B. Laks (ed), *Congrès Mondial de la Linguistique Française – CMLF'08*, Paris, Institut de Linguistique Française, p. 87-101.
- Langacker, R. W., 1987, Mouvement abstrait, *Langue française*, 76, p. 59-76.

- Lenci, A., 1995, The semantic representation of non quantificational habituals, in P. M. Bertinetto, V. Bianchi, J. Higginbotham, & M. Squartini (ed), *Temporal reference, aspect and actionality. Semantic and syntactic perspectives*, vol. I, Turin, Rosenberg and Sellier, p. 143-158.
- Levin, B., 1993, *English Verb Classes and Alternations: a Preliminary Investigation*, Chicago, The University of Chicago Press.
- Levin, B. & Rappaport Hovav, M., 2005, *Argument Realization*, Cambridge, New York, Cambridge University Press.
- Lux-Pogodalla, V. & Polguère, A., 2011, Construction of a French Lexical Network. Methodological Issues, in: *Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011. An ESSLLI 2011 Workshop*, Ljubljana, p. 54–61.
- McCawley, D., 1976, Remarks on What Can Cause What, in: M. Shibatani (ed.) *The Grammar of Causative Constructions*, New York, Academic Press, p. 117-129.
- Martin, F., 2011, Revisiting the Distinction between Accomplishments and Achievements, *Cahiers Chronos*, 22, p. 43-64.
- Mel'čuk, I., Clas, A. & Polguère, A., 1995, *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve : Duculot.
- Mittwoch, A., 1991, In Defence of Vendler's Achievements, *Belgian Journal of Linguistics*, 6, p. 71-85.
- Mourelatos, A., 1981, Events, Processes and States, in P. Tedeschi & A. Zaenen, (ed), *Syntax and Semantics. Tense and Aspect*, 14, New York, Academic Press p. 191-212.
- Parsons, T., 1985, Underlying events in the logical analysis of English, in E. Lepore & B. McLaughlin (ed), *Actions and events: perspectives on the philosophy of Donald Davidson*, Oxford, Blackwell, p. 235-267.
- Piñón, Ch., 1997, Achievements in an Event Semantics, in A. Lawson & E. Cho (ed), *Proceedings of Semantics and Linguistic Theory*, 7, Ithaca, New-York, CLC Publications, Cornell University, p. 273-296.
- Polguère, A. & Sikora, D., 2013, Modèle lexicographique de croissance de vocabulaire fondé sur un processus aléatoire, mais systématique, in C. Garcia-Debanc, C. Masseron & Ch. Ronveaux, (ed), *Enseigner le lexique*, Namur, Presses universitaires de Namur, p. 35-63.
- Sénéchal, M. & Willems, D., 2007, Classes verbales et régularités polysémiques : le cas des verbes trivalenciels locatifs, *Langue française*, 153, p. 92-110.
- Smith, C., 1991, *The parameter of aspect*, Dordrecht, Boston, London, Kluwer.
- Steedman, M., 2012, Computational Linguistics, in R. Binnick (ed), *The Oxford Handbook of Tense and Aspect*, Oxford, New York, Oxford University Press, p. 102-120.
- Vendler, Z., 1967, Verbes and Times, *Linguistics in Philosophy*, p. 97-121.
- Verkuyl, H.J. 1972, *On the Compositionnal Nature of Aspect*, Dordrecht, Reidel.
- Wilkins, D. P., 1996, Natural Tendencies of Semantic Change and the Search for Cognates, in M. Durie & M. Ross (ed), *The Comparative Method Reviewed. regularity and Irregularity in Language Change*, New York, Oxford, Oxford University Press, p. 264-304.
- Willems, D., 2007, Typologie des procès et régularités polysémiques, in D. Bouchard & I. Evrard (ed), *Représentations du sens II*, Bruxelles, De Boeck, p. 162-177.

II. Sémantique lexicale et modélisation du sens

Cette partie est de loin la plus volumineuse tant pour ce qui est du nombre de pages que pour la quantité des articles réunis. Ils couvrent une période de plusieurs années, en témoignant ainsi d'un intérêt constant et renouvelé pour le sens des unités du lexique qu'elles soient lexémiques ou phraséologiques. On y trouvera, en [9], l'article de 2008 sur les verbes *teżknić* en polonais et *manquer* en français, ré-étudiés dans les chapitres 4 et 5 du volume de synthèse de mes travaux de recherche.

Les travaux de cette seconde partie forment trois axes de recherche. Le premier d'entre eux réunit les études sémantiques des unités lexémiques, souvent dans une perspective comparée, avec en filigrane, des préoccupations didactiques en enseignement de langues étrangères. Cet intérêt pédagogique ne se manifeste pas encore dans des conclusions proposant une conceptualisation des unités du lexique en tant qu'objet de l'enseignement ou une réflexion sur les démarches didactiques, mais plutôt dans l'identification des problèmes et des difficultés auxquels celles-ci auront à répondre dans la pratique de classe de langue. L'article consacré à la combinatoire lexicale est une analyse des composants du sens lexical responsables des associations de lexèmes dans une collocation.

Le deuxième domaine d'intérêt évoqué dans cette partie est représenté par deux études des locutions. Ces travaux s'intéressent à la fois aux aspects sémantiques et syntaxiques de ces phrasèmes non compositionnels. Ils constituent une première approche des thèmes récurrents de mes recherches actuelles sur la phraséodidactique, présentées dans la troisième partie de ce volume.

Les sept articles du premier axe ([8] à [14]), avec les deux études des unités

phraséologiques du deuxième, montrent une orientation méthodologique de plus en plus marquée vers l'approche néo-structuraliste des réseaux lexicaux.

Enfin, les trois dernières études rapportent ([17] et [18]) des propositions de description et de modalisation des informations lexicales largement inspirées de mon travail dans le cadre du projet *Relief – RL-fr*. L'introduction au volume des *Cahiers de Lexicologie*, co-rédigée avec A. Polguère ([19]), est consacrée spécifiquement au rôle et à la place de la définition, *ipse eo*, d'une paraphrase du sens dans une structure de données complexe telle qu'un réseau lexical. Mon article paru dans ce même numéro se trouve dans la troisième partie du présent volume, car il aborde la question de définition lexicale dans sa dimension applicative en didactique des langues.

Étude du sens des unités du lexique

- [8] Pourquoi revient-on à ses moutons? Regard croisé sur la sémantique du retour en français et en polonais. *Studia Romanica Posnaniensia*, 33, 2006, 199-218.
- [9] Éléments d'une sémantique pour la description d'un sentiment. Approche contrastive franco-polonaise. *Arena Romanistica*, 1, 2008, 124-142.
- [10] *Arriver* et *venir* – quand la deixis fait (ou ne fait pas) la différence. *Pratiques*, 141/142, 2009, 138-149.
- [11] Le verbe COURIR et ses extensions sportives. *Études de linguistique appliquée*, 165, 2012 95-106.
- [12] NATION et NARÓD : analyse sémantique d'un couple de faux-amis franco-polonais. Dans : Degani, M. Frassi, P. et M. I. Lorenzetti (éds) *The Languages of Politics/La politique et ses langages*, vol. 1, Newcastle upon Tyne : Cambridge Scholars Publishing, 2016, 209-232.
- [13] Peut-on formaliser l'excentricité? Quand une description sémantique s'organise en réseau. *Cahiers du Littoral*, à paraître.
- [14] Combinatoire lexicale et polysémie : le cas de manquer. *Les Cahiers du Dictionnaire*, Classiques Garnier, à paraître.

Sens des unités phraséologiques

- [15] Identification d'unités phraséologiques et équivalence sémantique dans la traduction. Dans : Corpas Pastor, Gloria (ed.) *Computerised and Corpus-*

based Approaches to Phraseology : Monolingual and Multilingual Perspectives/Fraseología computacional y basada en corpus : perspectivas monolingües y multilingües. Geneva : Tradulex. [ISBN 978-2-9700736-8-0], 2016, 218-231.

- [16] Polysémie et homonymie dans un modèle lexical. *5^e Congrès Mondial de Linguistique Française, SHS Web of Conferences*, 27, 2016, 05012 (avec M.-S. Pausé)

Description et modalisation du sens lexical

- [17] Définir le sens dans un réseau lexical. *Studia Romanica Posnaniensia*, 39, 3, 2012, 63-79.
- [18] Modélisation de connaissances lexicales dans une ressource, *Revue Française de Linguistique Appliquée*, XXII, 1, 2017, 61-74.
- [19] Introduction du numéro thématique *La définition*. Dans : Polguère, A., Sikora, D. (dir) «La définition», *Cahiers de Lexicologie*, 109, 2, 2016, 9-11 (avec A. Polguère).

DOROTA SIKORA-POUIVET

Université Nancy 2

POURQUOI REVIENT-ON À SES MOUTONS ? UN REGARD CROISÉ SUR LA VALEUR SÉMANTIQUE DU RETOUR EN FRANÇAIS ET EN POLONAIS

Abstract. Sikora-Pouivet Dorota, *Pourquoi revient-on à ses moutons ? Un regard croisé sur la valeur sémantique du retour en français et en polonais* [Why "revient-on à ses moutons"? A cross look at the semantic value of the concept of "return" in French and in Polish]. *Studia Romanica Posnaniensia*, Adam Mickiewicz University Press, Poznań, vol. XXXIII : 2006, pp. 199-218. ISBN 83-232-1643-6, ISSN 0137-2475.

French verbs *revenir*, *rentrer*, *retourner* refer to the same kind of location situations as their Polish « doubles » *wrócić/wracać*. This paper analyzes their meanings through examples presented with their contextual data. *Revenir* and *rentrer* are morphologically analysable as complex lexemes derived from the corresponding *venir* and *rentrer* by adding the prefix *RE-*. *RE-* derivation provides a semantic feature of iterativeness to these complex items. From a synchronic point of view, *retourner*, in the sense investigated in this article, cannot be considered anymore as *RE-* derived, but it lexicalizes similar values. *Wracać/wrócić* share some of them. Some differences in the lexical meaning of all these verbs are explained, especially by showing how they are used as pragmatic indexes inside the relation between the speaker and a reference to a location.

Pourquoi revient-on à ses moutons plutôt que de retourner les voir ? La question semble frivole, mais l'intérêt d'une réponse, tant linguistique que pédagogique, est certain. Quantité d'étudiants en Français Langue Étrangère trébuchent depuis toujours sur les emplois de *revenir*, *rentrer* et *retourner*. Ces verbes sont ressentis, plutôt que compris, comme des synonymes ou, du moins, comme des équivalents discursifs. Dans cet article, nous nous proposons d'observer le fonctionnement de ces items dans différentes situations de communication afin de le comparer ensuite à celui du couple polonais *wrócić/wracać*.

Dans un premier temps, nous allons nous arrêter sur les trois verbes français mis parfois dans le même panier pour cause de ressemblance. Une réflexion sur leurs propriétés morphologiques, largement inspirée par l'article de Denis Apothéloz sur les différentes manifestations d'itérativité liées au préfixe *RE-* (voir Apothéloz 2005), devrait permettre de saisir d'une part en quoi consiste le faux air de famille entre ces items lexicaux, et d'autre part de repérer ce qui les distingue. Le

préfixe *RE-* y est pour beaucoup, à tout le moins en ce qui concerne *revenir* et *rentrer*. Considéré comme foncteur appliqué à la base de dérivation ou à certains de ses éléments, il détermine en grande partie des effets de sens produits par le mot construit. Nous entendons montrer par la suite qu'en choisissant, dans une situation de discours, le mot « juste », adéquat à ce que nous voulons dire, il convient de prendre en compte non seulement les paramètres du référent, mais également de se situer soi-même, en tant que locuteur, par rapport au déroulement du procès. Il s'agirait d'un phénomène déictique. Une troisième partie sera consacrée au couple polonais *wracać/wrócić*¹, dont la sémantique sera comparée à celle de ses « cousins » français. Dans notre démarche, nous adoptons résolument une perspective référentielle. Les exemples sont accompagnés d'une description aussi fidèle que possible de la situation d'énonciation. C'est le seul moyen, nous semble-t-il, de ne pas passer à côté des paramètres contextuels², essentiels pour le maniement des verbes qui nous occupent ici.

1. UN FAUX AIR DE FAMILLE ET QUELQUES FAITS MORPHOLOGIQUES

1.1. Certains énoncés tels que

- (1) Je rentre à la maison.
- (2) Je retourne à la maison.
- (3) Je reviens à la maison.

pris en dehors de toute situation d'énonciation, peuvent apparaître comme équivalents. De là, il n'y a plus qu'un pas pour conclure à la synonymie de *rentrer*, *revenir* et *retourner*. Il est vrai que, dans ce cas précis, ils lexicalisent les mêmes acceptions que les verbes polonais *wracać/wrócić*. Cependant, quand il s'agit de rendre compte d'un événement, nous ne pouvons nullement nous fier à cette prétendue synonymie. Elle se trouve rapidement battue en brèche car la situation extra-linguistique dont nous voulons parler force notre choix lexical. Il n'est alors pas indifférent d'employer l'un ou l'autre. Imaginons un parent polonais, en période de vacances, profitant de la plage en compagnie de son enfant. Cédant à ses demandes répétées, il lui donne une pièce et l'autorise à aller s'acheter une glace. Il lui dit alors :

- (4) Idź, kup sobie loda i wracaj.

Une substitution pure et simple conduit à des résultats pour le moins curieux. On obtient ainsi :

¹ Faute de place, nous ne pouvons nous lancer ici dans une analyse de différences aspectuelles et temporelles des doublets polonais.

² Nous parlerons ici de contexte au sens large du terme, qui ne se limite pas à l'environnement linguistique de l'énoncé, mais comprend toutes les informations (linguistiques, pragmatiques, socio-culturelles) à la disposition du locuteur pour interpréter le discours (Voir Moeschler J. & Reboul A., 1994).

- (5) Va l'acheter une glace et reviens.
 ? (6) Va l'acheter une glace et rentre.
 ? (7) Va l'acheter une glace et retourne (ici).

(6) et (7) paraissent décidément inadéquats à la situation même si c'est pour des raisons différentes. Nous les analyserons dans la section 2.

1.2. En français, *revenir*, *rentrer* et *retourner* présentent une structure morphologique similaire. Tous les trois relèvent du même paradigme dérivationnel par préfixation en *RE-*³. Dans la mesure où il s'agit d'une classe dérivationnelle des plus productives, on aboutit à une multitude d'items lexicaux en *RE-* aux significations multiples et variées. On aurait tendance à voir dans *RE-* un morphème polysémique. En effet, peut-on soutenir le contraire lorsqu'on songe à *refaire*, à *ranimer* ou encore à *rappeler* ?

Apothéloz (2005) défend justement la thèse opposée en proposant de réduire cette polysémie souvent encombrante. L'auteur examine le préfixe *RE-* en tant que foncteur opérant sur un argument *X*. Cet argument n'est autre que la base de dérivation dans la totalité de sa structure actancielle ou bien l'un de ses éléments. *RE-* n'est pas polysémique, bien au contraire. Appliqué à son argument *X*, il apporte à sa base de dérivation l'élément « à nouveau » en produisant un sens itératif. Celui-ci peut ensuite, selon le cas donner lieu soit à une interprétation itérative, soit à une interprétation annulative. Il y aurait ainsi un sens propre au préfixe *RE-*, avec deux effets de sens possibles.

1.2.1. Si elle se laisse résumer par une seule formule « à nouveau *X* », la valeur itérative de *RE-* peut s'interpréter de différentes manières selon ce qui tombe sous la portée du foncteur. En suivant le raisonnement et les analyses de l'auteur, nous allons à présent regarder de plus près ces différents cas de figure. Soit le verbe *reprendre* et les énoncés suivants :

- (8) Je reprends de ton côté, il est excellent.
 (9) Il ne comprend rien à ce que je dis. Reprends tout dès le début, moi, j'abandonne !
 (10) Le rétablissement a été très rapide, il reprend le mois prochain.

L'itérativité véhiculée par *RE-*⁴ peut s'interpréter de trois façons différentes selon la portée du foncteur. En paraphrasant (8), on dira : « à nouveau *A* prends *x* ». Le foncteur opère ici sur tous les éléments de la structure actancielle, l'agent comme le patient, ce qui donne lieu à l'itération du procès. (9) présente un deuxième type

³ Nous suivons la convention adoptée dans Apothéloz (2005) : le morphème-type *RE-* est transcrit en petites capitales, alors que ses différentes réalisations *re-*, *ré-* et *r-* seront notées en minuscules.

⁴ Comme le remarque l'auteur, en interprétant les dérivés en *RE-*, il faut bien se garder de confondre la valeur sémantique du foncteur et les multiples facettes de son argument, c'est-à-dire une polysémie éventuelle de la base de dérivation. Si le *reprendre* de (8) a un sens différent de celui de (9) ou (10), ce n'est pas à cause de *RE-*, mais justement à cause de *prendre*.

d'itérativité. Dans cet énoncé, le foncteur porte sur un seul paramètre x (en l'occurrence *tour*), élément de la structure actancielle de l'argument X -prendre. Cela signifie que « à nouveau il y aura x », en sachant que $x = \text{une/des explication(s)}$. Ce qui se reproduira, ce n'est pas le procès intégral, mais un événement qui en fait partie. L'agent de (10) reste inchangé, c'est lui qui reprendra, i.e. répètera quelque chose, une activité quelle qu'elle soit. Nous sommes donc face à une itération d'activité qui est la troisième interprétation possible d'itérativité présentée par *reprendre*. Lorsque nous avons affaire à un dérivé en *RE-* dans son acception itérative, nous présupposons qu'une action similaire ou au moins l'un de ses paramètres, exprimés par le verbe-base de dérivation a déjà eu lieu. « *A reprend x* » signifie qu'auparavant « *A a déjà pris x* » : j'avais déjà pris du rôti, quelqu'un a déjà « pris » tous les arguments, il avait déjà « pris » une activité au moins une fois dans sa vie.

1.2.2. Il en est autrement lorsque *reprendre* (ou un autre dérivé en *RE-*) véhicule une valeur annulative, comme c'est le cas de (11) et (12)

(11) « Rendez-moi, dit-il, mes chansons et mon somme. Et reprenez vos cent écus » (La Fontaine).

(12) Le Royaume Uni a repris la Présidence du Conseil de l'Union européenne.

Dans ce type d'emplois, *reprendre* signifie « *prendre pour entrer à nouveau en possession de quelque chose qu'on avait donné* ». Dans son acception annulative, *reprendre* est en relation d'antonymie directionnelle avec *donner* (11) ou *laisser, abandonner* (12), dont il annule les états résultatifs : le foncteur *RE-* porte sur l'objet censé changer d'emplacement et d'appartenance. Pour qu'il y ait valeur annulative, on doit donc présupposer l'accomplissement préalable de ce que Denis Apothéloz appelle un « *procès transformateur (...) directionnellement inverse de celui dénoté par le verbe base* »³.

1.2.3. Arrêtons-nous un instant sur les critères qui permettent de distinguer les dérivés en *RE-* des formes qui, malgré une ressemblance de surface, n'en sont pas (ou plus). Il s'agit de faire la différence entre les items morphologiquement construits (comme *reprendre, renaître, remonter, redire*) et ceux que, dans une perspective synchronique actuelle, nous devons traiter comme entièrement lexicalisés (par exemple *remplacer, rassasier, rallier, rencontrer*). On ne peut plus calculer leur sens à partir des apports du préfixe et de la base. A la différence des premiers, ces derniers ont perdu leur compositionnalité. Leur sens effectif s'étant éloigné de celui que l'on pourrait déduire de leur structure, ils sont mémorisés comme lexèmes sémantiquement autonomes.

Les lexèmes construits par préfixation *RE-*, au contraire, gardent leur diagrammaticité et, par conséquent, leur sens est plus ou moins prédictible à partir des éléments formels qui les composent. Rappelons, avec Apothéloz (2005), les trois con-

³ Apothéloz, 2005, p. 61.

ditions que doit remplir un item lexical en *RE-* pour qu'on puisse le considérer comme construit :

- (i) il doit contenir dans son sémantisme, à quelque niveau d'analyse que ce soit, l'élément « à nouveau » ;
- (ii) cet élément « à nouveau », lié à l'apport sémantique du préfixe foncteur *RE-*, doit porter sur un argument relevant de la structure sémantique de la base de dérivation ;
- (iii) le dérivé doit être hyponyme de sa base de dérivation.

2. *REVENIR, RENTRER, RETOURNER* EN TANT QUE LEXEMES CONSTRUITS (OU NON ?) EN *RE-*

Dans cette section, nous nous proposons de regarder de plus près ces trois verbes. Nous allons vérifier s'il correspondent aux critères établis par Apothéloz et, si tel est le cas, quelle est, pour chacun d'entre eux, la portée du foncteur *RE-*. Ces renseignements devraient jeter une lumière sur l'éventuelle valeur itérative et/ou annulatrice de ces items.

2.1. Le cas de *revenir*

2.1.1. Y a-t-il un composant « à nouveau » dans la structure sémantique de *revenir* ? La question est, on l'a compris, de savoir si l'item satisfait à la condition (i) évoquée ci-dessus. Soit les exemples (13) et (14) :

(13) Toni Braxton *revient* bientôt.

La diva soul américaine sera bientôt de retour avec un nouveau (ric ?) opus et un nouveau single. (presse)

(14) La Coccinelle *revient* ! (affiche de cinéma)

(13) comme (14) se laissent réduire à une forme minimale « *A revient* », intuitivement paraphrasable par « *A vient à nouveau* ». La condition (i) est ainsi satisfaite. Dans l'univers du discours, les deux énoncés présupposent valide l'information « *A est déjà venu auparavant* » et appellent de ce fait une interprétation itérative. Nous reprendrons dans la section suivante la réflexion sur la nature de cette itérativité (s'agit-il de l'itération du procès, de l'événement ou d'activité ?) en essayant de délimiter la portée du foncteur *RE-* dans chacun de ces exemples. Il convient d'abord de reprendre celui que nous avons déjà évoqué dans la section 1. Peut-on trouver ce même composant « à nouveau » dans (5) ?

(5) Va t'acheter une glace et reviens !

Pour éviter les difficultés liées à la forme injonctive de (5), que nous n'avons pas la possibilité d'examiner ici, appelons un autre exemple où *revenir* s'interprète de manière similaire et dont la forme assertive rend la démonstration plus claire :

(15) *Je suis revenu chez Wanadoo. Je viens de quitter Neuf Télécom, son deux mégas et son téléphone illimité à 24,90€ pour revenir à mon bon vieux 512 extense sans bonus au même tarif* (écrit électronique).

Dans ces emplois, (5) – abstraction faite de sa forme injonctive – comme (15) peuvent être ramenés à l'énoncé minimal « *A revient* » mais celui-ci appelle une interprétation différente. Nous pouvons la formuler de façon suivante : « *A vient pour être à nouveau dans le lieu qu'il a auparavant quitté* ». On y retrouve l'élément « *à nouveau* » imposé par la condition (i). L'information présupposée valide dans l'univers du discours est donc : « *A était parti* » ou bien « *A avait quitté un lieu P* », ce qui permet de conclure à un effet de sens annulatif. *Partir, quitter* ou encore *s'en aller*, antonymes lexicalisés de *revenir* dans les exemples (5) et (15), dénotent le procès transformateur dont nous avons parlé plus haut.

2.1.2. Conformément à la condition (ii), le foncteur *RE-* avec sa valeur « *à nouveau x* », dont nous avons constaté la présence dans *revenir*, doit opérer sur un argument relevant de la structure sémantique de la base de dérivation. *RE-* porte soit sur la structure dans sa totalité (l'exemple 8), soit sur l'un des paramètres qui la compose (9). Si nous voulons comprendre les différentes interprétations possibles de *revenir*, il nous faut repérer les paramètres liés à la valence de *venir* susceptibles de tomber dans le champ d'action du foncteur.

2.1.2.1. Les verbes transitifs expriment une relation qui s'établit entre l'agent et le patient ou l'objet du procès qu'ils dénotent. Les verbes de déplacement mettent en rapport un agent (ou dans d'autres cas, une entité) et un point de référence spatiale *P*. Le premier de ces arguments, que Langacker (1990)⁶ appelle « *mover* » a un statut particulier : il peut être en même temps agent et patient du déplacement. Décrire un déplacement suppose de prendre également en compte le facteur temporel (*T*). Fillmore (1975)⁷ explique ainsi le rapport entre tous ces paramètres : « Nous disons que quelque chose est en mouvement – par mouvement – j'entends un déplacement – lorsque cet objet est localisé dans un endroit à un moment donné et dans un endroit différent à un autre moment ». Certains verbes de mouvement – poursuit – il, comme ceux qui nous intéressent ici, et notamment *venir*, désignent un mouvement délimité. Ils supposent un lieu *P*₁, point de départ où le déplacement commence au moment *T*₁, et un lieu *P*₂, point final où il prend son terme au moment *T*₂. Ces paramètres relatifs au référent, c'est-à-dire à l'événement auquel renvoie l'énoncé, peuvent apparaître explicitement sous forme de circonstants :

⁶ Voir Langacker R., 1990, p. 236.

⁷ Fillmore Ch. J., 1975, p. 51 : « We say of something that it has moved, in the « locomotion » sense of movement that I have in mind, if it is at one location at one time and at another location at another time ». Cette définition est tout à fait satisfaisante si, comme Fillmore, on s'intéresse au mouvement délimité, en laissant de côté, par exemple les rotations des corps terrestres et célestes.

- (16) Il est clair que depuis qu'il est arrivé *de Metz* (P_1) pour environ 3ME en juin 2000 (T_1), ses statistiques laissent rêveur (écrit électronique).
- (17) Il est arrivé *à la gare de Tokyo* (P_2) avec trois minutes de retard (écrit électronique).
- (18) *Il était 10h00* (T_1) quand nous sommes parés voir une cascade dans les bois et la grotte Notre Dame de Lourdes (écrit électronique).

Ces énoncés renvoient à un déplacement extra-linguistique. Dans le même texte, Fillmore remarque que l'acte d'énonciation peut, lui aussi, être considéré comme un déplacement, un transfert de message entre des interlocuteurs. Liés à la situation d'énonciation, les paramètres de ce transfert sont alors identifiables à l'intérieur de celle-ci. Le message est codé par le locuteur en un lieu (que nous allons noter P_{en1} pour marquer la différence d'avec le P_1 de référence), à un moment T_{en1} qui caractérisent ensemble « le départ » du message. Celui-ci est ensuite décodé par son destinataire en un lieu P_{en2} au moment T_{en2} . Il arrive souvent, surtout dans un échange oral qui se déroule en tête à tête, que le codage et le décodage se suivent immédiatement de telle façon qu'il suffit de parler du temps de l'énonciation (du message) T_m en seul et même lieu P_m . Or, il existe des verbes qui codent différentes configurations de paramètres référentiels et déictiques à la fois sans recours aux circonstants. C'est le cas de *venir* qui nous intéresse ici en tant qu'argument de *RE-*.

Imaginons un café avec son lot d'habitues dont l'un a pendant un moment manqué à l'appel. Ses comparses pourront dire avec un brin de reproche bien compréhensible :

- (19) Il n'est pas venu depuis un bon moment.

Ajouter un complément « au café » serait redondant. *Venir* n'en a pas besoin pour orienter l'interprétation de (19) vers le lieu où se termine le déplacement, c'est-à-dire vers le paramètre P_2 du référent. Le locuteur et le destinataire peuvent se trouver à bonne distance l'un de l'autre ($P_{en1} \neq P_{en2}$) et, s'il est écrit, le message pourra être lu à un moment T_{en2} distinct de T_{en1} ($T_{en1} \neq T_{en2}$). Dans ce cas précis, la situation d'énonciation, avec son P_m et son T_m , importe peu, du moment que P_2 est présent dans l'univers du discours. L'exemple (20) montre néanmoins que *venir* exprime quelquefois des configurations de paramètres relatifs à l'acte d'énonciation.

- (20) Salut les vieux... *je viens à Grenoble* du 12/05 au 15/05, y'a moyen de s'organiser un pot ?????? (forum électronique)

Dans cet emploi, *venir* est en effet suivi d'un circonstant, mais en même temps, il véhicule l'information capitale pour l'interprétation de (20) : Grenoble, c'est P_{en2} . *Venir* localise ainsi les destinataires du message et le lieu où celui-ci sera décodé. Cette information disparaîtra, si nous remplaçons *venir* par un autre verbe :

- (21) Salut les vieux... *je vais à Grenoble* du 12/05 au 15/05, y'a moyen de s'organiser un pot ??????

L'information concernant le déplacement, c'est-à-dire l'évènement auquel réfère (21), n'est pas différente de (20), mais *Grenoble* n'y a pas le statut de P_{act} .⁴

2.1.2.2. Revenons (si je puis dire) à l'exemple du café évoqué plus haut, qui constituera de nouveau notre univers du discours. Le compagnon perdu de vue réapparaît, fait auquel on pourra référer par (22) :

(22) Il est enfin revenu.

RE- avec sa valeur « à nouveau » porte ici sur la totalité du schéma actanciel « *A* vient à P_2 » constitué par l'agent et la variable P_2 identifiable à l'intérieur de l'univers du discours. En d'autres termes, l'énoncé (22) exprime une itération du procès. Nous obtiendrons un résultat comparable en appliquant *RE-* au venir de l'exemple (20) :

(23) Salut les vieux... *Je reviens à Grenoble* du 12/05 au 15/05, y'a moyen de s'organiser un pot ??????

à une différence près. Dans l'interprétation de (23), à la place de l'argument *P*, nous trouverons le paramètre déictique P_{act} , relatif à l'emplacement du destinataire du message au moment du décodage. L'information présupposée valable dans l'univers du discours peut être formulée de manière suivante : « *A* est déjà venu à Grenoble ». Nous admettons cependant volontiers que, pris en dehors de tout contexte (au sens large du terme), et cela malgré la spécification temporelle du circonstant, (23) est ambigu. À côté de l'interprétation itérative, il peut prêter à une glose annulative : « *A* vient à nouveau pour être dans un lieu (*Grenoble*) qu'il avait auparavant quitté » et l'information présupposée valide dans l'univers du discours sera alors « *A* avait quitté *P* ». C'est aussi dans cette interprétation itérative, qu'il faut prendre l'exemple (5) dont nous parlerons ultérieurement.

Le foncteur *RE-* adjoint à *venir* conduit, comme nous venons de le voir, à des effets de sens itératifs et annulatifs. Il semblerait cependant que seuls deux des trois types d'itérativité apparaissent sous son action. Le cas de l'itération du procès nous paraît attesté par (22) et (23). L'itération de l'activité nous semble également possible : c'est cet effet de sens que l'on pourrait accorder aux exemples (13) et (14) ci-dessus. On peut les paraphraser de manière suivante « à nouveau *A* vient ». Le foncteur *RE-* porte seulement sur « *A* vient », et non pas sur P^0 .

2.1.3. Pour clore provisoirement ces réflexions autour du verbe *revenir*, il convient de remarquer qu'il satisfait également à la troisième condition évoquée à

⁴ Pour éviter de compliquer notre exposé, nous laissons (ici) de côté le jeu de ces paramètres autour du personnage considéré comme central dans les textes narratifs.

⁵ Dans la mesure où il s'agit de communiqués à caractère publicitaire, on peut se demander si le destinataire du message n'est pas à chaque fois celui ou celle qui le lit (en P_{act} au moment T_{act}) auquel cas, le paramètre P_{act} se trouverait également dans le champ du foncteur et nous serions face à une itération du procès.

la fin de la section 1. Tant dans ses acceptions itératives, que cette itérativité concerne le procès ou l'activité, que dans son interprétation annulative, *revenir* est bel et bien un hyponyme de sa base de dérivation, c'est-à-dire de *venir*. Les gloses que nous avons proposées montrent que grâce à l'apport sémantique de *RE-*, *revenir* véhicule une spécification supplémentaire par rapport à son hyperonyme. Nous pouvons donc le compter parmi les mots construits du lexique résultant du schème dérivationnel en *RE-* et, par conséquent l'analyser en ces termes.

2.2. Le cas de *rentrer*¹⁰

Devant l'initiale vocalique nasale de *entrer*, *RE-* apparaît sous sa forme *r-*, l'un de ses allomorphes. Ce n'est d'ailleurs pas sa seule variante auprès de *entrer*, puisqu'on note aussi des formes en *re-* et en *re-*.

2.2.1. Le cas de *rentrer* confirme l'observation d'Apothéloz (2005), selon laquelle une spécialisation sémantique accompagne les variations allomorphique du foncteur *RE-*. Ainsi, *rentrer* relève d'un micro-système d'opposition lexicale composé de couples de dérivés avec une même base. C'est notamment le cas de *raccompagner* vs *reaccompagner*, de *racheter* vs *reacheter*, etc. Les premiers membres de ces paires véhiculent surtout une valeur annulative et éventuellement itérative, alors que leurs doublons en *re-* expriment systématiquement une itération¹¹. Nos observations concernant *entrer* et ses dérivés en *RE-* vont dans le même sens. L'exemple (24) nous semble éloquent à ce titre :

(24) Les bureaux de direction ? Il faut ressortir d'ici et *rentrer* par la porte principale.

L'interprétation que l'on donnera à (24) ne va pas de soi. On peut la concevoir comme un effet de sens itératif avec une paraphrase telle que : « il faut que *A* entre à nouveau dans *P* ». Cette interprétation ne sera cependant acceptable qu'à une seule condition : on doit la considérer comme une itération du procès. *Entrer* et *rentrer* doivent avoir le même paramètre *P*. Dans le cas (24), ce sera bien évidemment non pas la porte, mais le bâtiment, c'est-à-dire l'entité dont la ou les portes incriminée(s) font partie. En effet, (25) paraît difficilement, voire pas du tout acceptable :

? (25) Les bureaux de direction ? Il faut ressortir et *rentrer* dans le bâtiment derrière.

La contrainte d'un seul et même *P*, qui pèse sur (24) est à la base de l'ambiguïté que présente cet énoncé. À côté de l'interprétation itérative que nous venons de proposer, une lecture annulative est également possible : « *A* doit entrer pour être à nouveau dans *P* ». Pour lever l'ambiguïté, le locuteur peut recourir à *re*¹², allomorphe du préfixe *RE-*, à valeur itérative :

¹⁰ Dans le cadre de cet article, seuls les emplois intransitifs de *rentrer* et de *entrer* nous intéressent. Aussi n'allons nous pas nous arrêter sur les emplois tels que *rentrer sa voiture*.

¹¹ Voir Apothéloz (2005), p. 64-67.

¹² Sans prétendre ériger cette observation au rang de système, il semblerait que l'allomorphe *re-* à valeur itérative soit plus fréquent quand il s'agit de *entrer* dans les emplois intransitifs : « Si vous

(26) Une fois dans le parc, peut-on ressortir et *ré-entrer* ?

Oui, vous pouvez quitter le parc, un tampon sera mis sur votre main pour pouvoir *ré-entrer* dans la même journée. (site électronique d'un parc Walibi)

D'ores et déjà, nous pouvons constater que les conditions (i) et (iii) sont satisfaites. *Revenir* et *ré-entrer* sont des mots construits appartenant au même paradigme dérivationnel en *RE-*. Ils produisent bien des effets de sens itératif et annulatif, qui à un niveau d'analyse ou un autre comprennent l'élément « à nouveau » (i). Par ailleurs, leurs sens apporte des spécifications supplémentaires par rapport à celui de leur base de dérivation, ce qui nous permet de conclure à une relation d'hyponymie que *rentrer* et *ré-entrer* entretiennent avec *entrer* (iii). Avant de vérifier si la condition (ii) est également remplie, il nous paraît intéressant de s'arrêter sur la question des types d'itérativité véhiculés par *rentrer*.

Dans le cas de *rentrer* intransitif, comme d'ailleurs dans celui de *revenir*, la structure actancielle met en jeu les mêmes paramètres. Dans les exemples que nous avons vus jusqu'ici, l'agent du déplacement est en même temps le patient ou le bénéficiaire, c'est-à-dire respectivement celui qui le subit ou qui en profite. Il paraît donc difficilement imaginable de séparer les deux. Logiquement, le foncteur *RE-* ne saurait porter uniquement sur le patient, sans opérer sur l'agent. Une itération de l'événement, comme dans l'exemple (9), semble impossible en ce qui concerne *revenir* et *rentrer*. Et pourtant certains emplois de ce dernier conduisent bel et bien à cet effet de sens. L'exemple (27) montre bien de quoi il s'agit, si nous l'envisageons à travers une situation d'énonciation et un univers du discours, c'est-à-dire à travers des caractéristiques pragmatiques qui accompagnent le message. La première pourrait, pour les besoins de notre démonstration, se résumer de cette façon : *A* et *B* sont nos interlocuteurs, tour à tour auteur et destinataire des énoncés qui composent l'échange (27). Pour compléter l'image, précisons que *A* sonne pour la première fois de sa vie à la porte de *B*. Comme on l'a entendu maintes fois, *B* peut alors dire en ouvrant :

(27) *B* : *Reentre* !

A : Une autre fois volontiers. Là, je n'ai pas le temps d'*entrer*.

De toute évidence, *A* et *B* réfèrent à une même situation extra-linguistique. Les items différents qu'ils emploient à cette fin ne peuvent s'expliquer qu'à l'intérieur de leurs univers du discours respectifs. Dans la cas de *A*, il s'agit bien d'une première réalisation de l'activité consistant à entrer³³. Dans son univers du discours, *B* n'a pas la même conceptualisation des choses. Compte tenu du paramètre P_{in} du message, c'est-à-dire de l'endroit où il l'émet, il envisage l'action consistant à entrer dans son appartement comme un événement reproduit plus ou moins souvent

supprimez vos cookies, vous devez *re-entrer* tous vos mots de passe de connexion et de messagerie = (écrit électronique).

³³ S'il compte la reproduire dans certaines conditions, il pourra toujours recourir à *ré-*.

par des agents différents. L'invitation à entrer dans (27) accepte, sans rien enlever à l'hospitalité de *A*, la paraphrase : « à nouveau il y a une entrée ».

2.2.2. Suivant la condition (ii), le foncteur *RE-* opère sur un ou plusieurs paramètres de la structure sémantique de *entrer*. Sa portée détermine ensuite l'une des trois interprétations itératives ou bien annulatives à donner. Nous avons vu, dans la section 2.1, que les paramètres discursifs de *venir* sont également présents dans *revenir*. Si un prédicat comme *aimer* établit une relation entre un agent et un patient, les verbes de déplacement construisent un rapport entre « cible l'entité localisée et site le point de repère par rapport auquel sa situation est fixée »¹⁴. Il nous semble que le sens du terme *site* employé par Vandeloise n'est pas sensiblement différent de ce que nous désignons par *P*. S'il ne permet pas de faire la différence entre les types de *P* distingués par Fillmore, sa définition n'en saisit pas moins la nature du rapport entre ces paramètres des verbes de déplacement. Le verbe lui-même impose certaines contraintes quant à *P*. Aussi *venir* comme *revenir*, orientent-ils la cible vers P_2 , que ce soit P_2 ou P_{act} , la seule façon pour eux de référer à P_1 étant de recourir, sur le plan syntaxique, à un complément de lieu. *Entrer*, quant à lui, tout en se focalisant sur P_2 , fait peser des contraintes supplémentaires :

- (28) Ce matin, je suis entré dans la bibliothèque-fouillis dont j'ai l'habitude (écrit électronique).
 (29) Je suis entré au club de Valenciennes comme responsable administratif (écrit électronique).
 (30) Depuis le jour où je suis entré sur le marché du travail, je n'ai accompli que des boulots vains, sans intérêt réel (écrit électronique).

Certaines prépositions, par exemple *vers* n'introduisent quasiment jamais le complément de lieu de *entrer*. C'est parce que celui-ci, par son sémantisme, conceptualise *P* en tant qu'entité bi ou tridimensionnelle délimitée dans l'espace, de manière à ce que l'agent puisse, au terme du procès, se retrouver à l'intérieur de celle-ci. En d'autres termes, *entrer* impose une relation contenu-contenant entre l'agent (ou la cible, si l'on préfère) et *P*, le site, qu'il soit référentiel ou pragmatique¹⁵. On pourrait, certes, objecter que du point de vue matériel, (30) contrevient à ce que nous venons de dire : le marché du travail n'est pas – heureusement pour les demandeurs d'emploi – spatialement délimité. Nous répondrons alors, en suivant l'étude de Vandeloise (1987) consacrée à la préposition *à*, que ce rapport locatif peut s'étendre, par métaphore, à des sites institutionnels et à la routine qui leur est associée. Ils sont alors conceptualisés comme entités capables de contenir ou d'inclure, indépendamment de leur caractère spatial ou non dans la réalité. La contrainte reste valide lorsque le paramètre *P* tombe dans la portée de *RE-*. Nous tenons donc la réponse à la question de savoir pourquoi, l'exemple (6) est inacceptable dans la

¹⁴ Vandeloise C. (1987), p. 77.

¹⁵ Borillo qualifie cette relation de rapport d'inclusion (voir Borillo, 1998).

situation d'énonciation décrite dans la section (1). Si, dans l'exemple (5), on ne peut pas substituer tout bonnement *rentrer* à *revenir* :

- (5) Va t'acheter une glace et reviens.
 ? (6) Va t'acheter une glace et rentre.

Dans cet emploi, *revenir* s'interprète de façon annulative : « viens pour être à nouveau à l'endroit que tu quittes ». Le site auquel réfère le locuteur-auteur de (5) est celui où il se trouve au moment de prononcer le message décodé aussitôt par le destinataire au moment T_m . *RE-* porte donc sur le paramètre P_m de venir qui est, certes, un point de référence spatiale, mais qui n'est pas conceptualisé comme contenant. *Revenir* et *rentrer* redeviennent substituables dès qu'une telle conceptualisation est possible. Nous supprimerions sans hésiter le point d'interrogation de (6), si P désignait une maison, un appartement ou ne serait-ce qu'une tente.

2.3. Le cas de *retourner*¹⁶

2.3.1. Malgré un air de famille morphologique, *retourner*, contrairement à *revenir* et à *rentrer*, ne peut pas (plus ?), dans la synchronie actuelle, être compté parmi les mots construits en *RE-*. Certes, l'élément « à nouveau » est facile à y trouver, comme le montrent les exemples (31) et (32) :

- (31) Je suis revenu de Minsk
 Paris/varsovie en avion
 ensuite varsovie/minsk en train
 aucun problème dans mon hôtel
 je suis resté 2 semaines
 je pense y retourner bientôt (frédéric, Paris, écrit électronique, orthographe de l'original)
- (32) Crespo retourne à Chelsea (dépêche AFP)

Dans l'exemple (31), *retourner* produit un effet de sens itératif que l'on peut gloser par « y aller à nouveau ». Le schème actanciel est reproduit dans sa totalité, ce qui permet de conclure à une itération du procès. Dans (32), nous avons une valeur annulative paraphrasable par : « A retourne à l'endroit qu'il avait auparavant quitté ». Le contexte situationnel de l'énoncé confirme cette interprétation puisque la même dépêche de l'Agence France-Presse précise que le footballeur-vedette avait été prêté pendant un an au Milan AC.

2.3.2. La présence du composant « à nouveau » ne suffit pas pour compter *retourner* parmi les construits en *RE-*, alors que les conditions (ii) et (iii) ne sont pas

¹⁶ Dans le cadre de cet article, seuls les emplois intransitifs de *retourner* paraphrasables par *aller* à nous intéressent. La place nous manque pour en faire un exposé complet, mais disons, en résumé, que *retourner* remplit les trois conditions, et donc est bien un mot construit dans les exemples suivants : (a) « Il se *retournait* dans tous les sens sans trouver le sommeil » ou (b) « Ce n'est pas la première fois qu'elle a *retourné* sa veste. » Le foncteur opère alors sur la base (*se*) *tourner* en produisant des effets de sens intransitifs (a) ou annulatifs (b).

remplies. Les valeurs itérative et annulatrice de (31) et de (32) ne sauraient être considérées, dans une perspective synchronique actuelle, comme issues de *RE-* opérant sur la base *tourner*, ni sur aucun de ses paramètres. Ils ne sont pas non plus hyponymes de *tourner*, ni lorsque celui-ci désigne un mouvement rotatif, ni quand il s'agit d'un changement de direction. Si relation d'hyponymie il y a, c'est plutôt par rapport à *aller*.

Les conditions (ii) et (iii) n'étant pas remplies, nous ne pouvons pas (plus ?) voir en *retourner* un construit en *RE-*. S'il l'a été, à un moment ou un autre de sa diachronie, il a perdu sa compositionnalité et ne peut plus être analysé en termes des apports sémantiques relatifs au préfixe et à la base de dérivation. Il a subi un processus de lexicalisation en codant dans son signifié ce que nous interprétons aujourd'hui comme effet de sens itératif et annulatif.

2.3.3. Cependant, ce constat ne permet pas d'expliquer pourquoi, malgré les effets de sens apparemment comparables, *revenir* et *retourner* ne sont pas substituables, par exemple dans (5) et (7). En cherchant la réponse, observons d'abord une régularité syntaxique, qui à notre avis, constitue un indice important. Or, *retourner*, contrairement à *revenir* et *rentrer*, appelle systématiquement un complément de lieu soit sous forme d'un syntagme nominal, soit sous une forme pronominale ou adjectivale indexicale (*ici* dans l'exemple 7). Ce complément, présent au niveau syntaxique, est un complément directionnel indiquant P_2 , c'est-à-dire le lieu où le mouvement auquel réfère *retourner* prend sa fin. Or, dans les cas de *revenir* et de *rentrer* une telle complémentation apparaît quelquefois sur le plan syntaxique, mais elle n'est pas systématique. Rien d'étonnant à cela, puisque, rappelons-le, ces verbes orientent l'interprétation vers un P_2 identifié dans l'univers du discours et/ou vers un P_{act} lié au locuteur au moment d'énonciation (exemple 5) ou bien au locuteur au moment où le message est supposé lui parvenir (exemple 23). Si *retourner* appelle systématiquement un complément directionnel, c'est que, de par son sémantisme, il ne renvoie pas automatiquement au paramètre P_2 quel qu'il soit. Notre hypothèse est que, au contraire, le paramètre central de *retourner*, c'est P_1 . L'exemple (33) va tout à fait dans ce sens :

(33) Discovery *retourne* dans l'espace. (titre de presse)

puisqu'il est possible de le paraphraser intuitivement par : « *A* quitte P_1 pour aller à nouveau à/dans/sur P_2 ». Dans ce cas, il semblerait que les deux paramètres P_1 entrent en jeu : le P_1 du référent, c'est-à-dire le lieu initial où le déplacement commence, et le P_{act} relatif à la position de l'auteur de l'énoncé. C'est d'ailleurs ce dernier qui se révèle décisif. Remplaçons, dans l'exemple (33), le verbe *retourner* par *revenir*. Nous avons alors l'énoncé (34) :

(34) Discovery *revient* dans l'espace.

qui est difficilement acceptable, sauf ... s'il est énoncé par un membre de l'équipe à bord de la Station internationale qui signifiera ainsi « *A vient à nouveau à/dans/sur P_{ind}* », c'est-à-dire à l'endroit où se trouve le locuteur.

Si avec *revenir* et *rentrer*, nous accordons un statut privilégié à P_2 , l'entrée en jeu de P_1 introduit un élément nouveau dans les interprétations de *retourner*. Puisque nous faisons référence, dans un même énoncé à P_1 et à P_2 , nous ne pouvons ignorer la distance qui les sépare, même si, dans la plupart, elle n'est pas des cas précisée. L'état résultatif prévu du procès se laisse résumer de manière suivante : « *A n'est plus à P_1 , A est à P_2* ». Pour *revenir* et *retourner*, on dira seulement : « *A est à P_2* ».

2.3.4. Un dernier point mérite, à notre avis, l'attention, puisqu'il s'avère décisif pour la distribution de *retourner*. Dans la section 2.2.1, nous avons constaté, en ce qui concerne *revenir* et *rentrer*¹⁷, que l'agent du déplacement en est en même temps patient ou bénéficiaire. En d'autres termes, le sujet grammatical renvoie à l'agent. Ce n'est pas toujours le cas pour *retourner* :

- (35) Tu peux prendre mon dictionnaire, il est sur l'étagère, mais après, il doit *retourner* à sa place.

Le sujet grammatical de (35) renvoie non pas à l'agent du déplacement mais à son bénéficiaire. Nous tenons ainsi un autre élément de réponse à la question pourquoi la substitution pure et simple des trois verbes n'est pas toujours possible.

3. LE CAS DE *WRÓCIĆ*/*WRACAĆ*

Ce couple aspectuel est sémantiquement proche des verbes français que nous venons d'analyser. Si, dans une approche synchronique actuelle, on analyse ces deux verbes d'un point de vue constructionnel, ce n'est pas en tant que mots construits, mais uniquement dans la mesure où ils servent de base de dérivation par préfixation en *po-* et *od-*. En d'autres termes, *wrócić* et *wracać* sont des items lexicalisés, au sens où ils codent un signifié qui n'est pas calculable à partir de leur structure morphologique. Ils appartiennent au lexique mental des usagers de la langue qui y accèdent par voie mémorielle.

3.1. Une itération est-elle possible ?

Les énoncés dont l'un des deux est le prédicat rendent compte du même type d'événement comprenant un déplacement. Ils partagent la plupart des valeurs de *revenir*, *rentrer* et *retourner*, sans pour autant être leurs miroirs fidèles. Pour un francophone, le maniement de *wracać*/*wrócić*, qui pourrait paraître plus simple pour

¹⁷ Nous pouvons, certes, dire que la voiture *rentra* dans le garage, mais nous allons alors référer à un rapport locatif dimensionnel et non pas à un déplacement.

des raisons quantitatives (les valeurs relatives à trois unités sont à répartir entre deux items), présente cependant quelques embûches. (36) a été énoncé par un francophone parlant polonais. Ajoutons que le locuteur est un provincial effectuant des séjours occasionnels dans la capitale.

(36) Niedługo wrócimy do Paryża.

Si l'on ne peut rien reprocher à sa grammaticalité, il n'en est pas pour autant plus adéquat à la situation. Compte tenu de P_1 , point de référence spatiale relatif au locuteur, c'est-à-dire au lieu initial de ce déplacement, différent de P_2 , son terme, (36) devait être la traduction de :

(37) Nous allons bientôt retourner à Paris.

dans son acception itérative, effet de sens qui n'est pas celui que produisent *wracać* ou *wrócić*. Le procès susceptible d'être reproduit se présente sous forme d'un schéma simple (fig. 1) :



Or, pour exprimer une itération du procès, le polonais recourt à une forme adverbiale adjointe au verbe :

(37') Niedługo znowu pojedziemy do Paryża.

Il existe cependant des emplois de *wracać/wrócić* que l'on pourrait très bien interpréter comme itératifs. C'est le cas de (38) et (39) :

(38) Raz przyjedziesz – wrócisz na pewno.

(Une fois venu, tu reviendras à coup sûr – slogan publicitaire sur le site internet de la ville de Dźwirzyno, petite station balnéaire au bord de la Baltique).

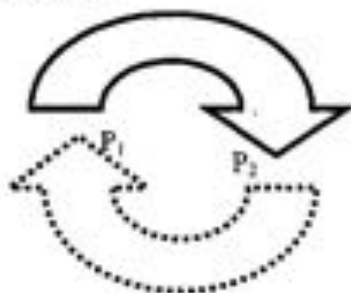
(39) Atmosfera w mieście już się uspokoiła, trwa sprzątanie i myślę, że turyści mogą tu już do nas wracać. (Dans la ville, l'atmosphère s'est rassérénée, on est en train de nettoyer et je pense que les touristes peuvent déjà revenir ici, chez nous – journal télévisé, relation en direct de Shams el Sheikh par la représentante d'une agence de voyages résidant sur place, peu après les attentats d'août 2005).

Une interprétation itérative ne saurait être exclue à la légère. On dirait alors que *wrócić* (38) et *wracać* (39) signifient ici *przyjechać/przyjeżdżać ponownie*. Nous pourrions ainsi ramener ces énoncés à des formes plus simples, exprimées en français par « venir à nouveau ». Il nous semble cependant que, une fois de plus, il faut prendre en compte les éléments contextuels tels que le statut du locuteur, la convention sociale liée aux messages publicitaires, les habitudes culturelles (séjours de vacances). Dans ce contexte, une interprétation annulative n'est pas moins justifiée. Elle s'exprime par la paraphrase « venir pour être à nouveau à P_2 ». S'agit-il

de « venir à nouveau » ou plutôt de son état résultatif « être à nouveau à P_2 »? La seconde interprétation nous semble avoir tout pour elle, ce qui nous fait pencher pour un effet de sens annulatif. Nous reconnaissons néanmoins une certaine ambiguïté due, à notre avis, au rapport entre cause et effet qui relie les deux procès.

3.2. *Wrócić/wracać* et la valeur annulative

Les effets de sens produits par *wracać* et *wrócić* sont donc, malgré quelques cas douteux, à classer comme annulatifs. C'est cette valeur que le couple lexicalise et c'est cette interprétation qu'un destinataire polonophone donnera à (36). Dans l'univers du discours, il présuppose valide non pas l'information selon laquelle (a) « A est déjà allé auparavant de P_1 à P_2 », mais celle que l'on peut intuitivement énoncer sous la forme (b) : « A avait auparavant quitté P_2 pour un P_1 ». Il active ainsi un tout autre schéma (fig. 2) :



Le mouvement de P_2 vers P_1 que représente la flèche en pointillés, correspond à l'information présupposée (b). Il s'agit du procès transformateur qui, dans l'univers du discours, a précédé l'événement auquel réfère le verbe annulatif, en l'occurrence *wracać/wrócić*, représenté dans notre schéma (fig.2) par la flèche du haut. Ils ne peuvent s'interpréter que par rapport à leur antonymes discursifs tels que *odejść*, *wyjechać*, *wyjsść* (équivalents de *partir*, *sortir*, *quitter*,...). Ces derniers, relatifs à un déplacement en direction opposée, n'ont pas besoin d'être énoncés ; il suffit qu'ils soient présupposés.

Wrócić/wracać privilégient le P relatif au terme du déplacement. Cependant, comme le laisse voir le schéma ci-dessus, ce P final présente une particularité : en tant que point de référence spatiale, il est en même temps le lieu initial du procès exprimé par l'antonyme discursif, celui-là même qui se rapporte au procès transformateur présupposé. C'est ainsi que nous allons interpréter l'exemple (4) de la section 1. C'est également en fonction de ce statut privilégié du P final, que (40) ne peut être compris que dans un sens annulatif.

(38) Po chwili *wrócił*. Pomyślałem, że musi mieszkać bardzo blisko.

Le P final du déplacement est identifié par rapport à la position du locuteur au moment de référence et éventuellement au moment d'énonciation. Puisque nous ne

disposons d'aucune information concernant les caractéristiques dimensionnelles de P_1 , pour la traduction, il nous faut recourir à *revenir* :

(40') Elle est *revenue* au bout d'un moment. Je me suis dit qu'elle devait habiter tout près.

3.3. *Wrócić* et *wracać* versus *retourner*

3.3.1. Les propriétés actancielles de *wrócić* et *wracać* d'un côté, de *retourner* de l'autre, présentent une certaine ressemblance. Dans le rôle de leur sujet grammatical, nous trouvons aussi bien l'agent du déplacement, comme dans les exemples (40) et (41) que son bénéficiaire comme dans l'exemple (42).

(41) Tym razem wrócić z wakacji opalona i wspaniale wypoczęta (*Cette fois, je reviendrai de vacances bronzée et merveilleusement reposée – pressée*)

(42) Codex Iuris Lubusciensis powiniem wrócić do Głubczyc (= *Le CIL devrait retourner à Głubczyce – presse*).

3.3.2. Une conceptualisation différente

Un énoncé avec *wrócić* ou *wracać* dans le rôle du prédicat peut spécifier P_1 par un complément de lieu, comme cela a lieu dans (41), mais les deux verbes restent orientés vers P_2 . C'est ce paramètre qu'ils ont en commun avec *revenir* et *rentrer*, même si, contrairement à ce dernier, en polonais aucune contrainte n'est imposée quant aux caractéristiques dimensionnelles de P_2 . Qu'en est-il donc de P_1 dans le couple *wrócić/wracać*? Encore une fois, des éléments contextuels qui accompagnent la situation d'énonciation seront d'une grande utilité. L'exemple qui suit a été puisé dans la vie quotidienne. Sur la porte fermée d'une boutique, il n'est pas rare de tomber sur l'écriteau dans le genre de (43) laissé par employé qui a dû s'absenter :

(43) Zaraz wracam.

Laissons de côté les références temporelles concernant le message pour n'envisager que les paramètres spatiaux. Dans (43), *wracać* fait référence à P_2 , qui est à la fois le terme du déplacement de l'agent-locuteur et le lieu où son message parviendra à l'allocutaire. Dans une situation similaire en France, nous avons trouvé (44) sur la porte d'une boutique de photocopies.

(44) Il est 10 heures 30. Je reviens dans 20 minutes.

Quelle que soit la raison de cette défection, le locuteur-agent se trouve à un endroit P_1 , différent de P_2 , où son mouvement annulatif commencera. Au moment de se mettre en route pour rejoindre son lieu de travail, il dit en polonais :

(45) No dobra, muszę wracać/wrócić do sklepu

résolution louable que l'on exprimera en français par :

(46) Bon, il faut que je *retourne* au magasin.

Le verbe polonais *wracać* et son doublon perfectif *wrócić* rendent compte tous les deux des situations envisagées dans (44) et (46) pour lesquelles le français dispose de deux verbes différents. Il oriente l'activité vers son pôle final P_2 sans subir les

contraintes qui, en français, obligent le locuteur à choisir entre *revenir* et *rentrer*, ce dernier étant marqué par les caractéristiques dimensionnelles et le rapport d'inclusion imposé à P_2 . Une autre différence réside dans la conceptualisation des paramètres situationnels en jeu lorsqu'il s'agit de *retourner*. Contrairement au polonais, le français prend en compte la position initiale de *A* et/ou celle du locuteur par rapport au déroulement du procès. Peut-on pour autant parler d'une polarité initiale pour *retourner* ? Il nous semble que non. *Retourner*, comme ses « proches » sémantiques français et polonais à l'exception de *wracać* qui désigne une activité, est un verbe télique supposant, certes, une durée, mais porté essentiellement vers le résultat qui sera son aboutissement. D'un point de vue cognitif, un procès de déplacement suppose un scénario. En transposant cette idée de scénario dans un domaine théâtral, on pourra dire que les éléments de la scène (actancielle) prévus dans le script, se trouvent mis au premier plan grâce à un jeu de lumières savamment orchestré, alors que d'autres restent dans l'ombre. Dans le cas de *wrócił/wracać*, comme dans celui de *revenir* et *rentrer*, c'est P_1 qui est mis en avant. Avec *retourner* il n'est nullement perdu de vue mais on le regarde, en tant que *A* ou en tant que locuteur, par rapport à P_1 . La « gestion » de tous ces paramètres et le choix lexical qu'elle soutend se compliquent inévitablement quand le nombre de protagonistes augmente dans un texte ou dans un discours. La dépêche de l'Agence France-Presse dont nous avons donné le titre dans l'exemple (32), continue en introduisant des points de référence spatiale liés aussi bien à son auteur qu'aux deux protagonistes : Hernan Crespo et Peter Kenyon :

(32) Crespo *retourne* à Chelsea

(32') « Nous sommes ravis que Hernan *revienne* » a déclaré le directeur général du club londonien Peter Kenyon sur www.Chelseafc.com. « Il a réalisé une grande saison pendant son prêt à Milan (...) et il est heureux de *revenir* à Stamford Bridge... »

Le recours à *retourner* dans le titre, puis à *revenir* lorsqu'un autre protagoniste intervient, permet ainsi de localiser avec plus ou moins de précision tous les intervenants. En ce qui concerne le locuteur, cette localisation est négative : il ne se situe pas à Londres, et encore moins à Chelsea. Hernan Crespo, au moment d'énonciation, se trouve encore à Milan, et Peter Kenyon est bien chez lui, lieu qui signifie le terme du mouvement P_2 . Bien évidemment, il ne s'agit pas de la localisation réelle des protagonistes au moment du discours mais à leur positionnement par le locuteur par rapport au procès. L'analyse est simpliste, voire grossière, mais son unique but est de montrer la complexité de ce jeu des paramètres spatiaux que l'on doit affronter lors du processus d'apprentissage du français langue étrangère.

4. POUR RÉCAPITULER

Selon les langues, nous parlons d'un même événement ou état des choses sans forcément mettre en avant les mêmes paramètres. Pour un Polonais, choisir entre *revenir* ou *retourner* suppose de se situer par rapport à un paramètre P_1 qui, pour le

type de procès décrit, n'a pas la même saillance cognitive dans sa langue maternelle. Le risque est alors d'aboutir à un hiatus entre ce que Donnellan (1978) appelle la référence du locuteur (*what the speaker had in mind*) et la référence sémantique, c'est-à-dire le sens qu'a conventionnellement une expression dans une langue. Le sens descriptif conventionnel est sans doute plus facile à saisir dans une situation d'apprentissage. Or, pour manier correctement les verbes dont nous venons de parler, il faut maîtriser en plus leur sens instructionnel, c'est-à-dire celui qui indique les procédures indispensables pour identifier le référent.

Pour terminer, essayons de réunir dans un tableau les caractéristiques sémantiques et pragmatiques des verbes qui nous ont occupés jusque là.

| verbe valeur | wracać/wrócić | revenir | rentrer | retourner |
|---|-----------------------|---------|--|--------------------------------|
| valeur(s) itérative(s) | - | + | + | + |
| valeur annulative | + | + | + | + |
| point de référence spatiale visé | P_2 | P_2 | P_2 | P_2 envisagé depuis P_1 |
| contraintes sémantiques sur P_2 | - | - | rapport d'inclusion entre P_2 et x | - |
| cas profond du sujet grammatical | agent bénéficiaire | agent | agent | agent bénéficiaire |

Il nous reste à répondre à la question qui ouvre cet article. On revient à ses moutons dans un sens annulatif, vers un P_2 où il nous faut se trouver à nouveau. On y rentre pas, car le point de référence spatiale ne suppose pas un rapport d'inclusion (il ne s'agit pas de rentrer dans une bergerie). On ne peut pas retourner à ses moutons pour deux raisons. Premièrement pour cause d'agrammaticalité : à ses moutons, tout en étant un point de référence spatiale, indépendamment de ses emplois métaphoriques, ne saurait être un complément de lieu, le test en où faisant foi :

- (47) – Où voulez-vous revenir ?
– ? A nos moutons.

Deuxièmement, le procès annulatif est focalisé sur P_2 final, et non pas sur P_1 point de départ.

BIBLIOGRAPHIE

- Apothélos D. (2005), *RE- et les différentes manifestations de l'itérativité*, *Pratiques*, 125-126, p. 48-71.
- Apothélos D., Fekarek Doehler S. (2003), *Nouvelles perspectives sur la référence : des approches informationnelles aux approches interactionnelles*, *Verbum XXV*, 2, p. 109-136.
- Borillo A. (1998), *L'espace et son expression en français*, Ophrys, Paris.
- Donnellan K. (1978), *Speaker Reference, Descriptions and Anaphora*, P. Cole éd., *Pragmatics, Syntax and Semantics*, 9, 47-68.
- Fillmore Ch.J. (1975), *Santa Cruz Lectures on Deixis : Coming and Going*, Indiana University Linguistics Club.
- Kleiber G. (1999), *Problèmes de sémantique. La polysémie en questions*, Presses Universitaires du Septentrion.
- Langacker R.W. (1990), *Concept, image and symbol. The cognitive basis of grammar*, Mouton de Gruyter, Berlin - New York.
- Moeschler J. (1994), *Dictionnaire encyclopédique de pragmatique*, Seuil, Paris.
- Przybylska R. (2002), *Polisemia przyimków polskich w świetle semantyki kognitywnej. (Polysémie des prépositions polonaises dans une perspective sémantique cognitive)*, Universitas, Kraków.
- Slobin D. (2003), *Od « myśli i języka » do « myślenia dla mówienia »*. Trad. O. Kubínska, W. Kubínski. Dans : E. Dąbrowska, W. Kubínski (éd.), *Aktiwizacja języka w świetle językoznawstwa kognitywnego*. Wydawnictwo Universitas, Kraków, p. 361-402.
- Vandeloise C. (1987), *La préposition à et le principe d'anticipation*, *Langue Française*, 76, p. 77-111.

Éléments d'une sémantique de la description d'un sentiment¹

Dorota Sikora

Université de Nancy

UFR Langues et Cultures Étrangères

Département de Polonais et de Tchèque

« Le plus difficile, après mon départ de Pologne, fut de me transposer moi-même, ma personnalité, dans la façon américaine de voir et de saisir le monde. Notre personnalité est formée par notre culture et par notre langue. Or, pour moi, la langue et le monde avaient été séparés. Ce sentiment de séparation me pesait d'autant plus qu'il se faisait sentir à chaque pas, dans chaque situation. », Ewa Hoffman, *Let in Translation*.

En évoquant ainsi ses premières années outre-Atlantique, Ewa Hoffman, aujourd'hui critique littéraire et écrivain de langue anglaise, décrit ce qui est le lot commun d'un émigré, d'un étudiant boursier ou d'un touriste un tant soit peu curieux de contacts avec la population du pays qu'il visite : s'exprimer et se faire comprendre exactement comme il le veut. Le domaine des sensations et des sentiments est, de ce point de vue, particulièrement délicat. Choisir le mot juste qui permettra à l'interlocuteur de comprendre ce que l'on ressent est de loin plus difficile que de lui signifier son désir d'acheter du pain. Klein (1986) distingue deux types d'informations qu'un apprenant doit saisir lorsqu'il procède à l'analyse des énoncés de la langue cible. Il s'agit de l'information linguistique à proprement parler et de ce qu'il appelle l'« information parallèle », composée principalement de signaux visuels, tels que gestes, mimiques et éléments contextuels de la situation d'énonciation. Or, lorsqu'on parle de sentiments ou d'émotions, les signaux visuels sont naturellement déterminés (cf. Wierzbicka 2003) et, par conséquent, difficiles à identifier. Dans cet article, nous nous proposons de nous concentrer sur le premier type d'informations, celles qui sont transmises par les moyens dont dispose la langue. Sans prétendre faire le tour complet du problème, nous nous proposons de réfléchir sur un exemple précis, celui des verbes *spłóć* en polonais et *manquer* en français.

¹ Je tiens à remercier les/les/leurs anonymes. Leurs remarques ont permis d'améliorer ce texte. Elles sont tout utiles pour la suite de mes recherches.

Il s'agira de savoir quels sont les moyens employés par chacune des deux langues pour décrire une émotion qui, au départ, du moins le croyons-nous, est la même. Dans nos réflexions, nous avons utilisé uniquement des exemples authentiques et représentatifs, selon nous, de quelques dizaines d'émotions qui composent nos mini-corpus polonais et français. La première section sera consacrée à un tour d'horizon lexicographique pour vérifier quelles facettes du sens lexical des deux items sont habituellement saisies dans les définitions de dictionnaires. Dans la deuxième section, nous nous efforcerons de décrire *spisać* et *manquer* en tant que prédicats, c'est-à-dire de présenter leur modèle(s) de réalisation argumentale qui leur sont attachés et de comparer les structures syntaxiques sur lesquelles ils débouchent. Nous verrons que les propriétés aspectuelles liées aux deux prédicats tiennent aux différences sur le plan argumental et au niveau des fonctions sémantiques que l'un attribue aux arguments. Cette section se terminera sur une remarque de nature pragmatique concernant le verbe *manquer*. La troisième et l'avant-dernière section posera la question de l'organisation de l'information qu'imposent les structures argumentales et thématiques des deux verbes. De qui ou de quoi pouvons nous parler à l'aide de l'un et de l'autre ? Quelle est la saillance situationnelle des éléments mis en relation par les deux prédicats ? Toutes ces étapes nous conduiront, l'espérons-nous, à la quatrième et dernière partie, celle qui, pour conclure, devrait nous permettre d'expliquer au moins en partie les raisons de ce « sentiment de séparation » entre la langue et le monde évoqué par Ewa Hoffman.

Une dernière précision s'impose avant de commencer : nous sommes loin de défendre des thèses whorfianes de relativisme culturel ou de poser l'intraduisibilité des sentiments. Là où le sens de *spisać* ne peut pas être rendu par *manquer*, le français dispose, bien évidemment, d'autres moyens tels que les verbes *se languir de*, *s'ennuyer de*, *avoir le mal du pays*. Cependant, dans cet article, plutôt que de nous demander quand *spisać* est traduit par tel ou tel autre verbe français, nous avons voulu comprendre pourquoi il ne peut pas être considéré comme équivalent de *manquer* et vice versa.

1. D'un point de vue lexicographique

1.1. Le verbe polonais *spisać* fonctionne, d'un point de vue lexicographique, comme un lexème autonome. Il constitue à lui seul une entrée de dictionnaire.

Ainsi, Domaszewski (1967)², explique-t-il que *spoknić* signifie, dans une première acception, « ressentir du regret suite à l'absence de quelqu'un, au manque de contact avec quelqu'un ou avec quelque chose, souhaiter fortement voir une personne ou un objet ». Une deuxième acception de ce verbe élargit la portée en y ajoutant un élément supplémentaire, à savoir un désir de possession ou de réalisation d'un objet connoté en objectif : « souhaiter fortement obtenir, atteindre quelque chose, vouloir fortement quelque chose »³.

Une poignée d'exemples authentiques présentement nous les sens véhiculés par *spoknić*. Les traductions en français que nous en proposons doivent, en premier lieu, rendre compte du contenu sémantique de l'original, quitte à ne pas préserver la structure formelle à laquelle nous reviendrons plus loin.

- (1) Ale jak każdy tegdy tegdy jednak za domem, za rodziną (passe)
[Mais comme à tout navigateur, sa maison, sa famille lui manquent quand même]⁴ ou bien
- (2) Czy nie spoknię do takiego spokoja, szczęścia i dobrobytu ? (beaucoup religieux)
[N'aspirez-vous pas à une telle sérénité, à un tel bonheur et à un tel bien-être ?] ou bien
[Une telle sérénité, un tel bonheur et un tel bien-être ne manquent-ils pas ?]
- (3) Lato w Mińskowie jest przepiękne. A czasem spoknię tramwaj do Krakowa, ale nie tak bardzo już. Przypomni się całowick. (échange oral)
[L'été à Mińskowick est plus que beau. Quelquefois, Cracovie me manque tout de même, mais plus tant que ça. On s'habitué.]
- (4) Ja wyjeżdżam do dziewczyny i na 100% za komputere spoknię nie będę.
Pasa tym, jak bym miał spoknię za kompera, to bym się za odwykiem

2 Domaszewski, W. 1967. Słownik Języka Polskiego (Dictionnaire de la langue polonaise). Warszawa, Państwowe Wydawnictwo Naukowe. C'est l'auteur du présent article qui traduit les entrées citées.

3 Domaszewski, W. 1967. vol. IX, p.139.

4 On peut objecter que le français dispose du verbe *se languer de* qui permet de rendre compte de même type de situation, notamment en ce qui concerne l'exemple (1). Mais comme tout navigateur, il se languerait même de sa maison, de sa famille. Si nous avons choisi de le laisser de côté, c'est pour deux raisons : premièrement, les emplois qui nous intéressent ici, c'est à dire se morbidité, être malheureux de l'absence de quelqu'un, de quelque chose sont qualifiés par le TLF de *regretueux*, il semble également qu'ils nous obligent à quitter le registre dont on a l'habitude dans les échanges quotidiens. Nos réflexions sur Google sont dans le même sens : aucune occurrence pour le *se languer de* de monnaie, de une à quatre occurrences à peine pour le *se languer de* Paris et le *se languer de* soleil. Deuxièmement, comme le montrent nos exemples (3) et (4), il n'existe pas d'équivalence sémantique entre *spoknię* et le *se languer de*, mais cela, faute de place, devra faire l'objet d'une autre étude comparative.

scopladal (échange sur Internet)

(Je pars chez ma copine et je suis sûr à 100 % que mon ordinateur ne me manquera pas. D'ailleurs, si l'ordi devait me manquer, je commencerais à chercher une cure de désinvolte)

(3) Motna opikni! na tylo ritojych sposobow. Motna opikni! na slobotni, na pogodę, na głośni albo na mroczni. Jest opiknora na sposobem i ciesz albo na muzyką (échange sur Internet)

(5a) i. On peut manquer/ he languir de tant de manières différentes. ii. On peut se languir du soleil, du beau temps, des montagnes ou de la mer. iii. Il existe un sentiment de manquer par rapport à la sérénité et au silence ou à la musique/une nostalgie de la sérénité et du silence ou de la musique)

(5b) (On peut manquer de tant de manières différentes. Le soleil, le beau temps, les montagnes ou la mer peuvent [nous] manquer)

(5c) (On peut manquer de tant de manières différentes. On peut manquer de soleil, de beau temps, de montagnes ou de mer)

Dans (3), (5), (4) et (5), *opikni!* actualise la première acception, tandis que dans (2), il actualise la seconde. Les périphrases définissives adoptées par différents auteurs ne présentent pas de désaccord ni d'écart sensible et le statut de *opikni!* en tant que verbe psychologique est ressort clairement.

Les descriptions lexicographiques diffèrent cependant souvent en ce qui concerne les choix terminologiques dans leurs périphrases définissives. Avant de nous installer au verbe *manquer*, il nous semble donc utile de réunir les propositions lexicales de *opikni!* en une seule définition établie en termes de MSN. Cette démarche devrait nous permettre de repérer plus facilement les contrastes sémantiques entre les deux verbes. En revenant au métalangage de primitives sémantiques, Wierzbicka (1992) propose une gloss susceptible de reprendre ainsi les composants essentielles du sens de *opikni!*:

X *opikni* do/za Y³ →

(1) X thinks something like this:

(2) I am far away from Y

3. Le choix des éléments définissives pourrait être discuté, notamment en ce qui concerne le signe (2). Il nous semble que la référence à un bien-être personnel n'est pas une condition d'usage *non* qui non de *opikni!*. Nous ajoutons par ailleurs la possibilité de constructions prépositionnelles avec *za* que nous analyserons plus loin.

- (3) when I was with Y I felt something good
 (4) I want to be with Y now
 (5) if I were with Y now I would feel something good
 (6) I cannot be with Y now
 (7) because of this, X feels something bad⁶

Ajoutons par souci de précision que l'affect décrit par *tykwaś* peut également être conceptualisé dans une forme nominale *tykwaś*. Ses valeurs, telles qu'elles apparaissent dans les dictionnaires, reprennent les deux acceptions du verbe correspondant. *Tykwaś* peut ainsi désigner « un sentiment provoqué par la séparation d'avec quelqu'un ou d'avec quelque chose, doublé du désir de le retrouver » ou bien, par une restriction atmosphérique, « un désir intense d'obtenir quelque chose ». Dans les deux cas, l'existence d'un lexème nominal permet de désigner l'émotion ou le sentiment en question, « la *pep-chone* », dont parle Boucheard (1995), en tant qu'*entité* « tout comme *señle* ou *chalye* décrit un certain type d'*trouté* masculinelle, qu'elle que soient les différences de statuts ontologiques que nous sommes loin d'oublier.

1.2 En français

Le « *Treasure de la Langue Française informatisé* » (Le TLFi) consacre plusieurs pages à la description du, ou plutôt, des sens du verbe *manquer*. La définition lexicale est organisée autour de l'idée commune à tous les emplois répertoriés, à savoir celle d'absence. Analyser toutes les définitions lexicales que le TLFi en propose n'est pas l'objet de cet article. Nous allons donc nous concentrer sur celles qui se rapprochent de *apwəś*. Ainsi, l'idée d'absence se retrouve, effectivement, aussi bien dans :

- (6) Les enfants lui manquent, il les voudrait une semaine sur deux.
 (7) Les enfants me manquent et je leur manque.
 (8) Je suis espérié français à Bangkok et la France me manque parfois.
 (9) Serge Besson regrette cependant l'absence du demi d'ouverture du XV de la Rose, Jersey Wilkinson, blessé à la cuisse droite. « Oui, il me manque, a déclaré le Biarrot, je meerais fait une jute de pouvoir le rejouer pour la première fois depuis cette demi-finale perdue en Australie ».
 (10) Les enfants me manquent de respect.

⁶ See Wierzbicka, 1992: 121.

⁷ Szymczak, M. 1983, vol. 1, p. 202.

(10) J'aimerais beaucoup faire une formation, mais le temps et l'argent me manquent.

Le TLFi indique, en nos exemples le *confirmer*, que la construction *manquer à* peut signifier « faire cruellement défaut sur le plan affectif ». C'est justement cette interprétation psychologique que nous donnons à (6), (7) et (8). Inutilement, on lui ferait cependant à l'accorder à (9), même si, d'un point de vue formel, rien ne s'y oppose. En revanche, dans (10) et (11), nous avons affaire à des publics qui décrivent un rapport de possession que l'on pourrait qualifier de négative. Ce rapport est dépourvu de valeur psychologique. En effet, nous pourrions les paraphraser par *ne pas avoir* :

(10') Les enfants n'ont pas de respect pour moi.

(11') J'aimerais beaucoup faire une formation, mais je n'ai pas de temps ni d'argent.

Par souci de précision, ajoutons qu'une autre construction de *manquer*, celle suivie de la préposition *de*, peut aussi un sens psychologique. *Manquer de* quelque chose signifie, conformément au TLFi, « souffrir de son absence ». Cet emploi est cependant indubitablement rare. En effet, nos recherches sur Google ont débouché sur des sites consacrés aux chanteurs de musique populaire et aux paroles de leurs chansons. Par ailleurs, nous n'avons trouvé aucune occurrence d'une construction telle que :

Je manque de mes enfants.

ou

Je manque de la France

susceptible d'exprimer la même émotion que (6), (7) et (8).

Un dernier constat vient s'ajouter à ce survol lexicographique : dans les emplois qui nous intéressent ici, *manquer* (différents *manquer_{adv}*) n'a pas de nom de sentiment qui soit morphologiquement lié à lui. C'est d'ailleurs ce que nous a laissé supposer notre exemple (3) et nos tentatives maladroites de traduction dans la séquence (5a,14).

En recourant au métalangage proposé par Wierzbicka, nous pourrions décrire ainsi le sens de la construction *Y manque à X* :

(1) X thinks something like this :

(2) Y is not with me

- (3) when Y was with me I felt something good
 (4) I want Y to be with me now
 (5) if Y was with me I would feel something good
 (6) Y cannot be with me now
 (7) because of this, X feels something bad.

Le sens psychologique de *manquer_{ps}* est rendu par les éléments (3), (5) et (7) de la définition, c'est-à-dire ceux qui contiennent le prédicat primitif *feel*. Mais on remarquera aussi que cette même construction *Y manquer à X* peut très bien rendre compte de (1), exemple dans lequel *manquer* n'est pas un verbe psychologique.

Ces observations nous conduisent à deux constats :

1. en polonais, l'émotion qui nous intéresse est lexicalisée par un verbe ainsi que par un nom de sentiment et quelques adjectifs dérivés (*szczęśliwy, smutny, etc.*), tandis que
2. en ce qui concerne le français, *manquer* réalise un sens psychologique dans l'une de ses constructions syntaxiques possibles. Nous dirons donc que l'émotion en question est grammaticalisée.

Ces premières observations fournissent un élément de réponse à la question que nous nous sommes posée au début au sujet du sentiment de séparation que l'on peut ressentir entre le monde de ses propres émotions et l'environnement étranger. Or, habitués à l'exprimer dans notre propre langue, nous nous heurtons à la difficulté qui consiste à employer des moyens différents que la langue cible met à notre disposition. Entre le polonais et le français, il s'agit de passer d'un sentiment « lexicalisé » à un sentiment « grammaticalisé ».

2. Les structures argumentales et leurs réalisations syntaxiques.

La structure syntaxique, aussi bien dans le cas de *szczęśliwy* que dans celui de *manquer_{ps}*, peut être représentée par la même formule :

$$SN_1 V Prep SN_2$$

Dans les deux cas, il s'agit de verbes intransitifs avec un complément prépositionnel. Mais ces similarités de surface ou plutôt de façade n'en cachent pas moins des différences notables.

Nous ne tardons pas à le constater dès que nous changeons de niveau d'analyse et nous intéressons aux fonctions sémantiques des arguments sujets et

objets. En réalité, comme nous avons pu le constater en comparant les définitions en termes de primitifs sémantiques ci-dessus, *spénié* et *manquer_{sent}* relèvent de deux types différents de verbes psychologiques. Le premier est un verbe de type SubjExp : l'argument que nous trouvons à la place du sujet sémantique est un expérimenteur - sujet expérimentiel, c'est-à-dire celui en qui « se déroule l'expérience psychologique » (Bouchard, 1995 : 6) ou, selon les termes d'Anscombe, celui qui constitue « le lieu psychologique, le support du sentiment » (Anscombe 1995 : 43). *Manquer_{sent}*, au contraire, est un verbe de type ObjExp où le rôle thématique d'expérimenteur est réalisé par l'argument en position d'objet syntaxique, tandis que le sujet sémantique est le thème, l'objet de l'émotion.

L'organisation thématique de *spénié* et celle de *manquer_{sent}* entraînent les formes suivantes :

- spénié* (s)(y)
- spénié* (expérimenteur)(thème)
- manquer_{sent}* (y)(s)
- manquer_{sent}* (thème)(expérimenteur)

Cette pronomatation a des conséquences sur le plan aspectuel. De plus, elle n'est pas neutre pour la représentation de l'événement : de toute évidence, les deux prédicats en fournissent une description fort différente. Si les rôles sémantiques permettent, comme on le pense aujourd'hui, d'identifier les classes d'arguments qui entretiennent un même type de relation sémantique avec leur prédicat, l'étape suivante de notre réflexion consistera à nous poser la question du rapport entre l'organisation thématique des arguments et la description de l'événement réalisé par la prédication. Dans les deux sous-sections suivantes, nous nous intéresserons de près à la manière dont les deux arguments, l'expérimenteur et l'objet de l'émotion, profitent la relation psychologique entre eux par l'intermédiaire de leurs réalisations syntaxiques.

2.1 *Ţpănié* : les arguments et leurs réalisations syntaxiques

Ţpănié présente deux modèles de réalisation argumentale. Il peut apparaître avec un seul argument ou avec deux arguments. Dans aucun des deux cas de figure, il ne quitte la catégorie de verbes SubjExp.

2.1.1 *Ţpănié* en tant que prédicat à un argument

Un exemple de ce modèle de réalisation argumentale a été donné plus haut dans (5) :

- (5) *Măsimă Ţpănié na ryle rătărych spărobów.*
- (5a) *Ń. On peut manquer de tant de manières différentes.*

(12) en est un autre :

(12) *Tpikni pan ! Porwiedził skiniżnieta głowy*
[Vous marquez ? Il confirme d'un hochement de tête]

L'unique argument en position de sujet syntaxique est celui d'expérienceur. (5) et (12) diffèrent certes par la lecture générique que nous faisons dans le premier cas et une interprétation spécifique que nous attribuons à (12). Mais qu'il s'agisse de toute une classe d'individus, comme dans (5), ou d'un particulier de type (12), seule une lecture statique non-progressive est possible. Il s'agit d'un état, en l'occurrence psychologique, que nous asserions au sujet de l'expérienceur. Or, en parlant de noms de sentiment et d'attitude, Anscombe (1995) remarque que les sentiments sont joints aux individus, ce qui les rapproche des propriétés. Il nous semble que ceci est vrai également des verbes qui décrivent un état psychologique, comme celui de *tpikni* : ils attribuent une propriété, fix-elle accidentelle et transitoire, à l'expérienceur. On parlera, dans ce cas, de propriété extrinsèque, à la différence de propriétés intrinsèques à caractère permanent et continuatif pour un individu.

Anscombe propose par ailleurs un classement de noms de sentiment en deux catégories : les noms de sentiment endogènes, c'est-à-dire ceux « ayant une origine qui se confond avec le lieu psychologique » (1995 : 47), et les noms de sentiments exogènes, c'est-à-dire ceux dont l'origine est conceptuelle comme relative à l'expérienceur. Transposée aux verbes, cette typologie nous permet de classer *tpikni* comme un argument décrivant un sentiment endogène.

Le fonctionnement de *tpikni*, nom de sentiment correspondant, vient confirmer cette observation. Nous constatons en effet qu'il n'y a aucune résistance dans des constructions qui permettent d'exprimer un rapport de contenu à contenu ou celui de possession. Il s'agit d'une part des SN au génitif subjectif, c'est-à-dire ceux dont le génitif se substitue au sujet du verbe correspondant :

Ewa tpikni → *tpiknieta Ewy*
Enoncé prédicatif SN, V → Syntagme nominal : SN_{possesseur} SN_{possédé}

et d'autre part, des SN composés d'un adjectif possessif et d'un nom de sentiment :

Ewa tpikni → *jej tpiknieta*
Enoncé prédicatif SN, V → Syntagme nominal : Adj.Poss. SN₂

Ainsi, l'état psychologique auquel réfère *opóźnił* est une propriété accidentelle, conceptualisée comme endogène et séparable de son objet, dont l'expérimencé est à la fois l'origine et le déterminant.

2.1.2 *Opóźnił* est tout que prédicat à deux arguments

Dans les exemples (U) à (4), nous avons affaire au même prédicat, avec l'expérimencé en position de sujet et un deuxième argument réalisé, sur le plan syntaxique, par un syntagme prépositionnel de forme *Prép N*. La présence de ce complément indirect modifie sensiblement l'événement décrit. Nous ne sommes plus face à une propriété asserée au sujet de l'expérimencé, mais d'un événement qui est à décrire en termes de relation entre ses deux arguments.

Le syntagme prépositionnel que nous trouvons en position de complément indirect est la réalisation syntaxique de l'argument thème. D'un point de vue morphosyntaxique, les deux constructions prépositionnelles admises par *opóźnił* au *SV_{trans}* et de *SV_{inacc}* sont identiques à celles qui apparaissent dans les constructions spatiales dynamiques polonaises. Elles construisent alors des réalisations syntaxiques des arguments dont la fonction principale consiste à servir de point de référence spatial permettant de localiser l'argument « figure » arrivant en position de sujet syntaxique. Il existe une identité formelle entre les compléments prépositionnels de *opóźnił* et ceux qui, dans des constructions spatiales, indiquent à la fois la direction et le terme du déplacement. Aussi, dans (3) trouvons-nous le même complément argument de *Krakowa* que celui qui prend la place du complément argument de lieu* dans la valence d'un verbe de déplacement, par exemple *jechał* (aller par un moyen de transport).

(3) A czasami spóźniał transporty do Krakowa, ale nie tak bardzo jak.

[Quelquefois, Cracovie me manquait tout de même, mais plus tard que ça]

(U) Wczora miałem jechał do Krakowa

[Hier, je devais aller à Cracovie]

Grimshaw (1990) et bien d'autres chercheurs (Levin & Rappaport Hovav 2005) adoptent un postulat de départ selon lequel le ou les modifié(s) de réalisation argumentale d'un prédicat est/ont le reflet de son sens lexical. Dans cette optique, il faut supposer alors que, dans le cas de *opóźnił*, la présence d'un second argument est liée à une modification sémantique.

Tout d'abord, comme nous l'avons suggéré plus haut, le prédicat ne désigne plus un état-propriété, mais une relation qui est orientée d'un point de

* voir Grimshaw

vue directionnel. Elle va de l'expérencieur qui en est la source vers son thème. Une autre modification, liée cette fois à la valeur sémantique de la préposition, contribue à la description de l'événement. Or, selon Berthoinneau et Cadot (1993), les prépositions ont avant tout un sens instructionnel. Celui-ci s'actualise dans la prédication. Aussi bien *de* que *au*, permettent de localiser les deux arguments l'un par rapport à l'autre, en indiquant la direction : l'un vers l'avant, l'autre en arrière. Cet élément de séparation spatiale, introduit par le complément prépositionnel, organise la relation de la manière suivante :

(3) A crasmi opikniq trosneriq de Krakova, ale nie rak buadro jut.
 Trosnit de SV_{crasmi}
 (exp(thème))



Figure 1

Le complément prépositionnel de SV_{crasmi} apporte un élément spatio-temporel en indiquant que la séparation entre les entités représentées par les deux arguments consécutifs, pour l'expérencieur, à laisser le thème derrière lui. Cette relation est visualisée par la figure 2 :



Figure 2

Précisons que les figures 1 et 2 ne représentent que les modifications spatio-temporelles dues à la présence du complément prépositionnel. Celui-ci ne change rien cependant à l'organisation thématique du prédicat. L'expérencieur en position de sujet syntaxique reste l'argument prédominant, puisqu'il reste l'argument le plus saillant à la fois dans la hiérarchie des arguments où, selon Grimshaw, il prime sur le thème et dans l'organisation causale de la relation dont il est la source. De ce fait, il se rapproche des verbes agentifs. Le sujet expérencieur ne répond pas aux tests traditionnels d'agentivité, tels que celui du contrôle volitionnel. Mais il présente toutes les caractéristiques de ce que Rawet (1995) appelle *sujet intentionnel* : « [l]e sujet intentionnel est dirigé ou orienté vers un contenu ou objet » (Rawet, 1995 : 32). Une telle lecture devient effectivement possible lorsque, au lieu de désigner simplement une

relation d'ordre psychologique, *sybnit' de SN_{psych}* et *sybnit' za SN_{psych}* passent dans la catégorie de verbes d'attitude. Ils peuvent dès lors être considérés comme décrivant une activité, dans la mesure où un certain nombre de manifestations extérieures, y compris physiques, peuvent accompagner la relation psychologique.

2.2 Manquer et l'actualisation du sens psychologique

Manquer, comme nous l'avons vu, lexicalise un sens général d'absence. Sa valeur psychologique n'est que l'une de ses acceptions possibles et elle est actualisée par un seul modèle de réalisation argumentale. *Manquer* est interprété comme verbe psychologique uniquement lorsqu'il a deux arguments dans sa valence. Dans cette configuration argumentale particulière, le concept général d'absence peut être interprété de manière plus spécifique comme celui de séparation.

2.2.1 *Manquerpsych* en tant que prédicatif à deux arguments

Les exemples (6), (7) et (8) décrivent un événement psychologique en termes de relation qui s'établit entre ses deux participants. Les deux participants sont représentés par les deux arguments du prédicatif :

- (6) Les enfants lui manquent, il les voudrait une semaine sur deux.
- (7) Les enfants me manquent et je leur manque.
- (8) Je suis expatrié français à Bangkok et la France me manque parfois.

Comme nous l'avons vu plus haut, *manquer_{psych}* et *sybnit'* relèvent de deux classes différentes de verbes psychologiques, justement parce qu'ils diffèrent par le type d'argument en position d'objet syntaxique. *Manquer_{psych}* est un verbe de type ObjSubj : l'argument réalisé par le sujet syntaxique est l'objet de l'émotion ou thème.

Or, Grimshaw (1990) remarque que les constructions qui placent l'objet en position de sujet sont, de ce fait, doublées d'une relation causale. Effectivement, les *sybnit'* de (6) et (7), la *France* en (8) fonctionnent comme causes de l'état psychologique vécu par l'expérienteur. Une première conséquence de cette répartition thématique et aspectuelle consiste à considérer *manquer_{psych}* comme un verbe *exogène* : contrairement à *sybnit'*, l'origine de la « psy-chose » (Bouchard, 1995) a une existence indépendante de l'expérienteur. La source de l'émotion et son « support » (Anscombe, 1995) sont séparés.

Que devient l'expérienteur dans une telle configuration ? Premièrement, il se retrouve en position argumentale nettement moins forte, puisque, s'il

reste l'argument le plus saillant dans la structure thématique (Grinstead, 1998), il ne l'est pas dans la structure causale. Deuxièmement, il n'est pas intéressant de constater que, sur le plan syntaxique, il est réalisé par un complément prépositionnel datif. Dans ses études sur la sémantique des cas, Rudzka-Ostyn (2000) parvient à la conclusion selon laquelle chaque cas grammatical « profile » d'une manière très précise l'objet désigné par le syntagme nominal fléchi par rapport aux autres participants de l'événement. Dans cette optique, les constructions datives ont un sens prototypique qui se résume au concept de transfert d'un objet vers un bénéficiaire, le transfert étant initié par l'agent. C'est le cas d'un énoncé comme (14) :

(14) Il a donné à son élève le goût de la solidité et de la performance (Gentile).

Le SN datif désigne le bénéficiaire passif du transfert. On peut raisonnablement admettre que, même si *manquer_{pass}* n'est pas une structure ditransitive, son complément datif est associé au même rôle sémantique. Certes, (14) ne peut recevoir qu'une lecture non-stative avec un changement d'état à la clé, ce qui n'est pas le cas de *manquer_{pass}* interprété seulement d'une manière stative comme état mental ou psychologique. De ce fait, plutôt que de décrire l'expérimenteur comme bénéficiaire, il faudrait qualifier sa fonction sémantique de TISM, c'est-à-dire de « Target Subject Matter ».

Les facteurs dont nous venons de parler, c'est-à-dire l'organisation argumentale, la répartition des rôles thématiques et la valeur aspectuelle du prédicat, imposent une construction de la scène événementielle par *manquer_{pass}* radicalement différente de celle fournie par *avoir* et *changer* : l'orientation et le caractère de la relation psychologique. Nous la représentons ci-dessous par la figure 3 :



Figure 3

Une dernière précision s'impose avant de quitter cette section. Nous savons que les prédicats à caractère causal sont généralement ceux qui décrivent un événement complexe. Une relation causale s'établit entre deux sous-événements désignés par un prédicat complexe, comme dans le cas de (15) :

(15) Il a cassé la vitre au lieu d'ouvrir la porte.

représenté par la formule (15') :

[[X AGR]] CAUSE [SUBVANT «Y ÉTAT RESULTATIF»]

Or, nous venons de dire que *manquer_{sub}* décrit une mise en relation active. Il s'agit donc d'un événement simple, ce qui est effectivement le cas. Ce fait s'explique, à notre avis, par la nature de cette relation causale. Dans le cas de *manquer_{sub}*, la relation causale relie non pas deux sous-événements, comme dans (15) et (15'), mais une cause, représentée par l'argument thème et un état résultant localisé dans TISM.

2.2.2 Un « complément » pragmatique

Nous avons essayé de montrer, jusqu'ici, que *manquer* ne lexicalise pas un état psychologique, mais qu'il devient un *manquer_{sub}* dans un modèle de réalisation argumentale particulière. Cependant, il s'accorde avec l'exemple (9) que cette condition est nécessaire, mais non suffisante. Rappelons cet exemple :

(9) Serge Besson regrette cependant l'absence du demi d'ouverture du XV de la Ross, Jonny Wilkinson, blessé à la cuisse droite. « Oui, il ne me manque », a déclaré le Biarrot. Je me serais fait une joie de pouvoir le rejouer pour la première fois depuis cette demi-finale perdue en Australie ».

(9) remplit sous les conditions qui font de *manquer* un verbe psychologique. C'est un prédicat à deux arguments auxquels nous pourrions attribuer les fonctions sémantiques de thème et de l'expérienteur. Or, on donnerait difficilement à (9) une interprétation psychologique comme celle qu'on n'hésite pas à attribuer à (6), (7) et (8). De ce point de vue, (9) est beaucoup plus proche de (10) ou (11) où l'absence n'entraîne pas un état psychologique. D'ailleurs, dans (9), nous pouvons aisément substituer à *manquer* une expression dépourvue de tout sens psychologique sans que l'énoncé en soit sémantiquement modifié :

(9') Serge Besson regrette cependant l'absence du demi d'ouverture du XV de la Ross, Jonny Wilkinson, blessé à la cuisse droite. « Oui, il ne me fait défaut », a déclaré le Biarrot. Je me serais fait une joie de pouvoir le rejouer pour la première fois depuis cette demi-finale perdue en Australie ».

Or, une même substitution effectuée dans (6), (7) et (8) entraîne une modification sémantique sensible :

- (6) Les enfants lui font défaut, il les voudrait une semaine sur deux.
- (7) Les enfants me font défaut et je leur fais défaut.
- (8) Je suis expatrié français à Bangkok et la France me fait défaut parfois.

Pourquoi une telle différence de lecture ? La seule explication valide, nous semble-t-il, consiste à recourir aux facteurs pragmatiques liés à notre savoir sur le monde extérieur⁸. La relation psychologique se construit à la base de propriétés conventionnellement accordées à une catégorie d'objets. Nous savons que les relations entre un parent et son(s) enfant(s), entre un(e) expatrié(e) et son pays d'origine ne sont pas celles qui existent (ou que nous imaginons exister) entre deux rhythmes.

3. Structure communicationnelle

Les propriétés grammaticales de l'énoncé ne sont pas sans influence sur l'organisation du discours. Si *spănu* et *manșur*_{pres} peuvent décrire le même type d'émotions, les différences constatées sur le plan de réalisation argumentale, celui des rôles sémantiques et leurs répercussions syntaxiques nous obligent à nous poser la question des relations thématiques telles qu'elles sont organisées par les deux prédicats dans le discours.

On admet généralement que la position frontale dans un énoncé est celle qui correspond au thème, c'est-à-dire à l'objet du discours, dont on assure quelque chose. Or, l'une des différences majeures entre *spănu* et *manșur*_{pres} est justement la nature de l'élément en position frontale de l'énoncé. En polonais, l'argument-thème, celui qui constitue l'objet du discours en l'expérimencant, comme nous l'avons vu dans nos exemples, à commencer par (1) :

- (1) *Ala jak každy reglars tekni jednak za domem, za rodziną (potrac).*

⁸ Comme nous l'a fait remarquer l'un des lecteurs anonymes, le besoin de chercher des explications en dehors de faits linguistiques peut équivocal à un aveu de faiblesse. Cela ouvre un débat sur les limites de la théorie. Une approche théorique est-elle réellement pertinente si elle n'est pas capable de faire face, à elle toute seule, à tous les phénomènes qu'elle affronte ? Si nous décidons de nous en tenir pour comprendre et décrire les faits langagiers, quelle marge de tolérance faut-il avoir ?

Du point de vue de la structure de l'information, c'est l'expérientiel qui constitue « l'information donnée », celle qui est présumée connue de l'interlocuteur. L'expérientiel est donc l'élément le plus saillant dans la situation du discours. L'information nouvelle, non rhématisée à son sujet, est apportée par le prédicat avec ou sans son argument interne directionnel.

Il suffit de comparer (1) avec sa traduction française recourant à *manque_{non}* pour constater une situation inverse :

(1) *Mais comme à tout navigateur, sa maison, sa famille lui manquent.*

L'argument-sujet est un thème ou l'objet de l'émotion sur le plan actantiel. Sa fonction communicationnelle est celle de thème, objet du discours. C'est à propos de lui que l'on assure quelque chose. Dans la structure de l'information, il est considéré comme l'élément donné, connu de l'interlocuteur, celui dont la saillance situationnelle est la plus forte. L'information nouvelle à son sujet est apportée par le prédicat avec son argument interne qui n'est autre que l'expérientiel-DSM.

Une remarque triviale pourrait dès lors résoudre le problème. Lorsque j'utilise le verbe polonais *tydnąć*, je dis quelque chose de l'expérientiel-support de l'émotion, alors qu'en employant *manque_{non}* en français, j'assure quelque chose de l'objet qui cause une émotion.

4. Conclusion

« Notre personnalité est formée par notre culture et par notre langue ». Nous avons sciemment laissé de côté les facteurs culturels, décrits abondamment par des auteurs tels que Wierzbicka, pour nous apercevoir que les langues elles-mêmes mettent en place des moyens différents pour décrire une même situation. Un locuteur natif polonais disposant dans sa langue maternelle d'un lexème spécifique pour l'émotion qui nous intéresse dans ces pages est exposé à une tâche d'analyse complexe des énoncés décrivant un état psychologique similaire à l'aide de *manque_{non}*. En français, cette description est effectuée et complétée sur plusieurs niveaux de prédication. Tout d'abord, elle est tributaire d'un certain mode de réalisation argumentale du prédicat avec une répartition inversée des rôles rhématiques. Un complément pragmatique est nécessaire pour actualiser le sens d'absence, lexicalisé par *manque* en tant qu'émotion. Qui plus est, l'organisation différente des prédicats *tydnąć* et *manque_{non}* met en avant un autre participant de l'événement : en polonais, le discours

apporte une information au sujet de l'expérience, c'est-à-dire celui qui assure l'émission. Le français, au contraire, dans l'objet de l'émission constitue l'élément saillant du point de vue communicationnel. N'est-ce pas l'une des explications possibles de ce « sentiment de séparation » entre la langue et le monde environnant ?

Bibliographie

- Anacibar, J.-C. 1995. Morphologie et représentation énonciative : le cas des noms de sentiment et d'attitude, dans : *Langue Française* 105, 40-54.
- Berthommieu A. & Cadot P. (éd.). 1993. *Les prépositions – méthodes d'analyse*. Presses universitaires de Lille.
- Bialy, A. 2005. Polish Psychological Verbs at the Lexicon-Syntax Interface in Cross-Linguistic Perspective. Frankfurt a/Main : Peter Lang.
- Bouchard, D. 1995. Les verbes psychologiques, dans : *Langue Française* 105, 6-16.
- Grinstead, J. 1990. *Argument Structure*. MIT Press.
- Gross, M. 1995. La notion de lieu argument du verbe, dans : Bas-Zorek Szydkiewicz H., Kupferman L. (éd.). *Tematowe studia w lingwistyce francuskiej i generalnej*. *Linguisticae Investigationes Supplementa* 20, 173-200. Amsterdam/Philadelphia : John Benjamins Publishing Company.
- Grzegorzczakowa, R. 1999. Nazwy 'emocji' w różnych językach, dans : *Semantyka i morfologia języków 2*, Warszawa, Polska Akademia Nauk, Instytut Słowistyczny.
- Klein, W. 1986. *Second Language Acquisition*, Cambridge University Press.
- Levin, B. & Rappaport Hovav, M. 2005. *Argument Realization*. Cambridge University Press.
- Moeschler, J., Riboul, A. 1994. *Dictionnaire encyclopédique de pragmatique*. Paris : Éditions du Seuil.
- Nowakowska-Konopa, I. 1995. Koncepcjonalizacja świata w języku polskim, Poznań.
- Rudka-Owya, B. 2000. *Z rozważań nad kategorią przypadku*. Kraków : Universitas.
- Roux, N. 1995. Les verbes de sentiment peuvent-ils être agencés ?, dans : *Langue française*, 105, 28-39.
- Van de Velde, D. 1997. Ces objets objets du désir. L'objet des verbes de sentiment, dans : *Travaux de linguistique* 35, 67-78.
- Vandeloist, C. 1986. *L'espace en français : sémantique des prépositions spatiales*. Paris : Éditions du Seuil.
- Wierzbicka, A. 1992. *Semantics, Culture and Cognition. Universal Human Concepts in Culture-Specific Configurations*. Oxford University Press.
- Wierzbicka, A. 2003. *Cross-Cultural Pragmatics. The Semantics of Human Interaction*. Berlin, New York : Mouton de Gruyter.

Dictionnaires

Doroszewski, W. (ed.) 1967. *Słownik języka Polskiego*. Warszawa : Polska Akademia Nauk, Państwowe Wydawnictwo Naukowe.

Szymczak, M. (ed.) 1983. *Słownik języka Polskiego*. Warszawa : Państwowe Wydawnictwo Naukowe.

Traflet de la Langue Française informatif (TLFI), <http://atilf.atilf.fr/TLFI.htm>.

Arriver et venir – quand la deixis fait (et ne fait pas) la différence

Dorota Sikora

ATILF – CNRS (UMR 7118) – Nancy – Université

1. Introduction

Venir et arriver, arriver et venir : est-ce bonnet blanc et blanc bonnet ? Non – réponses certainement plus d'un étudiant de français langue étrangère, repris plusieurs fois pour avoir employé l'un pour l'autre, là où il ne fallait surtout pas le faire. Oui – répliquera un autre, se souvenant d'un exercice dans lequel il devait, avec le souci légitime d'élargir son vocabulaire, « remplacer le mot entre parenthèses par son synonyme ». Dans cet article, nous nous proposons d'observer les relations synonymiques qui s'établissent entre ces deux verbes dans leurs emplois spatiaux, c'est-à-dire ceux qui réfèrent à des situations de mouvement. Il convient donc de préciser ce que nous entendons par *mouvement* ⁽¹⁾.

La notion de mouvement (déplacement)

En suivant Asher et Sablayrolles (1995 : 166), nous admettons que « [L]e mouvement lui-même est à chaque fois un type d'événement et il consiste dans un changement de position spatiale à travers le temps » ⁽²⁾. Talmy (2000/01 : 25) définit les quatre éléments conceptuels qui forment une situation ou un événement de mouvement. *Figure* est un objet mobile (ou bien conceptualisé comme tel, et c'est ainsi que nous employons ce terme). *Site* (*Ground*) est un point de référence par rapport auquel le mouvement de figure peut être décrit et il recouvre, selon le cas, la source (lieu initial) ou bien le terminus (lieu final) du déplacement. L'ensemble des positions adoptées par figure dans le cours de son déplacement forme sa *trajectoire*. Dans cet article, nous allons donc employer le terme de *mouvement* comme synonyme de *déplacement*.

(1) Certains auteurs intègrent dans cette classe de sens (Ashby et Kupferman 2008) des verbes comme *dévisser*, *eslever*, etc. (cf. Boons, 1985, 1987).

(2) Asher et Sablayrolles (1995 : 166) : « Motion is a property of objects (i.e. objects move), but motions themselves are a type of event. Motions are changes of some sort of spatial position over time ».

2. Qu'en est-il de la synonymie entre *arriver* et *venir* ?

2.1. Quelques remarques sur la synonymie « en général »

Depuis Duchacek (1979), un accord semble établi sur la notion de synonymie : sont synonymes « les unités lexicales de sens identique ou presque identique ou du moins aussi proches de sens qu'elles sont interchangeables dans certains contextes » (Duchacek, 1979 : 9). C'est aussi à Duchacek que nous devons l'idée selon laquelle il existe des degrés de synonymie allant de l'identité totale (absolue) à une quasi-identité de sens ou synonymie partielle plus ou moins forte.

La distinction généralement opérée entre la synonymie totale (ou absolue) et la quasi-synonymie est le reflet du caractère dynamique des relations synonymiques. La synonymie totale (absolue) suppose une identité « parfaite » de sens entre les expressions A et B. Elles sont alors substituables dans tous les contextes grammaticalement bien formés : là où A est approprié, B l'est aussi et vice versa. De tels cas de synonymie sont extrêmement rares, sinon inexistantes. Le concept de synonymie absolue sert plutôt de point de référence cognitive permettant de considérer la synonymie comme un phénomène scalaire, plus ou moins éloigné de ce cas idéal.

D'un autre côté, la synonymie partielle recouvre une quantité de relations intuitivement perçues comme celles de ressemblance, mais différant en réalité d'un point de vue tant sémantique (degrés de prototypicalité variés, différences aspectuelles entre les lexèmes concernés) que stylistique (Duchacek 1979, Cruse 2004).

Si l'étude de Duchacek (1979) constitue, depuis trente ans, un point de référence, voire un point de départ, théorique et terminologique pour de nombreuses réflexions sur la synonymie, la problématique n'en reste pas moins riche et stimulante. Il ne saurait en être autrement dans la mesure où les relations plus ou moins synonymiques entre deux unités lexicales ont un caractère dynamique : l'identité rescatée par les locuteurs n'est pas une constante. L'expérience d'un étudiant de FLE, quelque peu perdu face aux relations synonymiques instables entre les occurrences de *venir* et celles d'*arriver*, le montre bien. Elles s'établissent dans certains emplois, elles disparaissent lorsque le co-texte, voire le contexte, change. Cette instabilité rend leur acquisition particulièrement difficile.

Mais c'est grâce à ce caractère dynamique et fluctuant que la synonymie ne manque pas de sel. « Si nous voyons dans la synonymie simplement une identité de sens, elle apparaîtra comme une relation somme toute peu intéressante. Si nous considérons au contraire que les synonymes sont des mots dont les ressemblances sémantiques sont plus saillantes que les différences susceptibles de les opposer, une vaste sphère d'investigation s'ouvre. Quel est le type de différences qui ne met pas en cause l'intuition d'identité ? Comment se fait-il que les synonymes soient si fréquents ? (Alors qu'une identité absolue ne semble pas motivée d'un point de vue fonctionnel) Existe-t-il des domaines du lexique où ils sont particulièrement nombreux ? »⁽³⁾ (Cruse 2004 : 154).

(3) « If we interpret synonymy simply as sameness of meaning, then it would appear to be a rather uninteresting relation ; if, however, we say that synonyms are words whose semantic similarities are more salient than their differences, then a potential area of interest opens up. What sort of difference do not destroy an intuition of sameness ? Why are such synonyms so frequent ? (Absolute sameness of meaning would seem to be functionally unmotivated). Do they proliferate in particular areas of the vocabulary ? » (Cruse 2004 : 154).

2.2. Synonymie dans un cas spécifique : *arriver* et *venir* en tant que quasi-synonymes en langue

Arriver et *venir*, sont-ils réellement des synonymes ? Deux réponses semblent s'esquisser selon que l'on s'en tienne aux définitions lexicographiques ou que l'on s'en remette aux faits linguistiques, tels que le test de substituabilité. En effet, les emplois synonymiques de *arriver* et de *venir* sont largement confirmés : parmi les soixante expressions répertoriées par DicoSyn⁽⁴⁾ comme plus ou moins équivalentes de *venir*, *arriver* vient en troisième position. Inversement, *venir* n'est que quatrième sur la liste des cinquante quatre synonymes possibles d'*arriver*⁽⁵⁾.

Si l'on s'en tient à des définitions lexicographiques qui, par leur nature, sanctionnent ce qu'offre la langue, l'éventuelle synonymie de *venir* et de *arriver* devrait paraître plutôt incertaine.

2.2.1. Arriver

Le *Treasure de la Langue Française informatisé* (désormais TLFi) précise que les emplois spatiaux de ce verbe ne montrent aujourd'hui aucune préférence étymologique pour les déplacements sur l'eau. *Arriver* signifie « parvenir à l'étape fixée, parvenir à atteindre (qqn) après des difficultés, parvenir à destination »⁽⁶⁾. Par ailleurs, la destination (terminus dans la terminologie que nous avons adoptée) peut être explicitement spécifiée par un complément dit de lieu. *Arriver* apparaît également dans des emplois absolus, c'est-à-dire sans que le terminus soit clairement identifié par un complément.

2.2.2. Venir

La définition du/des sens spatiaux de *venir* présente une structure différente, selon que le verbe est suivi ou non d'un complément de lieu. Construit avec un complément de lieu, le verbe signifie « se déplacer dans la direction d'un lieu ou d'une personne de référence »⁽⁷⁾. Deux périphrases sont proposées pour les constructions dans lesquelles *venir* est accompagné d'un « complément de lieu indiquant le terme du déplacement » : a. « aller auprès d'une personne, dans un lieu, pris comme référence », comme dans *venir à l'école* et b. « aller, se diriger vers », par exemple dans des constructions telles que *venir au devant de quelqu'un*. Si dans les deux cas le terminus est explicitement nommé, conformément à la définition b., il n'est pas forcément atteint par figure.

Outre les ressemblances que l'on repère dans les constructions grammaticales des deux verbes, les définitions du TLFi mettent en avant des points importants pour notre discussion des relations synonymiques entre *arriver* et *venir*. Si *venir* peut être traité comme un verbe de déplacement dans la mesure où il réfère à un événement de mouvement (figure change d'emplacement, en adoptant des localisations successives par rapport au terminus), *arriver* semble se focaliser non pas sur le déplacement, mais sur la mise en relation de figure et du terminus. En simplifiant, les deux verbes ne réfèrent donc pas à la même chose.

(4) <http://www.ceril.fr/synonymie/venir/verbe>, <http://www.ceril.fr/synonymie/arriver>

(5) Il nous est impossible, faute de place, de discuter cette asymétrie des rapports synonymiques entre *arriver* et *venir*. Il nous semble cependant qu'une des raisons possibles serait à chercher parmi les contraintes pesant sur la nature de l'argument sujet de *venir*. *Arriver* est est dépourvu et accepte bien des arguments non agentifs.

(6) <http://atilf.atilf.fr/dendien/scripts/tlfiv5/advanced.exe?8;1=2172557160>

(7) <http://atilf.atilf.fr/dendien/scripts/tlfiv5/vissuel.exe?12;1=117020265;1=1;page=201-3>

Cependant, les données fournies par DicoSyn indiquent un lien synonymique fort entre les deux. A notre avis, il s'explique par une quasi-synonymie reposant sur un rapport métonymique de type partie-tout entre les sens de *arriver* et de *venir*, le premier désignant la phase finale d'un déplacement.

2.3. « ...lieu ou personne pris comme référence » ou le caractère déictique de *venir*

La comparaison des deux définitions permet de remarquer une autre différence de taille. Outre le sens descriptif qui fait de *arriver* et de *venir* des candidats pour une relation de quasi-synonymie, *venir* présente une caractéristique en plus. Nous pensons notamment aux termes de *lieu* ou de *personne de référence*. Avec ou sans complément de lieu, le terminus est identifié à travers un *lieu* ou une *personne de référence*. On constate donc qu'en dehors du sens descriptif de mouvement paraphrasable par *aller*, *venir* a aussi un sens instructionnel ou procédural permettant d'effectuer une opération d'identification de l'entité qui servira de référence. Une telle procédure ne semble pas nécessaire dans le cas de *arriver*.

On comprend dès lors mieux pourquoi le maniement et les rapports entre ces deux verbes présentent tant de difficultés en cours d'apprentissage de français langue étrangère. Il s'agit de comprendre la nature d'une opération dont le déroulement n'est pas simple à établir. Pour le faire, le locuteur doit suivre une certaine procédure. La question serait donc double : 1) comment trouve-t-on « le lieu ou la personne de référence » ? 2) comment se fait-il que dans certains co(n)-textes, le sens procédural n'empêche pas les relations de synonymie entre *venir* et *arriver* ? Ce qui est une autre façon de s'interroger, avec Cruse (2004), sur les ressemblances plus fortes que les différences.

Afin de saisir la démarche qui conduit le locuteur vers la personne ou le lieu de référence, nous proposons deux exemples d'emplois du verbe *venir* :

- (1) Viens !
- (2) Viens au concert !

Dans (1), facilement accompagné d'un geste, l'énonciateur constitue le terminus du déplacement. Plus précisément, le lieu est identifié à travers sa position dans l'espace. Dans (2), le complément de lieu *au concert* identifie le lieu-terminus indépendamment de l'énonciateur. Cela ne signifie pas pour autant que l'énonciateur ne joue aucun rôle dans la signification de l'énoncé : en employant *venir* dans (2), il signale que lui-même compte se rendre au concert. En d'autres termes, il se situe au terminus du déplacement.

Ainsi, les exemples (1) et (2) montrent-ils les deux modes d'implication déictique de l'énonciateur :

1. le lieu de référence est identifié comme celui où se trouve l'énonciateur (« lieu englobant » chez Kleiber, 2008)⁽⁸⁾.
2. le lieu de référence est identifié indépendamment de l'énonciateur, mais celui-ci n'est pas indifférent pour la signification de l'énoncé, puisqu'en recourant à *venir*, il indique sa présence au terminus du déplacement.

Ces observations confirment le caractère déictique largement reconnu de *venir*.

(8) Ce n'est sans doute pas un hasard si, dans ce type d'emplois, *venir* est suivi de *ici* à valeur englobante (cf. Kleiber, 2008 : 133).

Elles s'inscrivent, nous semble-t-il, dans les remarques de Kleiber (2008) qui propose une conception *token*-réflexive de la deixis :

Une expression *token*-réflexive est une expression qui renvoie à son référent par l'intermédiaire d'éléments spatio-temporellement reliés à son occurrence.

Kleiber, 2008 : 124

Pour la suite de notre réflexion, nous adopterons la thèse de travail selon laquelle l'énonciateur en tant qu'« élément spatio-temporellement relié » à certaines occurrences aussi bien de *venir* qu'à celles d'*arriver* joue un rôle important, voire décisif, pour les relations synonymiques (ou pour leur absence) entre ces deux verbes. En observant les possibilités de substitution dans des co(n)-textes précis, nous nous reposerons sur le classement proposé par Fuchs (1997).

2.4. Les types de co-textes (Fuchs 1997)

L'existence ou non des relations synonymiques entre deux lexèmes, ou plus généralement, entre deux expressions, est vérifiée à l'aide du test de substitution. Il consiste à remplacer, dans un co-texte, un lexème (une expression) par celui (celle) que l'on considère comme équivalent(e), pour vérifier si cette opération produit une modification dans sa signification. La possibilité d'une commutation témoigne de la synonymie, son impossibilité l'exclut. En observant les effets de telles substitutions dans les environnements linguistiques, Fuchs (1997) distingue deux principaux types de co-textes. Il existe, en effet, des co-textes dissimulateurs et des co-textes assimilateurs. Chacun des deux types possède une variante forte et une variante faible. Au total, quatre cas de figure doivent être envisagés :

1. un co-texte dissimulateur fort interdit toute substitution ;
2. alors qu'un co-texte dissimulateur faible l'autorise, mais les deux énoncés ainsi obtenus ne sont pas des paraphrases. En d'autres termes, leurs sens sont différents et l'on ne peut pas considérer les lexèmes (ou les expressions) substitués(e)s comme synonymes.
3. Les co-textes assimilateurs faibles autorisent une substitution, mais celle-ci aboutit à une quasi-synonymie, c'est-à-dire à un recouvrement partiel de la signification des énoncés.
4. Les co-textes assimilateurs forts sont ceux dans lesquels une substitution produit des énoncés qui sont des paraphrases l'un de l'autre.

3. *Venir* et *arriver* en co-texte⁽⁹⁾

3.1. ...dissimulateur fort

Nous commençons nos analyses par des co-textes qui excluent la substitution synonymique, car repérer la nature de l'élément responsable du blocage nous semble tout aussi instructif que cerner une similitude. Observons l'exemple (3) :

- (3) « - Nous avons décidé d'aller voir la tombe de papa. Viens si tu veux, il t'aimait bien ». Cette dernière phrase est ridicule mais je vais avec elles. (Frantex1)

(9) Je tiens à remercier chaleureusement tous les membres de l'équipe Lesique (ATILF-CNRS UMR 7118) qui ont bien accepté de participer à l'atelier de travail du 14 mars 2008. Leurs avis sur les exemples que j'ai pu recueillir et leurs remarques m'ont été précieuses. Je suis seule responsable d'éventuelles erreurs et incompréhensions des idées exposées lors de nos discussions.

Remplacer *venir* par *arriver* est, en effet, impossible, comme le montre (3') :

(3') « - Nous avons décidé d'aller voir la tombe de papa. *Arrive si tu veux, il t'aimait bien ». Cette dernière phrase est ridicule mais je vais avec elles.

L'ensemble de (3) forme un co-texte dissimulateur fort, mais – pourrait-on objecter – cela n'est pas lié au caractère déictique de *venir*. On pourrait simplement faire remarquer que *arriver*, de manière générale, résiste à la forme impérative, tout comme certains autres lexèmes verbaux susceptibles de fonctionner en synonymie avec lui, par exemple *parvenir*. Prenons donc un autre exemple :

(4) Je te donne mon disque mais tu viens à mon concert. (forum Internet)

La substitution de *venir* par *arriver* produit l'énoncé (4') qualifié de douteux par nos informateurs :

(4') ? Je te donne mon disque mais tu arrives à mon concert.

Dans (4), le co-texte dissimulateur fort ne peut pas être expliqué par l'emploi de l'impératif, puisque *venir* est employé à l'indicatif.

Indépendamment de la forme impérative, dans l'exemple (3) tout comme dans (4), l'engagement de l'énonciateur empêche la substitution. Dans (3), l'énoncé contenant l'occurrence de *venir* est enclavé dans un discours direct pris en charge par une énonciatrice (l'une d'« elles »), alors que la séquence du discours dans son ensemble doit être attribuée à un autre locuteur, celui qui assume le *je*. L'énonciatrice de *venir* signale sa présence au terminus, c'est-à-dire à la tombe de papa. Nous avons donc affaire au second mode d'implication déictique.

L'exemple (4) présente un cas similaire, la forme impérative en moins. En faisant recours à *venir*, l'énonciateur qui, cette fois est également le locuteur, signifie qu'il sera au concert, lieu de référence, terminus du déplacement. On pourrait, certes, objecter que cela va de soi dans la mesure où le possessif *mon* ne laisse pas de doute à ce sujet. Il est difficile, en effet, d'imaginer qu'il soit absent de son propre concert. L'engagement de l'énonciateur reste cependant entier, si nous modifions le co-texte pour donner à croire qu'il s'agit d'un disque et d'un concert assuré par une tierce personne :

(4'') Je te donne son disque mais tu viens à son concert.

Les exemples (4) et (4''), ainsi que nos tentatives infructueuses de substitution par *arriver*, montrent que le recours au seul co-texte n'est pas suffisant, puisque l'identification du référent dépasse le cadre textuel. Il faudrait donc parler, dans ce cas, de contexte dissimulateur fort.

3.2. ...dissimulateur faible

L'exemple (5) et sa version (5') obtenue par substitution de *venir* par *arriver*, a fait l'objet d'un long débat⁽¹⁰⁾. Il est vrai que (5') est discutable à plus d'un égard :

(5) Message de A :

Bonjour

Je vais bientôt commencer mes travaux d'agrandissement et pour la dalle je compte faire sur un hérisson. Cependant où trouve-t-on ce genre de gros cailloux ? Merci.

(10) Atelier de travail du groupe « Lexique », ATELIF-CNRS (UMR 7118), le 14 mars 2008.

Message de B :

Va voir s'il n'y a pas de constructions dans ton coin, il y a souvent de beaux gros blocs et gros gravats ! Les proprios sont contents de s'en séparer gratuitement.

Message de C :

Si non tu viens chez moi, j'ai 10m3 de gravats (béton, carrelage, tuile) à prendre gratuitement.

(5') Message de A :

Bonjour

Je vais bientôt commencer mes travaux d'agrandissement et pour la dalle je compte faire sur un hérisson. Cependant où trouve-t-on ce genre de gros cailloux ? Merci.

Message de B :

Va voir s'il n'y a pas de constructions dans ton coin, il y a souvent de beaux gros blocs et gros gravats ! Les proprios sont contents de s'en séparer gratuitement.

Message de C :

Si non tu arrives chez moi, j'ai 10m3 de gravats (béton, carrelage, tuile) à prendre gratuitement.

D'un avis général, le message C de la version (5') fut considéré comme « moins bon » que la version originale. Il avait d'abord été rejeté, puisque contrairement à (5), il ne localisait pas les gravats, c'est-à-dire dissociait le lieu où ils se trouvaient de *chez moi*. Finalement, il a été jugé acceptable, mais uniquement à l'oral et sous réserve de disposer d'informations supplémentaires. De tels écarts de jugements ne peuvent qu'attirer davantage l'attention.

Les réserves émises à l'égard de la version (5') étaient, à chaque fois, liées aux problèmes de localisations, celle de gravats ou celle des protagonistes. Comparativement, *venir* dans (5) apparaît comme un élément fédérateur dans la construction de la signification du message C. Cet élément semble au contraire faire défaut dans la version (5'C). Ce qui manque à (5'), à notre avis, c'est l'indication de la procédure à suivre pour identifier la personne de référence, comme cela se fait dans le cas de *venir*. *Venir* dans le message (5C) est accompagné d'un complément de lieu *chez moi* désignant le terminus du déplacement que doit faire figure. Il en est de même de sa version modifiée dans (5'). Originellement cependant, le message de C dans (5) véhicule une indication supplémentaire : *venir* autorise une inférence au sujet de l'énonciateur. On comprend qu'au moment, quel qu'il soit, où figure (*tu*) atteindra le terminus *chez moi*, l'énonciateur s'y trouvera également. *Arriver* de (5') reste neutre à cet égard. Le recouvrement des significations n'est donc que partiel, car il ne concerne que le déplacement de figure. Aussi bien dans (5) que dans (5'), il lui faudra parcourir la même trajectoire pour atteindre le terminus spécifié par *chez moi*. C'est par l'implication de l'énonciateur que les énoncés incriminés diffèrent. Contrairement à (5), (5') désengage l'énonciateur quant à son association au terminus du déplacement.

3.3. ...assimilateurs forts

Deux types de situations forment des co-textes dans lesquels la substitution entre *venir* et *arriver* débouche sur des énoncés pleinement acceptés et considérés comme synonymes. La substitution s'y opère donc *salvo significatione*. L'analyse de notre corpus de travail a permis de distinguer deux types de configurations qui forment des co(n)-textes assimilateurs forts. Plus précisément, les co(n)-textes assimilateurs forts sont en grande partie tributaires du rôle assumé par *je*.

3.3.1. Énonciateur en tant que terminus

Dans l'exemple (6), le long co-texte gauche met en place tous les éléments nécessaires pour l'interprétation de l'énoncé qui nous intéresse. Il constitue un scénario qui dispose tous les participants de la scène, et surtout il permet d'identifier, par une procédure anaphorique, le référent du pronominal *il*, figure du déplacement, qui s'avère être le serpent, et non pas Yann, « ami des bêtes ». Comme le montre (6'), la substitution par *arriver* ne modifie pas la signification initiale de (6) :

- (6) *J'étais en train d'enjamber des arbres tombés, les autres étaient déjà passés et ne l'avaient pas remarqué.
Me retrouvant en équilibre sur un tronc, je le vois, je reste immobile, en criant : « Merde, un serpent ! »
Il était en train de se dénouer pour monter sur le même tronc que moi ! Le stress !!!
Dilemme : je m'éloigne vite fait, ou je fais comme on a appris et je fais l'arbre.
En plus, Yann, cet ami des bêtes, me disait « Vas-y, bouge pas, pour la photo et tout ! »
Tu parles... Comme il venait vers moi, et que je n'étais pas sûr de reconnaître l'espèce, une fois qu'il était à 40 cm de ma cheville, je me suis barré ! (blog Internet)*
- (6') *J'étais en train d'enjamber des arbres tombés, les autres étaient déjà passés et ne l'avaient pas remarqué.
Me retrouvant en équilibre sur un tronc, je le vois, je reste immobile, en criant : « Merde, un serpent ! »
Il était en train de se dénouer pour monter sur le même tronc que moi ! Le stress !!!
Dilemme : je m'éloigne vite fait, ou je fais comme on a appris et je fais l'arbre.
En plus, Yann, cet ami des bêtes, me disait « Vas-y, bouge pas, pour la photo et tout ! »
Tu parles... Comme il arrivait vers moi, et que je n'étais pas sûr de reconnaître l'espèce, une fois qu'il était à 40 cm de ma cheville, je me suis barré ! (blog Internet)*

Dans (6) comme dans (6'), le locuteur, c'est-à-dire celui qui dit *je*, et l'énonciateur prenant en charge l'énoncé ne font qu'un. Il se constitue explicitement « personne de référence ». C'est par son intermédiaire que nous identifions le terminus du déplacement par rapport auquel le mouvement de figure (*il* – *serpent*) est évalué. Plus simplement, le lieu « englobant »⁽¹¹⁾ l'énonciateur est en même temps le site du déplacement.

Le mécanisme de construction de la signification est cependant complexe, puisque l'énonciation en tant qu'acte individuel se fait dans un contexte qui efface son cadre spatio-temporel. Il s'agit, en effet, d'un écrit Internet qui remplit sa fonction communicative indépendamment de la situation spatio-temporelle de l'énonciateur au moment de son acte d'énonciation. Les caractéristiques spatio-temporelles de l'acte d'énonciation sont par ailleurs peu pertinentes pour le co-énonciateur, c'est-à-dire celui qui est en charge d'interpréter (6). En clair, il sait seulement que l'énonciateur était devant son ordinateur mais cette information est de peu d'utilité pour trouver le référent, en l'occurrence celui de *il venait vers moi*.

Jusque là, nous avons traité *venir* comme une expression *soi-même*-réflexive, en adoptant la définition formulée par Kleiber (2008 : 124). Elle stipule que le référent d'une telle expression doit être trouvé, *appelons-le*, « par l'intermédiaire d'éléments spatio-temporellement reliés à son occurrence ». Or, le seul lien spatio-temporel entre l'occurrence, qui relève de l'acte d'énonciation, et le référent, c'est-à-dire l'événement de mouvement, c'est l'énonciateur. En tant qu'individu, il est, certes, seul et unique, mais il peut aussi être appréhendé comme un ensemble d'instances spatio-tem-

(11) Cf. Kleiber, 2008.

porelles. Dans l'une d'entre elles, il a participé de l'événement décrit dans (6). C'est donc ainsi, en tant qu'ensemble instances spatio-temporelles, que l'énonciateur assure le lien entre l'événement et l'acte d'énonciation. Le contexte assimilateur fort de (6) est issu d'une configuration très particulière d'éléments sémantiques et discursifs : l'énonciateur, en tant qu'instance spatio-temporelle, est en même temps impliqué dans l'événement, puisque c'est lui qui permet d'identifier le point de référence spatiale par rapport auquel le mouvement s'effectue⁽¹²⁾.

3.3.2. Énonciateur en tant que figure

On remarque (Bloneste, 2005) que lorsque le verbe *venir* est employé à la première personne, il perd son caractère déictique. Une perte de ce caractère consisterait donc à affranchir le processus de référence du recours aux éléments spatio-temporellement reliés à l'occurrence de ce verbe. L'exemple (7) permet de vérifier si un tel affranchissement a réellement lieu.

- (7) Franchement, quand je viens à Lyon, je suis surpris par l'absence de bateaux-mouches entre Rhône et Saône pour visiter le centre de Lyon jusqu'à l'Île Barbe. (Forum du Grand Lyon)

Comme dans l'exemple (6), le *je* est assimilé, sur le plan discursif, à la fois par le locuteur et par l'énonciateur. Si nous abordons (7) du côté de son référent, c'est-à-dire par l'événement de mouvement désigné⁽¹³⁾, le *je* représente figure de déplacement. Le terminus est explicitement identifié par le complément *à Lyon*. Les particularités sémantico-discursives d'une telle configuration ressortent clairement, lorsque l'on compare (7) avec (4'') :

- (4'') Je te donne son disque mais tu viens à son concert.

Dans les exemples (4'') et (7), le terminus est explicitement spécifié, respectivement par *à son concert* et *à Lyon*. Comme nous l'avons vu dans la section 3.1., (4'') est un contexte dissimulateur fort, ce qui n'est pas le cas de (7). En (7), la substitution se fait *in vivo* signifiante :

- (7') Franchement, quand j'arrive à Lyon, je suis surpris par l'absence de bateaux-mouches entre Rhône et Saône pour visiter le centre de Lyon jusqu'à l'Île Barbe.

La différence essentielle entre (4'') et (7) réside dans le positionnement de l'énonciateur. En interprétant un énoncé tel que *tu viens à son concert*, c'est en vertu du choix de *venir* que nous associons l'énonciateur au terminus. Dans (7) l'énonciateur est en même temps figure de déplacement, il est donc supposé, au terme de celui-ci, rejoindre le terminus *Lyon*. Par conséquent, le mode d'association de l'énonciateur avec le terminus du déplacement tel que nous l'avons vu en (4) et en (4'') se trouve neutralisé par la sémantique du verbe *venir*.

Plutôt que de perdre du caractère déictique par *venir*, nous préférons parler de sa neutralisation. Si, comme nous l'avons admis, c'est l'énonciateur qui constitue l'élément spatio-temporel nécessaire pour identifier le référent, il n'a pas disparu avec l'emploi de la première personne. Il a, au contraire, renforcé sa présence en tant que participant de l'événement de mouvement, en autorisant une relation synonymique forte avec *arriver*.

(12) Et, plus loin, l'exemple (8).

(13) Dans la mesure où cette information n'a pas d'incidence pour notre propos, nous ne faisons que signaler qu'il s'agit d'un événement pluriel.

3.4. ...assimilateur faible

L'exemple (8) est intéressant, car le degré de synonymie entre *arriver* et *venir* en (8a) (8'a) n'est pas le même que celui qui relie *venir* et *arriver* dans la partie (8b) et (8'b). En ce qui concerne (8a) et (8'a), la synonymie peut, à la suite de Duchacek, être considérée comme parfaite. Sans être explicitement spécifié par un complément, le site, lieu « pris comme référence », est accessible à travers l'énonciateur. Il s'agit donc d'une information déictique de type I : le lieu de référence est identifié comme celui où se trouve l'énonciateur (qui l'englobe, dirait-on avec Kleiber). Pour ce qui est de (8a), il s'agit d'un co-texte assimilateur fort, comparable à celui de (6).

(8) Nous avons mangé notre jambon et nos compotes assis en tailleur devant l'exposition et il a dit : (8a) « Quand ils arrivent, les invités ? – (8b) Je ne crois pas qu'ils viendront ce soir », ai-je dit, le cœur arrêté. (Frantext)

(8') Nous avons mangé notre jambon et nos compotes assis en tailleur devant l'exposition et il a dit : (8'a) « Quand ils viennent, les invités ? – (8'b) Je ne crois pas qu'ils arriveront ce soir », ai-je dit, le cœur arrêté.

En (8b) et (8'b), le lien de synonymie entre *venir* et *arriver* semble, en revanche, moins fort. C'est parce que, comme remarque Fuchs (1997 : 37), « le type de co-texte ne peut se définir qu'à taille explicitée. En effet, si l'on change la taille du co-texte, on peut en modifier le type : selon les cas, on ajoute ou on supprime des opérations, en sorte que la nature de l'interaction entre A (ou B) [A, B représentent les expressions synonymiques, en l'occurrence (8b) et (8'b) – DS] et son environnement au sein de l'énoncé peut s'en trouver changée ».

Effectivement, pour *arriver* et *venir*, le co-texte n'est pas le même, selon qu'il s'agit de leurs occurrences en (8a) et (8'a) et celles qui apparaissent en (8b) et en (8'b). Même si (8a) et (8b) réfèrent à un seul et même événement de mouvement, (8b) se trouve dans la portée du modalisateur épistémique *je ne crois pas*. Ce modalisateur porte sur l'ensemble du contenu propositionnel *ils viendront ce soir/ils arriveront ce soir*. Le Querler (1996) qualifie une telle modalité d'extra-prédicative, traduisant une attitude extérieure au contenu propositionnel. Ce contenu propositionnel n'est pas identique. Il semblerait que le modalisateur actualise le sens lexical tel qu'il est défini en langue. Or, les définitions du TLFi présentent *arriver* et *venir* comme des quasi-synonymes : *venir* renvoie à un déplacement et peut être paraphrasé par *aller à*, alors qu'*arriver* sélectionne plutôt l'étape finale et la mise en contact avec le terminus. Nous y avons détecté une relation métonymique entre les sens lexicaux de *arriver* et *venir*, puisque l'étape finale fait bel et bien partie de la trajectoire, l'un des éléments conceptuels constitutifs d'un déplacement.

En (8b) et en (8'b), l'opérateur modal *je ne crois pas que* fait ressortir le sens propre à chacun des lexèmes. Par conséquent, avec *venir* l'énonciateur de (8b) met en doute le déplacement dans son intégralité. Il ne semble pas convaincu que les invités se soient tout simplement mis en route. Avec *arriver* dans (8'b), seul le fait d'atteindre le terminus est sujet à caution, d'autant plus que l'adverbiale temporelle *ce soir* localise cet événement dans le temps. Les invités (figure) ont bien entrepris le voyage, mais ils n'atteindront pas le terminus (*vous*) ce soir. Ainsi, pour ce qui est des relations synonymiques entre *venir* et *arriver* en (8b), l'exemple (8) représente-t-il un co-texte assimilateur faible où la commutation conduit à des énoncés « plus ou moins paraphrastiques, en ce sens que les valeurs prises par chacune des deux expressions A [*venir* dans l'énoncé (8b) de l'exemple (8) – DS] et B [*arriver*

dans l'énoncé (8'b) de l'exemple (8'b) – DS] ne se chevauchent que partiellement » (Fuchs 1997 : 33)

4. En guise de conclusion

Accusée de tous les maux, mise en cause jusqu'à dans son existence (Honeste 2007), la synonymie est pourtant un fait linguistique, attestée par de nombreux exemples. Le cas de *arriver* et de *venir* nous paraît intéressant en ce qu'il montre toute la richesse du phénomène. Loin de l'avoir explorée complètement, nous en avons évoqué à peine quelques aspects. Autorisée en langue, les relations synonymiques entre *arriver* et *venir* ont un caractère dynamique en discours. Dans cette étude, nous avons essayé de déterminer le rôle de l'énonciateur qui constitue « l'élément spatio-temporel relié à l'occurrence » verbale. Notre corpus montre que les co(n)-textes assimilateurs sont ceux dans lesquels l'énonciateur adosse un double rôle sur le plan discursif et sémantique, en tant que participant du déplacement. Il suffit cependant d'en modalisateur, comme dans (8), pour affaiblir le lien de synonymie. Autant dire que la synonymie aura toujours de quoi surprendre...

Bibliographie

- ASSER, N., SABLAYROLLES, P. (1995) : « A Typology and Discourse Semantics for Motion Verbs and Spatial PPs in French », *Journal of Semantics*, 12, 163-209.
- ASNÈS, M., KUPFERMAN, L. (2008) : « Événements, prédicats, arguments : quelques points de repère », *Langages*, 169, 7-33.
- BOONS, J.-P. (1985) : « Préliminaires à la classification des verbes locatifs : les compléments de lieu, leurs critères, leurs valeurs aspectuelles », *Linguisticae Investigaciones IX*, 2, 195-267.
- (1987) : « La notion sémantique de déplacement dans une classification syntaxique des verbes locatifs », *Langue française*, 76, 5-40.
- BOUCHARD, D. (1993) : « Primitifs, métaphores et grammaire : les divers emplois de *venir* et d'*aller* », *Langue Française*, 100, 49-66.
- BOURDIN, Ph. (1992) : « Constance et inconstance de la déicticité : la resémantisation des marqueurs anaditifs et ventifs », *La Dérive*, Paris, Presses universitaires de France, 287-307.
- (2005) : « *Venir* en français contemporain : de deux fonctionnements périphrastiques », Le Querler, N. et Bat-Zeev Shydkrot (éds), *Les périphrases verbales. Linguisticae Investigaciones Supplementa*, 25, 261-278.
- BOUVEROT, D. (2007) : « La synonymie dans le TLF », *Le français moderne*, t. LXXV, 1, 7-14.
- CRUSE, A. (2004). *Meaning in Language. An Introduction to Semantics and Pragmatics*, Oxford University Press.
- DUCHACEK, O. (1979) : « Synonymie en terminologie », *Études romanes de Brno*, X, 9-19.
- DUCROT, O., SCHAEFFER, J.-M. (1995) : *Nouveau dictionnaire encyclopédique des sciences du langage*, Paris, Éditions du Seuil.
- FUCHS, C. (1997) : « La synonymie en co-texte », Guimier, C. (éd.), *Co-texte et calcul du sens*, Presses universitaires de Caen, 31-41.

- HONESTE, M.-L. (2005) : « *Peut-il un verbe périphrastique ? Etude sémantico-cognitive* », Le Querler, N. et Bat-Zeev Shyldkrot (éds), *Les périphrases verbales. Linguistic Investigations Supplements*, 25, 295-310.
- (2007) : « *Entre ressemblance et différence : synonymie et cognition* », *Le français moderne*, t. LXXV, 1, 160-174.
- KLEIBER, G. (2008) : « *Comment fonctionne ICT* », Vuillaume, M. (éd) *Ici et maintenant. Cahiers Chronos*, 20, 113-145.
- LE QUERLER, N. (1996) : *Typologie des modalités*, Presses universitaires de Caen.
- NEVEU, F. (2004) : *Dictionnaire des sciences du langage*, Paris, Armand Colin.
- TALMY, L. (2000a) : *Toward a Cognitive Semantics : Concept Structuring System, Volume I*, Harvard, MIT Press.
- (2000b) : *Toward a Cognitive Semantics : Typology and Process in Concept Structuring, Volume II*, Harvard, MIT Press.
- VANDELOISE, C. (1987) : « *La préposition à et le principe d'anticipation* », *Langue Française*, 76, 77-111.

Ressources électroniques

DicoSyn/CRISCO, <http://www.ceril.fr/synonymie/> ;

Trésor de la Langue Française Informatisé, <http://atilf.atilf.fr/dendien>

PRATIQUES

PROCHAIN NUMÉRO
n° 143/144 (décembre 2009)



Écrits de savoir

Coordination : Mohamad Karim

L'écriture des savoirs engage un certain nombre de difficultés que ce numéro voudrait contribuer à caractériser pour mieux les appréhender et, le cas échéant, les didactiser. Un enjeu des écrits des savoirs, et non des moindres, procède de la nécessité à laquelle sont tenus les scripteurs de construire et de conduire des parcours épistémologiques propres à ce que l'on pourrait appeler la codification écrite des savoirs. Si les soubassements épistémologiques paraissent si pégnants dans les écrits de savoir, c'est que le référentiel y occupe une place prépondérante, fût-il du langage, dans la mesure où il constitue l'objet premier de la recherche. C'est en premier lieu la connaissance d'un objet qui motive l'initiation d'une recherche et par suite sa formalisation écrite. A cet égard, l'enjeu aujourd'hui est autant de communiquer le résultat de ses recherches que les recherches elles-mêmes, bref l'on peut dire que la *théorisation*, comme le rappelle Neveu est aussi importante que la théorie. Or l'écriture des savoirs est une activité nécessairement alimentée, à l'université comme dans le secondaire et dans le premier degré, par d'autres activités (l'observation d'un objet ou d'un phénomène, les lectures d'information ou d'investigation, les interactions orales entre élèves et professeurs, les discussions entre pairs, à titre d'exemple) qui en sont du même coup constitutives puisqu'elles contribuent à la genèse de la pensée scripturale. C'est la raison principale pour laquelle ce numéro en préparation associe le travail de *conception* à celui de *conceptualisation* des écrits complexes.

LE VERBE *COURIR* ET SES EXTENSIONS SPORTIVES

Résumé : L'article présente deux emplois du verbe courir qui relèvent clairement du domaine du sport. En effet, on dira d'un(e) athlète qu'il/elle court le marathon, alors que les participants d'un rallye automobile peuvent bien courir en moto. En partant de leur sens de base, qui est celui du déplacement, il s'agit de retracer les modifications sémantiques qui conduisent ce verbe dans le champ sémantique du sport. Il s'avère cependant que tout en relevant du domaine sportif, ces deux types d'emplois appartiennent à des classes verbales dont les membres se caractérisent par des propriétés aspectuelles et grammaticales fort différentes.

1. INTRODUCTION

Les dictionnaires font état de deux acceptions du verbe *courir* qui s'inscrivent dans le champ sémantique du sport. Les exemples (1) et (2) illustrent les emplois qui nous occuperont dans cet article.

- (1) *Michael Schumacher a encore couru en moto.* (Internet)
- (2) *Un des mineurs chiliens a couru le marathon de New York.* (Internet)

En analysant ces acceptions sportives qui constituent une modeste partie de la riche polysémie du vocable *COURIR*, nous nous efforcerons de croiser deux approches et deux cadres de travail, certes, éloignés, mais point incompatibles. D'une part, la Lexicologie Explicative et Combinatoire (désormais la LEC) – dont Mel'čuk, Clas et Polguère (Mel'čuk, Clas et Polguère, 1995) exposent les principes et que Mel'čuk et Polguère pratiquent, entre autres, dans *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français* (Mel'čuk et Polguère, 2007) – permet de suivre le chemin de la dérivation sémantique menant du sens de déplacement qu'illustre l'exemple (3) ci-dessous aux emplois sportifs dans les phrases (1) et (2).

- (3) *Quand elle franchit la barrière ouverte, Aimée courait presque, à pas raides.* (Frantext)

D'autre part, les analyses sémantiques développées par Levin (1993, 2005, 2006, 2009) et par Levin et Rappaport Hovav (2005) ont pour objectif d'élaborer des classes verbales qui réunissent des lexèmes sur la base des composantes du sens responsables des propriétés aspectuelles et grammaticales. La description qui en résulte présente une granularité nettement moins fine que celle élaborée dans le cadre de la LEC, mais en contrepartie elle contribue à mettre en relation certains types d'éléments sémantiques et les caractéristiques grammaticales des lexèmes verbaux. L'intérêt de rapprocher ces deux regards consiste, à notre avis, dans la possibilité d'observer le chemin par lequel les modifications subtiles explorées dans une perspective lexicologique conduisent à des transferts entre des classes verbales qui forment une typologie plus générale.

Notre réflexion est organisée en trois étapes. La première section sera consacrée, dans un premier temps, à la discussion des notions de champ sémantique et de champ lexical. Elle analysera ensuite le sens de *courir* en tant que verbe de manière de déplacement pour le comparer à celui de ses co-polysèmes sportifs. Dans la deuxième section, nous présenterons l'approche en termes de classes verbales pour identifier celles dont relève les acceptions sportives de *courir*. La troisième section nous permettra de conclure notre réflexion, en réunissant les informations que permettent de dégager les deux modes d'analyse.

2. COURIR ET SES ACCEPTIONS SPORTIVES DANS UNE APPROCHE LEXICOGRAPHIQUE

2. 1. Champs sémantiques et champs lexicaux

Les notions de champ sémantique d'une part, et celle de classe verbale s'inscrivent dans des modes d'analyse sémantique différents, mais qu'il est possible – à notre avis – d'aborder de manière complémentaire. Les deux permettent de proposer une structuration du lexique en microsystèmes par un traitement particulier des composantes identifiées dans le processus de décomposition du sens lexical. Lorsque l'analyse permet de dégager une composante commune à plusieurs lexies, qui les rattache toutes à un domaine d'expérience particulier, on admet qu'elles appartiennent au même champ sémantique. Mel'čuk, Clas et Polguère (Mel'čuk, Clas et Polguère, 1995 : 173) en proposent une définition formelle : dans une description lexicographique, les unités lexicales relevant d'un champ sémantique possèdent « une même composante identificatrice du champ », que celle-ci figure explicitement dans la définition ou non. Dans le cas qui nous intéresse, pour être admises dans le champ sémantique du sport, les acceptions de *courir* doivent être liées à la composante identificatrice de ce champ.

L'approche lexicographique développée dans le cadre de la LEC (pour une présentation, voir Mel'čuk, Clas et Polguère, 1995), distingue clairement la notion de champ sémantique et celle de champ lexical. À la différence du premier, le second regroupe non pas des lexies (acceptions), mais des

vocables dont l'acception de base relève d'un domaine d'expérience précisée¹. Ainsi, le substantif *match* appartient à la fois au champ sémantique et au champ lexical du sport dans la mesure où il se définit avant tout comme « rencontre entre deux concurrents ou deux équipes qui se mesurent en présence d'un arbitre »², acception que l'on classe sans hésitation, bien que sur des bases purement intuitives pour le moment, parmi les faits sportifs. On ne peut en dire autant de *courir*, dont le sens de base est paraphrasé sur les pages de la même ressource de la manière suivante : « se déplacer rapidement par un mouvement successif et accéléré des jambes ou des pattes prenant appui sur le sol »³. Si l'on s'en remet à cette définition, *courir* en tant que vocable, appartient au champ lexical du déplacement ou des activités physiques⁴, même si par ailleurs ses extensions polysémiques se situent dans des champs sémantiques relatifs à d'autres domaines de la vie, notamment dans celui du sport en ce qui concerne les emplois qui nous occupent ici.

En nous appuyant sur la notion de champ sémantique ainsi défini, nous serons donc en mesure de circonscrire de manière plus formelle celui du sport. Il nous faudra pour cela trouver la composante sémantique identificatrice de ce champ. Nous nous efforcerons de le faire dans la section suivante, en exploitant le travail lexicographique en cours dans le cadre du projet *Réseau Lexical du Français* développé actuellement au Laboratoire ATILF – CNRS (UMR 7118)⁵.

2. 2. À la recherche de la composante identificatrice du champ sémantique du sport

Afin de délimiter le champ sémantique du sport conformément à la définition que nous venons d'adopter, il nous faudra donc identifier une composante commune non seulement aux acceptions de *courir* attestées dans les exemples (1) et (2) ci-dessus, mais également à (beaucoup) d'autres unités lexicales telles que *match*, *jouer*, *stretching* et *jogging* intuitivement classées dans le domaine sportif. Précisons, en suivant le chemin tracé par Mel'čuk, Clas et Polguère (1995), que celle-ci n'occupe pas nécessairement une position centrale dans la paraphrase définitoire de l'unité lexicale examinée.

Il nous est, hélas, impossible de présenter d'une manière exhaustive tous les sens qui forment la polysémie du vocable COURIR. En partant de l'acception de base, nous ne présentons ensuite que ses co-polysèmes sportifs. Comme

1. Définition qui diffère sensiblement de celles proposées par d'autres chercheurs. À titre d'exemple, Neveu (2004 : 63) en propose la définition suivante « On parle fréquemment de *champ lexical* lorsque les lexèmes le constituant appartiennent à une même classe grammaticale (substantifs, adjectifs, verbes, etc.) ».

2. www.cnrtl.fr/definition/match

3. www.cnrtl.fr/definition/courir

4. Tout en nous appuyant sur la définition proposée par le Centre National des Ressources Textuelles et Lexicales (www.cnrtl.fr), il nous est impossible de ne pas évoquer de récents travaux en sémantique de l'espace avançant des arguments contre le traitement de *courir* et des verbes de manière de mouvement en termes de déplacement (cf. Aunargue, 2008).

5. <http://www.atilf.fr/spip.php?article908>

toutes les définitions du RLF, ces paraphrases sont structurées autour d'une composante centrale à laquelle s'articulent des composantes périphériques.

courir#I.a

L'individu X ~ :

<CC label=« se déplacer intentionnellement »>L'individu X se déplace</CC>

- <PC role=« milieu »>au sol</PC>
- <PC role=« vitesse.magnitude>très rapidement</PC>
- <PC role=« gestuelle »>en faisant des pas</PC>

courir#II.a

L'individu X ~ Y :

<CC label=« réaliser une épreuve sportive »>L'individu X fait une épreuve sportive Y</CC>

- <PC=« type.Y »>de course</PC>
- <PC=« distance parcourue »>consistant à *courir*#I.a sur la distance propre à Y</PC>

courir#II.b

L'individu X ~ dans Y

- <CC label=« participer »>L'individu X participe à une manifestation sportive Y</CC>
- <PC role=« type.Y »>de vitesse</PC>

Ces acceptations du verbe *courir* correspondent respectivement à nos exemples (3), (2) et (1). Les attributs des composantes périphériques permettent d'explicitier les modifications sémantiques qui s'opèrent dans les extensions polysémiques. Dans l'unité lexicale *courir*#II.a, le déplacement perd sa position centrale. Il reste présent à travers le sens I.a inclus dans la composante périphérique pour s'effacer ensuite quasi totalement dans la définition de *courir*#II.b. Il n'en reste plus que la vitesse dans le typage de la variable Y.

On peut admettre dès lors que l'entrée de ces deux sens de *courir* dans le champ sémantique du sport se fait par l'intermédiaire du second actant. On a donc affaire à des collocatifs dont le sens se définit par rapport à celui de la base de collocation que constitue le nom de l'épreuve sportive.

Une telle formulation des paraphrases définitoires permet de repérer la composante identificatrice du champ sémantique du sport à travers l'adjectif *sportif*, qui signifie *relatif au sport*. On remarque que sa présence est liée à celle du second actant Y, même si celui-ci peut être omis, comme cela a lieu dans l'exemple (1).

3. COURIR ET SES CLASSES VERBALES (LEVIN 2005, 2006, 2009)

3.1. Construction de classes verbales

La répartition des lexèmes verbaux en classes telle qu'elle est proposée par Levin (Levin, 2005, 2006, 2009) repose sur un autre mode de décomposition du sens lexical. L'objectif de cette démarche consiste non pas à construire des définitions analytiques, mais à identifier des composantes communes à un ensemble de verbes qui, d'une part, apparaissent comme sémantiquement apparentés, et de l'autre se caractérisent par des comportements grammaticaux similaires.

Pour identifier la classe à laquelle appartient un lexème verbal, il convient donc – selon Levin et Rappaport Hovav (Levin, Rappaport Hovav, 1999) et

Levin (Levin, 2005, 2006, 2009) – de décomposer son sens, en dégagant deux types d'éléments. Le premier, c'est le schéma, simple ou complexe, de la situation⁶ désignée par le verbe. Le schéma est simple, lorsqu'il comprend un seul élément structurel. Tel est le cas des activités (au sens de Vendler, 1967) illustré ci-dessous par *courir* dans l'exemple (3) :

- (3) *Quand elle franchit la barrière ouverte, Aimée courait presque, à pas raides.* (Frantext)

En (3), *courir* sélectionne un seul argument à caractère agentif qui syntaxiquement remplit la fonction de sujet. La structure du prédicat que forme le verbe dans cet emploi est donc monoargumentale et elle se voit associer une situation de type activité (Vendler, 1967, Dowty, 1979, Borillo, 1988, Récanati et Récanati, 1999), procès (Bach, 1986) ou processus (Caudal, 2006). Les activités, on le sait, sont duratives, structurées de manière homogène et, de ce fait, atéliques (Asnès, 2005). Remarquons que *courir* partage ce schéma de structure situationnelle avec les autres membres de la classe de verbes de manière de déplacement tels que *marcher* dans l'exemple (4) et *nager* dans (5) :

- (4) *L'ouvrier marchait et se déplaçait dans le décliv régulier de ses pas.*
 (5) *Hélène prête un maillot à Delphine qui nage longtemps, lentement, la tête bien droite, le regard vide.* (Frantext)

Une structure situationnelle, supposée commune à tous les membres d'une classe verbale, n'est évidemment pas indépendante du sens propre à chaque lexème. Elle apparaît comme tributaire de ce que Levin appelle la partie idiosyncrasique du sens du verbe. Si *courir#I.a* dans (3), tout comme *marcher* et *nager* dans (4) et (5), sont des activités, leurs paraphrases définitoires diffèrent par des composantes spécifiques. La démarche lexicographique adoptée dans le cadre du RLF est éclairante pour la comparaison :

marcher#I.a

L'individu X ~ :

<CC label=« se déplacer intentionnellement »>L'individu X se déplace</CC>

• <PC role=« milieu »>au sol</PC>

• <PC role=« géométrie »>debout</PC>

• <PC role=« gestuelle »>en faisant des pas</PC>

• <PC role=« background »>ce mode de déplacement ayant été appris par X</PC>

nager#I.b

L'individu X ~ (en utilisant la technique Y) :

<CC label=« se déplacer intentionnellement »>L'individu X se déplace</CC>

• <PC role=« milieu »>en suspension dans l'eau</PC>

• <PC role=« gestuelle »>par des mouvements de son corps (conformes à la technique Y)</PC>

• <PC role=« background »>ce mode de déplacement ayant été appris par X</PC>

Les définitions font ressortir les différences concernant le milieu, la position du corps, la gestuelle, voire le caractère inné ou appris du mode de

6. Dans la lignée de Smith (1991), nous employons le terme *situation* en tant que générique englobant les types de procès proposés dans les différentes approches classificatoires (activités, accomplissements, états, événements, etc.).

déplacement. Afin de construire de classes verbales, Levin propose de repérer ce qui est commun dans le sens de ces lexèmes, en les caractérisant par le type ontologique (*root*) qui les caractérise le mieux (par exemple, *contenant*, *instrument*, *résultat* et *manière*). Dans son sens de base, *courir*, tout comme *marcher* et *nager*, est ainsi qualifié de verbe de manière.

On pourrait, certes, poser la question de l'intérêt que présente une caractérisation aussi générale des sens spécifiques à un ensemble de lexèmes. Or, l'objectif en est d'articuler le sémantisme à la structure situationnelle d'une part, et aux caractéristiques grammaticales telles que la sélection des arguments de l'autre. Ces propriétés seraient difficiles, voire impossibles à associer à une composante spécifique, telle que, par exemple, la gestuelle que nous retrouvons régulièrement dans nos définitions. L'hypothèse adoptée dans l'approche proposée par Levin et Levin et Rappaport Hovav consiste donc à admettre un lien entre une configuration de composantes sémantiques caractérisée en termes de type ontologique et les traits en questions. On notera que ce n'est donc pas un trait particulier, par exemple la gestuelle, qui se voit associer un type ontologique. C'est une combinaison de composantes (gestuelle, milieu, géométrie, etc.) qui forme le type « manière ». Or, selon Levin, les verbes de manière sont essentiellement des activités, *ipse eo* des situations atéliques, qui sélectionnent de préférence des sujets agentsifs.

Il ne nous échappera pas cependant que la manière est une propriété relationnelle : elle n'existe pas *per se*, nous avons inévitablement affaire à la manière d'être ou de faire quelque chose. C'est la raison pour laquelle la classe des verbes de manière se répartit en sous-classes pour une granularité de description plus fine. On trouve, en effet, parmi les verbes de manière, ceux qui dénotent le déplacement, le mouvement, l'expression, etc. Il est important d'affiner ainsi le classement car c'est à ce niveau subordonné que l'on repère les traits tels que la transitivité ou l'intransitivité, les prépositions et le mode de réalisation syntaxique des arguments propres aux membres d'une sous-classe de verbes de manière.

La classe de verbes de manière est donc transversale dans la mesure où elle regroupe des lexèmes verbaux appartenant à plusieurs champs sémantiques. Dans la section suivante, nous réfléchirons à ce qu'une analyse en termes de classes verbales révèle des sens sportifs du verbe *courir*.

3. 2. Deux sens sportifs et deux classes verbales

3. 2. 1 *Courir#II.a en tant que verbe de réalisation*

Commençons par une poignée d'exemples tirés du Web qui illustrent le sens *courir#II.a* :

- (6) *Un des mineurs chiliens a couru le marathon de New York.*
- (7) *Le temps de Haggis au deuxième tour n'a pas été inscrit par les officiels. Au premier tour, il a couru le 200 mètres en 22,2.*
- (8) *Détail émouvant, Thierry avait couru à moto le 1er Rallye des Pharaons de Fenouil, il y a un quart de siècle, juste avant un terrible accident de la circulation, responsable de son handicap.*

Dans l'hypothèse lexicographique que nous avons émise à travers les définitions présentées dans la section 1.2, le sens en question apparaît comme plus proche de l'unité lexicale de base (*courir#I.a*) puisqu'il inclut celle-ci dans l'une de ses composantes périphériques. Par ailleurs, *courir#II.a* sélectionne aussi un sujet agentif contrôlant sa course.

Quant au second actant, l'expression qui dénote une épreuve sportive, il est syntaxiquement réalisé par un complément direct. *Courir#II.a* étant un verbe transitif, on ne s'étonnera donc pas qu'il accepte la transformation passive, aussi bien dans sa forme périprastique avec le complément d'agent explicitement mentionné, que pronominale qui interdit sa présence. Les exemples (9), (10) et (11) en témoignent.

- (9) *Le marathon de New York a été couru par l'un des mineurs chiliens/Le marathon de New York se court tous les ans.*
 (10) *Le 200 mètres a été couru par Haggis en 22,2./Le 200 mètres se court en 30 secondes.*
 (11) *Le 1^{er} Rallye des Pharaons de Fenouil avait été couru à moto par Thierry/Le Rallye des Pharaons se court à moto.*

En plus du passif périprastique, *courir#II.a* admet une construction en être + participe passé qui n'est pas un passif, mais qui véhicule une valeur aspectuelle et temporelle que ne possèdent ni *courir#I.a*, ni *courir#II.b*. Imaginons, en effet, qu'en se promenant dans les rues de New York le lendemain de sa performance, le mineur chilien rencontre une connaissance peu au fait du calendrier sportif. Celle-ci lui demande quand il va courir le marathon de New York. Une réponse que nous proposons sous (12) serait alors parfaitement adéquate, et surtout grammaticalement correcte. Dans le cas où quelques jours se seraient passés, il pourrait même prononcer (13) :

- (12) *Ah, le marathon est déjà couru.*
 (13) *Le marathon est déjà couru depuis longtemps.*

Les phrases (12) et (13) appellent une interprétation résultative. Nous avons, en effet, affaire à un état résultant d'une action accomplie, instauré au moment où le mineur chilien termine le marathon. La compatibilité avec les adverbiaux temporels en *depuis x temps* (13) renforce la lecture résultative.

Notons au passage la nécessité de distinguer un actant sémantique (*le marathon, le deux-cents-mètres*⁷ et *le Rallye des Pharaons*), qui est une expression dénotant une entité conceptuelle bien définie dans le contexte sportif et qui implique une réalisation de la part d'un agent, d'un circonstant tel que *vingt mètres* dans l'exemple (14) ci-dessous.

- (14) *Il y a un type qui a couru vingt mètres à côté de moi, dit le plus jeune.* (Frantext)

7. Notons que dans tous les exemples que nous avons trouvés le nom de l'épreuve est donné en chiffres et suivi d'abréviation (le 200 m), et non pas en toutes lettres. Nous proposons ailleurs (Sikora, 2012 : à paraître) une analyse de ce syntagme nominal qui nous conduit à le qualifier de nom composé formé selon les règles dérivationnelles du français comme *un deux-roues* ou *un trois-pièces*.

Le circonstant *vingt mètres* n'autorise, en effet, aucune des transformations effectuées sur les exemples (9) à (11). Il est par ailleurs facilement remplaçable par une expression d'évaluation quantitative, par exemple *un peu*. Il en ressort que le verbe *courir* employé dans l'exemple (14) relève de l'unité lexicale de base, c'est-à-dire de *courir#I.a.*

Revenons à présent sur le rôle qu'assume l'expression du second actant sur le plan de la structure situationnelle. La possibilité d'insérer, dans la phrase, l'adverbial temporel *en x temps*, comme nous l'avons fait dans l'exemple (10) et l'état résultant qui apparaît avec la construction être + participe passé indiquent une structure situationnelle télélique.

Il en ressort que *courir#II.a* dénote un type de situation bien différent de celui que l'on associe au prédicat monoargumental que forme la lexie *courir#I.a.* Dans les exemples (9), (10) et (11), nous avons affaire non pas à une activité, comme dans (3) et (14), mais à un accomplissement.

Or, dans l'approche développée par Levin, la structure situationnelle est décisive pour l'appartenance à une classe verbale, puisqu'elle doit être commune à tous ses membres. Il convient alors de se poser la question de savoir à quelle classe appartient le verbe *courir* dans son sens sportif II.a.

Pour proposer une réponse, observons les exemples (15) et (16) ci-dessous :

(15) *C'est lui qui fait du théâtre, il écrit et met en scène.* (Frantext)

(16) *Alexandre a écrit un roman et a même trouvé un éditeur.* (Frantext)

Dans (15), le verbe écrire ne sélectionne qu'un seul actant, tout comme l'unité lexicale *courir#I.a.*, en dénotant une situation que l'on classe parmi les activités. Avec l'introduction du second actant, *un roman* dans (16), la structure situationnelle est modifiée, car elle possède désormais un terme : elle se déroule à travers le temps tant que le roman est en cours d'élaboration et s'arrête quand le processus créatif arrive à son terme, c'est-à-dire lorsque le roman est écrit. Elle doit être classée non pas parmi les activités, mais parmi les accomplissements.

Vu sous cet angle, *courir#II.a* présente des traits aspectuels communs avec *construire*, *écrire*, *peindre*, *ré citer* dans leurs emplois biactanciels. Compte tenu de la nature des relations qui s'établissent alors entre les deux actants, on peut les qualifier de verbes de réalisation. Au fur et à mesure de l'activité exercée par X, la réalisation de Y progresse jusqu'à atteindre son terme.

Certes, des distinctions à l'intérieur de la classe de verbes de réalisation peuvent être affinées. On comprend clairement qu'un roman n'existe qu'une fois le processus de création achevé, ce qui n'est pas le cas d'un marathon ou d'une épreuve de deux-cents mètres qui préexistent à l'exécution en tant qu'entités définies en intension. Il serait alors légitime de parler de *verbes de création* dans le premier cas et de *verbes de performance* dans le second. Nous pensons, en effet, que les exemples de *courir#II.a* relèvent du second cas de figure, c'est-à-dire de celui où l'épreuve sportive, que ce soit le marathon, le deux-cents mètres ou le Rallye des Pharaons, est conceptualisée comme un objet « réalisé » par X. En courant le marathon, le coureur réalise

une performance d'un certain type préexistant, du moins conceptuellement, à ses réalisations particulières.

Plus qu'une ressemblance de famille, le transfert de *courir*#II.a dans la classe des verbes de réalisation se confirme sur le plan grammatical. En effet, tout comme les autres membres de cette classe, *courir*#II.a est transitif et autorise la transformation passive.

3. 2. 2. *Courir*#I.b en tant que verbe de participation

Reprenons l'exemple (2), renuméroté en (17) ci-dessous, avec deux autres emplois du verbe *courir* dans son acception sportive #II.b que nous avons isolés sur Internet.

- (17) *Michael Schumacher a encore couru en moto.*
- (18) *Les poussines nées en 2001 ont couru au championnat des Mauges de cross. Cette course est réservée aux enfants de 6 à 13 ans.*
- (19) *Mis à l'eau pour la première fois en 2007, le bateau a couru sous le nom de « Mowgli » avec Jeremy Salvesen et David Thomson lors des « Portimao Global Ocean Race » de 2008 et de 2009.*

La description lexicographique que nous avons proposée de ce sens de *courir* présente un écart plus important par rapport à l'unité lexicale de base. Parmi les composantes qui forment le type ontologique *manière*, seule la vitesse reste saillante. L'exemple (18) est, certes, quelque peu ambigu dans la mesure où son sujet dénote un ensemble d'individus capables de *courir*#I.a. Cependant, il pourrait également être asserté des jeunes compétitrices même si, suite à une baisse de forme physique, elles avaient marché pendant une ou plusieurs étapes de la compétition.

Les exemples (17) et (19) illustrent les conséquences de ce que l'on peut considérer comme une neutralisation, du moins partielle, de la manière de déplacement. Il est désormais possible d'adjoindre au verbe *courir* un circonstant tel que *en moto* ou à vélo, dont nous avons également trouvé des attestations, incompatibles avec l'unité lexicale de base. Par ailleurs, le sujet n'est pas nécessairement agentif, comme le montre le cas du bateau dans l'exemple (19). Enfin, la phrase (17) montre que l'expression du second actant n'est point nécessaire pour actualiser le sens *courir*#II.b.

Que devient alors la manière de déplacement dans le sens de *courir*#II.b? Peut-on réellement admettre qu'il n'en reste que la vitesse? Le caractère quelque peu étrange des exemples (20) et (21) semble bien indiquer le contraire.

- (20) # *Émilienne Salmon a couru aux Championnats d'Europe de tir en Campagne.*
- (21) # *Viswanathan Anand a couru au Championnat du monde d'échecs 2010.*

Ce sens de *courir* ne peut s'employer qu'avec un nom de manifestation sportive qui se définit en référence au déplacement réalisé avec une certaine vitesse, décisive pour la victoire. On dira donc, en résumé, que le sens de déplacement est déterminant pour le second actant. Rappelons cependant

deux faits : d'une part, le second actant peut rester inexprimé (exemple 17), de l'autre *courir#II.b* n'autorise pas n'importe quel nom de compétition en position de second actant (exemples 20 et 21). Cela porte à croire que tout en dénotant le fait de participer, de concourir dans une manifestation sportive, le verbe joue un rôle dans la sélection de l'argument en position du second actant. Si tel est le cas, son sens n'est pas uniquement celui de participation, car dans ce cas, il n'y aurait pas de restriction quant à la nature de la manifestation sportive. Il est donc question, dans le cas de *courir#II.b*, d'une manière de participer ou de concourir.

La configuration des composantes sémantiques de *courir#II.b* ne se répercute pas sur le plan de la structure situationnelle. Tout comme l'acception précédente, ce *courir* sportif possède une structure biactancielle, mais celle-ci s'exprime de manière intransitive. L'impossibilité de former un passif n'a donc rien d'étonnant. Par ailleurs, aucune manipulation – construction en être + participe passé, adjonction de complément temporel *depuis x temps* – ne permet de mettre en avant l'existence d'un état résultant. Enfin, les exemples (17) à (19) refusent l'adverbial temporel *en x temps*, comme l'illustrent les phrases (22) à (24) :

- (22) * *Les poussines nées en 2001 ont couru au championnat des Mauges de cross en 30 minutes. Cette course est réservée aux enfants de 6 à 13 ans.*
- (23) * *Mis à l'eau pour la première fois en 2007, le bateau a couru en trois heures sous le nom de « Mowgli » avec Jeremy Salvesen et David Thomson lors des « Portimao Global Ocean Race » de 2008 et de 2009.*
- (24) * *Michael Schumacher a encore couru en moto en une heure.*

L'ensemble de ces caractéristiques conduit à penser que *courir#II.b* dénote des situations atéliques, dépourvues de bornes temporelles claires, même lorsqu'il est employé à un temps grammatical à valeur perfective. La comparaison avec un emploi de *courir#II.a* est éclairante à cet égard. Reprenons l'exemple (6) discuté dans la section 2.2.

- (6) *Un des mineurs chiliens a couru le marathon de New York.*

L'interprétation de cette phrase consiste à admettre que le mineur a couru le marathon dans sa totalité, autrement dit, il l'a réalisé en entier. On ne peut pas en être complètement assuré dans le cas de (17), de (18) et de (19) où, sans supplément d'information contextuelle, on pourra toujours se demander si l'épreuve de vitesse fut parcourue dans sa totalité ou en partie seulement.

Récapitulons les éléments dégagés dans ces analyses. Elles montrent, dans un premier temps, que *courir#II.b* possède les mêmes propriétés aspectuelles que l'unité lexicale de base – verbe de manière de déplacement. La structure situationnelle atélique qui les caractérise en tant qu'activités au sens de Vendler (1967), est responsable des effets sémantiques produits par certaines formes grammaticales. Nous pensons notamment à l'ambiguïté quant au fait d'exécuter la totalité ou juste une partie de l'épreuve de vitesse que nous venons de mentionner pour les exemples (17) et (19).

4. CONCLUSIONS

Les deux analyses – lexicologique et lexicographique menée dans le cadre de la LEC d'une part, celle en termes de classes verbales proposée par Levin – mettent en avant des aspects différents des co-polysèmes sportifs du verbe *courir*.

L'étape lexicologique et lexicographique est indispensable dans toute réflexion sémantique pour identifier les liens et les mécanismes par lesquels les deux lexèmes, *courir#II.a* et *courir#II.b*, se rattachent au champ sémantique du sport. Elle nous permet de suivre les évolutions sémantiques, par exemple celles qui mènent du champ sémantique du déplacement à celui du sport, en nous interrogeant systématiquement sur le statut et sur la place des composantes de sens.

Nous tenons à souligner que le second mode d'analyse, celui qui vise la répartition des lexèmes en classes verbales, ne saurait en aucun cas, se passer du premier. Cependant, cette démarche constitue une sorte de réhabilitation du statut verbal de nos collocatifs. En effet, si le rattachement au champ sémantique du sport se fait, dans le cas de *courir#II.a* et dans celui de *courir#II.b*, par le truchement du second actant, le verbe subit des modifications sémantiques qui se répercutent sur son fonctionnement grammatical. Sur ce plan, nos deux *courir* sportifs présentent des différences frappantes qui autorisent à les répartir dans deux classes distinctes : celle de verbes de réalisation (ou de performance) pour *courir#II.a* et celle de verbes de participation pour *courir#II.b*.

Dorota SIKORA

RÉFÉRENCES BIBLIOGRAPHIQUES

- ASNÈS, M. 2005. « Incompatibility between telicity and homogeneity in French », *Cahiers Chronos*, n° 13. Amsterdam-New York, Rodopi, p. 93-103.
- AUNARGUE, M. 2008. « Qu'est-ce qu'un verbe de déplacement ? Critères spatiaux pour une classification des verbes de déplacement intransitifs du français », dans *Actes du Premier Congrès Mondial de la Linguistique Française*. Paris, Institut de Linguistique Française, CMLF'08.
- BACH, E. 1986. « The Algebra of Events », *Linguistics and Philosophy*, n° 9. Dordrecht, Boston, Reidel, p. 5-16.
- BORILLO, A. 1988. « L'expression de la durée : constructions des noms et des verbes de mesure temporelle », *Lingvisticae Investigationes*, n° 12/2, p. 363-396.
- BORILLO, A. 1998. *L'Espace et son expression en français*. Paris, Ophrys, 170 p.
- CAUDAL, P. 2006. « Aspect », dans D. Godard, L. Roussarie, F. Corblin (dir.), *Sémantopédie : dictionnaire de sémantique*, GDR Sémantique & Modélisation, CNRS (<http://www.semantique-gdr.net/dico/>).
- CREISSELS, D. 2000. « L'emploi résultatif de être + participe passé en français », *Cahiers Chronos*, n° 6. Amsterdam-Atlanta, Rodopi, p. 133-142.
- DOWTY, D.R. 1979. *Word Meaning and Montague Grammar*. Dordrecht, Reidel, 415 p.

- . 1991. « Thematic Proto-Roles and Argument Selection », *Language*, n° 67. Washington DC, Linguistic Society of America, p. 547-619.
- LASKOWSKI, R. 1998. « Kategorie morfologiczne języka polskiego – charakterystyka funkcjonalna », dans R. Grzegorzczkowska, R. Laskowski, H. Wróbel (dir.), *Gramatyka współczesnego języka polskiego. Morfologia*. Warszawa, Wydawnictwo Naukowe PWN, p. 151-203.
- LEVIN, B. 1993. *English Verb Classes and Alternations : a Preliminary Investigation*. Chicago : The University of Chicago Press, 348 p.
- . 2005. « Semantic Prominence and Argument Realization V. Structuring Event Structure », Cours 123, MIT (www.stanford.edu/~bclevin/lisa05evstr.pdf)
- . 2006. « Lexical Semantics and Argument Realization I. Empirical and Conceptual Challenges of Argument Realization » (www.stanford.edu/~bclevin/dgfs06.html-4)
- . 2009 « Where do the verb classes come from? » Communication à la conférence : *Verb Typologies revisited. A Cross-linguistic Reflection on Verbs and Verb Classes*, Université de Gand, Belgique, 5-7/02/2009, (www.stanford.edu/~bclevin/pubs.html).
- LEVIN, B., RAPPAPORT HOVAV, M. 1999. « Two Structures for Compositionally Derived Events », dans T. Matthews, D. Strolovitch (dir.), *Semantics and Linguistic Theory 9*. Ithaca New York, Cornell University Press, p. 199-223.
- . 2005. *Argument Realization*. Cambridge, Cambridge University Press, 278 p.
- MEL'ČUK, I., CLAS, A., POLGUÈRE, A. 1995. *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve, Éditions Duculot, 256 p.
- MEL'ČUK, I., POLGUÈRE, A. 2007. *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20000 dérivations sémantiques et collocations du français*. Bruxelles, De Boeck, 525 p.
- NEVEU, F. 2004. *Dictionnaire des sciences du langage*. Paris, Armand Colin, 316 p.
- PARSONS, T. 1990. « Events in the Semantics of English : A Study in Subatomic Semantics », dans E. Le Pore, B. McLaughlin (dir.), *Actions and Events : Perspectives on the Philosophy of Donald Davidson*. Londres, Blackwell, p. 235-267.
- PUSTEJOVSKY, J. 1991. « The syntax of event structure », *Cognition*, n° 41. Amsterdam, Elsevier Science, p. 47-81.
- RÉCANATI, C., RÉCANATI, F. 1999. « La classification de Vendler revue et corrigée », dans S. Vogeleer, A. Borillo, M. Vuillaume, C. Veters (dir.), *Cahiers Chronos*, n° 4. Amsterdam-Atlanta, Rodopi, p. 167-184.
- SIKORA, D. (à paraître) « Compléments aspectuels et verbes de manière de déplacement en français : entre bornage et téléicité », dans *Études Cognitives/ Cognitive Studies*, n° 12/2012. Instytut Slawistyki Polskiej Akademii Nauk.
- SMITH, C. 1991. *The Parameter of Aspect*. Dordrecht, Boston, Londres, Kluwer.
- VENDLER, Z. 1967. « Verbs and Times », dans *Linguistics in Philosophy*. Ithaca New York, Cornell University Press, p. 97-121.

Ressources électroniques

Centre National de Ressources Textuelles et Lexicales (www.cnrtl.fr)

Base textuelle Frantext (www.frantext.fr)

CHAPTER NINE

NATION ET NARÓD : ANALYSE SEMANTIQUE D'UN COUPLE DE FAUX-AMIS FRANCO-POLONAIS

DOROTA SIKORA

UNIVERSITÉ DU LITTORAL – CÔTE D'OPALE HLLI EA 4030

Abstract

The aim of this paper is to bring into focus conceptual and semantic discrepancies between the French lexeme *NATION* and the Polish one *NARÓD*. According to bilingual dictionaries, they are perceived as semantically equivalent. Indeed, corpora provide a huge sample of similar uses, especially in the domain of legal discourse. Notwithstanding this fact, an analysis of the way French and Polish speakers deal with *NATION* and *NARÓD* in every day communication reveals significant differences with respect to both semantic content of these lexemes and the underlying concepts. Within the frame of Explanatory and Combinatorial Lexicology, the article explores semantic networks organized around these lexemes as well as their combinatorial properties. The results of the investigation lead to the conclusion that there is a crucial difference between the linguistic status of *NATION* in French and the linguistic status of *NARÓD* in Polish. While the former displays properties of an abstract noun with a lexical content determined by a set of features that the members of a national community are supposed to share, the semantics of the latter puts forward the relational nature of the group. As a matter of fact, *NARÓD* seems to be connected to the same conceptual paradigm as *RODZINA* (family) and it clearly functions as a concrete term in the Polish lexical system.

Résumé

Le présent article propose une analyse comparée des lexèmes nominaux NATION en français et NAROD en polonais. Conformément aux dictionnaires bilingues, ces deux lexies devraient être considérées comme sémantiquement équivalentes et de fait, on trouve fréquemment des exemples d'emplois confirmant cette équivalence, notamment dans le domaine juridique. Il est frappant de constater cependant que, en dehors du discours légal, l'usage que les locuteurs des deux langues font de ces termes permet de repérer des différences significatives sur le plan conceptuel et sémantique. En adoptant le cadre méthodologique de la Lexicologie Explicative et Combinatoire, l'article s'intéresse aux réseaux de relations paradigmatiques et syntagmatiques construits autour de ces lexèmes dans les systèmes lexicaux du français et du polonais. L'analyse de ces liens sémantico-syntaxiques révèle des différences suffisamment marquantes pour considérer que NATION et NAROD sont en réalité de faux-amis. En français, le nom collectif NATION se définit par un ensemble d'attributs partagés par les membres de la collectivité à une certaine étape de l'histoire sociale. Il possède ainsi toutes les caractéristiques d'un nom abstrait. Pour un polonophone, NAROD dénote une communauté définie par un ensemble de relations qui existent entre ses membres. Le lexème polonais s'inscrit dans le même paradigme conceptuel que RODZINA (famille) et fonctionne de ce fait en tant que nom concret.

1. Introduction

1.1 Présentation : NATION et NARÓD – amis ou faux-amis ?

Chacun le sait : en politique et dans le discours politique, on se méfie de faux amis tout autant qu'on s'efforce de distinguer des faux-amis, lorsque l'on apprend une langue étrangère. En didactique des langues, le terme de *faux-amis* désigne des paires de lexèmes, respectivement dans les langues L_1 et L_2 , qui, en dépit d'une similarité de signifiants, ont des signifiés différents. RECEPTA en polonais et RECETTE en français, abordés en synchronie, en sont un cas d'école : le premier dénote une 'ordonnance délivrée par un médecin', le second une 'indication détaillée de la manière de préparer un mets' (Le Petit Robert 2011).¹

¹ On remarquera cependant que RECETTE avait autrefois le sens médical de 'formule et manière de préparer (un remède)' qui indique un étymon commun.

Dans le présent article, nous nous intéressons aux vocables **NATION**² en français et **NARÓD** en polonais et, plus précisément, à leurs sens politiques, c'est-à-dire aux lexèmes **NATION** et **NARÓD** I dénotant une collectivité. Selon les dictionnaires bilingues courants, leurs sens ne diffèrent guère. Traduits l'un par l'autre, ils présentent effectivement un certain nombre d'emplois similaires, dans lesquels ils désignent un acteur collectif de la vie politique. Il suffit de comparer les Constitutions des deux pays pour s'en convaincre. Ainsi, le titre III, article 20 de la Constitution française stipule que « Le Gouvernement détermine et conduit la politique de la nation »,³ alors que le chapitre IV de Konstytucja Rzeczypospolitej Polskiej précise, dans l'article 104, que « Posłowie są przedstawicielami Narodu. » (*Les députés sont des représentants de la Nation*).⁴

En dehors de ces emplois juridiques, **NARÓD** est largement employé par les locuteurs du polonais dans des contextes qui ont de quoi étonner un francophone. « Czy kochasz swój naród ? » (*Aimes-tu ta nation ?*) – titrait récemment un quotidien polonais, en présentant un sondage mené auprès d'un public de jeunes qui, par ailleurs, ne s'étonnaient nullement d'avoir à répondre à une telle question. Or, traduite en français, la question semble provoquer, au contraire, un léger malaise dû à une forte coloration nationaliste.

L'objectif de cet article est de montrer que **NATION** et **NARÓD** malgré un air de famille certain, sont en réalité de faux-amis, même si pour cerner les différences de sens, moins patentes que pour le couple **RECETTE** – **RECEPTA**, il nous faudra aller au-delà des emplois constitutionnels, somme toute peu révélateurs de différences, car relevant de la terminologie, pour étudier ce que ce « mot », politique par excellence, signifie pour l'ensemble des utilisateurs des deux langues. Cela nous semble important, car comprendre les différences conceptuelles qui se manifestent dans le lexique d'une langue est essentiel pour éviter des projections souvent réductrices, responsables de jugements hâtifs.

² Le vocable, transcrit en petites capitales, est un ensemble d'unités lexicales (lexies, sens, acceptions), directement ou indirectement liées sur le plan sémantique, qui ont le même signifiant (voir : Mel'čuk, Clas et Polguère 1995 : 159). La Lexicologie Explicative et Combinatoire conceptualise la polysémie comme apanage du vocable, et non pas de l'unité lexicale. En bref, un vocable polysémique regroupe au moins deux, et souvent plusieurs sens (acceptions, lexèmes, lexies, unités lexicales).

³ En ligne : http://www.assemblee-nationale.fr/connaissance/constitution.asp#titre_2.

⁴ En ligne : <http://www.sejm.gov.pl/prawo/konst/polski/konl.htm>.

1.2 Cadre théorique et méthodologique

Notre réflexion s'appuie sur le travail lexicographique du Réseau Lexical du Français (RL-fr),⁵ base de données lexicales développée dans le cadre théorique et méthodologique de la Lexicologie Explicative et Combinatoire (Mel'čuk, Clas, Polguère 1995, désormais la LEC). Parmi plusieurs champs lexicaux en cours de traitement figure celui de la politique, avec le vocable **NATION**.

Cependant, une description lexicographique n'est qu'une dernière étape d'un travail complexe d'analyses menées sur plusieurs plans : conceptuel et sémantique, syntaxique. On est ainsi amené à traiter plusieurs types d'entités linguistiques. Dans un souci de clarté, nous nous efforcerons de différencier les entités linguistiques mentionnées par des choix typographiques. Ainsi, les vocables et les lexèmes sont transcrits en petites capitales, mais les premiers sont en plus indiqués en gras (voir section 1.1, note 2).

L'objectif de notre article est double. D'une part, il s'agit de mettre en évidence les différences conceptuelles lexicalisées par **NATION** et **NARÓD**. Pour ce faire, nous commencerons, dans la section 2, par un examen attentif des informations disponibles dans les dictionnaires courants de français et de polonais. La section suivante sera consacrée à l'analyse des moyens lexicaux que les deux langues mettent en œuvre pour désigner un ou plusieurs individus appartenant à la collectivité. La quatrième partie examinera les liens syntagmatiques qui structurent les collocations formées autour de **NATION** et de **NARÓD** (4.1) et les indices lexicaux fournis par la combinatoire libre de nos lexèmes (4.2) ; un bilan d'étape en sera proposé en 5. Ces analyses nous conduiront vers l'hypothèse exposée et discutée dans la section 5 : **NATION** et **NARÓD** sont tous les deux des noms collectifs, mais le premier est un nom abstrait, alors que le second présente toutes les caractéristiques d'un nom concret. Plutôt que de conclure, nous proposerons, dans la section 7, une ébauche des définitions lexicographiques pour les deux lexèmes.

Notre second objectif est plus général. Au-delà du cas particulier qui nous occupe ici, nous tenons à montrer une fois de plus⁶ à quel point les

⁵ La première phase du projet, réalisé sous la direction d'Alain Polguère (Université de Lorraine & ATILF-CNRS) au Laboratoire ATILF-CNRS (UMR 7118), s'étale de 2011 à 2014. Pour une présentation détaillée, voir : Lux-Pogodalla et Polguère (2011).

⁶ De nombreux travaux s'inscrivant dans le cadre théorique de la LEC, et plus largement dans celui de l'approche Sens-Texte, ont déjà montré à quel point l'appareillage conceptuel qu'elles proposent est précieux pour les analyses

concepts et les formalisations de la LEC sont des outils précieux pour l'analyse et description des données multilingues. Par sa rigueur, cet outillage sert de fil conducteur, en balisant une réflexion qui avance sur deux chemins parallèles.

2. Analyse de données dictionnairiques

Nous commencerons cette section par une comparaison des structures polysémiques des vocables **NATION** et **NARÓD**, pour nous intéresser ensuite aux composantes sémantiques qui permettent de définir leurs sens politiques, c'est-à-dire les lexèmes **NATION** et **NARÓD**.

2.1 Polysémie « fuyante » de **NATION** versus deux sens stables de **NARÓD**

Du côté français, il est frappant de constater combien les structures polysémiques du vocable **NATION** diffèrent d'un dictionnaire à l'autre. Le *Petit Robert* 2011 en identifie six sens. *Le Trésor de la Langue Française* dans sa version informatisée (TLFi) en distingue quatorze, parmi lesquels on trouve des acceptions que nous qualifierions d'historiques. En effet, elles correspondent à ce que ce nom signifiait pour les locuteurs du français à certaines périodes de l'histoire. Ainsi, au Moyen-Âge, on appelait **NATION** (acception B-1.b dans le TLFi) un « groupement de maîtres et d'étudiants à l'intérieur des universités, selon un découpage linguistique », alors que sous l'Ancien Régime, ce nom était réservé au Tiers État.

La description du vocable **NATION** dans le TLFi rend ainsi compte des formes que prend la vie en société. Les sens se construisent selon les conditions historiques et les régimes politiques successifs, et ils sont de ce fait directement liés au cadre institutionnel dans lequel ils fonctionnent. Se pose dès lors une question plus large, demeurant valide pour toute une série de notions dont le contenu évolue avec la vie sociale : faut-il, dans un dictionnaire, rendre compte de tous les changements que subit un sens à travers l'histoire ? La réponse est négative, et ceci pour au moins deux raisons. Premièrement, en élaborant une paraphrase définitoire, en l'occurrence celle du lexème **NATION**, le lexicographe est tenu d'y intégrer des informations sémantiques, et non pas historiques. Ces dernières intéresseront plutôt les auteurs d'entrées encyclopédiques. Deuxièmement,

contrastives dans un contexte multilingue. Voir, à titre d'exemple : Elnitsky et Mel'čuk (1984), Iordanskaja et Mel' čuk (1997).

un locuteur contemporain du français emploie le lexème NATION, sans nécessairement connaître les aléas de son évolution en diachronie.⁷

Comparativement, NARÓD est parfaitement consensuel sur le plan de sa structure polysémique. Les sens historiques sont absents de tous les dictionnaires que nous avons consultés. Comparée aux fluctuations du vocable NATION, la polysémie de NARÓD fait preuve d'une grande stabilité d'un dictionnaire à l'autre. On retrouve partout les deux acceptions suivantes, reliées par une relation de métonymie allant d'un tout à une partie :

- (I) 'Ogół mieszkańców pewnego terytorium mówiących jednym językiem, związanych wspólną przeszłością i kulturą, i mającą świadomość tej wspólnoty' (Sobol, 2003 : 502)
*Totalité des habitants d'un territoire, parlant une même langue, liés par un passé et par une culture communs, conscients de cette communauté.*⁸
- (II) 'Ludzie, tłum'
Des gens, une foule (Sobol, 2003 : 502)
 Exemple : Górnicy, to taki dziwny naród, co lubi pracować.
*Les mineurs, c'est un naród étrange, qui aime bien travailler.*⁹

2.2 Composantes définitionnelles

Parmi les quatorze acceptions de NATION décrites dans le TLFi, nous avons sélectionné celle qui semble la moins « chargée » du point de vue historique et de ce fait, indépendante des connaissances que peut (ne pas) avoir le locuteur actuel du français. Désormais, nous ne parlerons que de NATION au sens défini ci-dessous. Du côté polonais, nous prendrons en compte uniquement NARÓD I. En nous focalisant sur ces lexèmes, nous n'indiquerons plus leurs numéros. Dans le cas de NATION, il serait difficile

⁷ Loin de nous l'idée de nier l'intérêt des recherches en diachronie au sujet notamment de NATION. La démarche qui est celle du RL-fr repose sur l'analyse du fonctionnement des unités lexicales en synchronie, car celui-ci est décisif pour la combinatoire de la lexie et pour l'identification de sa place dans le réseau que forme le lexique.

⁸ C'est nous qui traduisons les paraphrases proposées par Sobol (2003). Par ailleurs, tous les exemples et citations sont traduits par nos soins.

⁹ Dans les traductions de nos exemples, nous ne remplaçons pas NARÓD par un équivalent français pour éviter, autant que faire se peut, l'influence du concept lié à NATION en français. C'est dans le souci d'en faciliter la lecture que nous abandonnons les gloses spécifiant des informations grammaticales telles que la partie du discours, le cas ou le temps grammatical.

d'en appliquer un d'entrée de jeu, dans la mesure où notre intention est de décider, en fonction des faits de langue exposés, si le vocable **NATION** est polysémique ou non. Concernant **NARÓD**, nous espérons que la présente mise au point suffira pour éviter les confusions.

En analysant le micro-champ sémantique que forment, en français, les lexèmes **ÉTAT**, **NATION**, **PATRIE**, **PAYS**, **PEUPLE**, Rémi-Giraud (1996) distingue trois traits fondamentaux – humain, géographique et structurel – que l'on retrouve dans les définitions des dictionnaires consultés. Le trait humain est instancié par des formules telles que « ensemble d'êtres humains » pour **PEUPLE**, « groupe humain » pour **NATION** et « groupement humain » pour **ÉTAT**. Les syntagmes « territoire défini » dans le cas de **PAYS** et « fixé sur un territoire déterminé » dans celui de **NATION** précisent le trait géographique, alors que le trait structurel s'exprime par des formules telles que « institutions », « communauté politique », etc.

Tout en gardant cette grille d'analyse, nous parlerons non pas de traits fondamentaux, mais de composantes définitionnelles. En effet, les paraphrases du RL-fr sont des définitions aristotéliennes, structurées autour d'une composante centrale (désormais CC), exprimant le genre prochain de la lexie définie, à laquelle se rattachent des composantes périphériques (CP), qui en précisent les différences spécifiques.¹⁰

Comparons à présent les définitions de **NATION** et de **NARÓD**. Les composantes humaine, géographique et structurelle sont indiquées en gras. Le tableau 1 réunit leurs formulations en français et en polonais.

NATION, nom féminin

Groupe humain stable, établi sur un territoire défini constituant une unité économique, caractérisé par une auto-conscience ethnique (marquée par l'idée de la communauté d'origine et de destinée historique), une langue et une culture communes, **formant une communauté politique personnifiée par une autorité souveraine et correspondant à un stade évolué du mode et des rapports de production.** (TLFi)

NARÓD, nom masculin

Ogół mieszkańców pewnego terytorium mówiących jednym językiem, związanych wspólną przeszłością i kulturą, i mającą świadomość tej wspólnoty. (Sobol 2003 : 502)

'Totalité des habitants d'un territoire, parlant une même langue, liés par un passé et par une culture communs, conscients de cette communauté.'

¹⁰ Certaines acceptions nécessitent l'introduction de composantes présuppositionnelles. Pour une présentation détaillée des définitions du RLF, voir Sikora (2012).

| composante | NATION (TLFi) | NARÓD (Sobol, 2003) |
|--------------|--|---|
| humaine | groupe humain | ogół mieszkańców <i>totalité des habitants</i> |
| géographique | établi sur un territoire défini | pewnego terytorium <i>d'un territoire</i> |
| structurelle | <ul style="list-style-type: none"> • [territoire défini] constituant une unité économique • formant une communauté politique personnifiée par une autorité souveraine et correspondant à un stade évolué du mode et des rapports de production | – |

Tableau 1. Composantes définitionnelles de NATION (TLFi) et de NARÓD (Sobol 2003)

Deux points importants doivent être soulignés. Premièrement, dans les deux cas, la CC porte le trait 'humain'. Nous tenons cependant à souligner une différence qui, tenue à cette étape de notre exposé, ne tardera pas à renforcer d'autres indices linguistiques mis en avant dans nos analyses. Il s'agit de la formulation adoptée par les lexicographes. En français, on recourt au syntagme « groupe humain », qui a une structure syntaxique « N Adj ». Le nom collectif NATION est défini par recours à un nom général ou métacollectif (Lammert 2008) tel que GROUPE. Or, les métacollectifs se caractérisent par une homogénéité interne liée au fait que tous les éléments qu'ils réunissent (un groupe est nécessairement un groupe des X, peu importe le moyen d'expression adjectival ou nominal de l'actant X) sont présentés comme appartenant à une seule et même catégorie.

En polonais, le syntagme « ogół mieszkańców », que nous traduisons par « totalité des habitants », a une structure que l'on pourrait représenter comme « N₁ N₂Génitif pl ». N₁ est un élément nominal, qui quantifie les individus dénotés par N₂, élément pivot de la CC. Autrement dit, la CC définit NARÓD non pas par un métacollectif catégorisant, mais comme une pluralité quantifiée d'individus.

Deuxièmement, la définition du lexème NATION révèle l'importance des liens institutionnels pour le concept correspondant. Dans la définition, la composante structurelle est de loin la plus développée, puisqu'elle est reprise par deux formulations (voir Tableau 1). Du côté polonais, la

composante structurelle est absente de la paraphrase. Il est donc possible de définir NARÓD sans y recourir.

L'analyse sémantique coïncide ainsi avec des observations de sociologues. Les propos de G. Gurvitch au sujet de l'idée même de NATION en tant qu'entité collective offrent un contraste frappant avec les remarques quelque peu amères de L. Dorn¹¹ à propos de celle de NARÓD :

Mais la présence de l'état en lequel la nation tend à prendre conscience d'elle-même reste un moyen commode de repérer une nation. Peut-on d'ailleurs parler de nation là où il n'y a pas d'état ? (Frantext)

Zawsze byliśmy narodem dążącym do niepodległości, ale jednocześnie apañstwowym. Nie antypañstwowym, ale właśnie apañstwowym. Dla Polaków własne państwo zawsze było czymś, co sobie cenimy, natomiast istniejące państwo staje się dla Polaków albo problemem, albo tworem, z którym nie czują głębokiego związku.¹²

Nous avons toujours été un naród_[Instrumental sg] aspirant à l'indépendance, et en même temps a-étatique. Non pas anti-étatique, mais justement a-étatique. Pour les Polonais, l'état a toujours été quelque chose d'appréciable, mais une fois qu'il existait, il devenait pour eux soit un problème, soit une construction à laquelle il ne se sentaient pas profondément attachés.

3. Noms collectifs et leurs X

3.1 De l'intérêt de la fonction lexicale Sing & S₁

En tant que noms collectifs, nos deux lexies, tout en restant grammaticalement au singulier, dénotent plusieurs entités X. Dans le cas de NATION et NARÓD, malgré le mode différent sous lequel cette pluralité se présente (définie par un métacollectif pour le premier et par un nom pluriel quantifié pour le second), les X sont avant tout des individus. On peut donc légitimement s'attendre à ce qu'il existe un lexème dénotant l'individu X. Autrement dit, compte tenu de la structure sémantique d'un lexème collectif L₁, on peut s'attendre à la présence, dans le lexique d'une langue, d'un lexème nominal L₂, dénotant l'individu X qui est membre de la collectivité. Nous avons ainsi affaire à une relation lexico-sémantique que la LEC modélise sous forme de la fonction lexicale standard S₁.¹³

¹¹ Homme politique et député, Ludwik Dorn est sociologue de formation.

¹² En ligne : <http://www.polskatimes.pl/arttykul/267468,ludwik-dorn-jestesmy-slabym-panstwem-bez-ducha,id,t.html?cookie=1>.

¹³ Pour une présentation des fonctions lexicales, voir Mel'čuk (2007), Polguère (2007).

Cette fonction lexicale (désormais FL) est considérée comme standard, car elle existe potentiellement dans toutes les langues.

Dans le cas des noms collectifs, on peut s'attendre néanmoins à deux façons de dénoter les membres d'une collectivité. Il est possible qu'une unité lexicale L_2 singularise un seul X faisant partie de l'ensemble dénoté par L_1 (en l'occurrence, NATION et NARÓD). La lexie L_2 est alors reliée à L_1 par un double lien lexico-sémantique : elle dénote à la fois son premier actant S_1 et une « unité minimale régulière » (Mel'čuk, Clas et Polguère 1995 : 134) qui forme le *designatum* d'un nom collectif. On parle ainsi de la FL complexe **Sing & S_1** .

3.2 National, citoyen, ressortissant ?

Le français dispose du lexème nominal NATIONAL (nom masculin) et de son quasi-synonyme NATIONALE (nom féminin)¹⁴ vers lesquels pourrait pointer la FL **Sing & S_1** , notamment lorsque l'on pense à l'importance de la composante structurelle que nous avons discutée dans la section 2.2. L'exemple (1), extrait du Web, illustre ce sens :

- (1) Un national français peut ne pas être résident fiscal français. (En ligne : http://www.nodula.com/Article_du_mois185.html)

Le fait qu'il s'agisse d'un dérivé morphologique semble renforcer cette hypothèse. Ce qui nous intéresse cependant, ce n'est pas le lien morphologique, mais la relation lexico-sémantique entre NATION (L_1) et la lexie nominale NATIONAL (L_2). La question est de savoir si l'on peut dériver sémantiquement L_2 de L_1 . La réponse serait positive, si le sens de L_2 était construit sur celui de L_1 . Dans ce cas, la définition de L_2 devrait recourir au sens de L_1 : on dirait qu'un NATIONAL (L_2) est un « individu qui fait partie d'une NATION (L_1) ». Or, il n'en est rien, puisqu'un NATIONAL se définit en tant que citoyen d'un état. Autrement dit, on n'emploie pas la lexie NATION (L_1) pour expliquer le sens de NATIONAL (L_2). Celui-ci est un quasi-synonyme de CITOYEN et RESSORTISSANT, dont la définition comprend une entité sociale telle que ÉTAT ou PAYS. La seule façon de désigner l'individu qui fait partie d'une nation consiste à parler de lui ou d'elle comme d'un « membre de la nation ». Or, la lexie MEMBRE n'est pas sémantiquement dérivé de NATION : elle dénote un individu appartenant à une collectivité. Cela signifie que le français ne dispose pas de lexème dénotant spécifiquement un membre de la nation.

¹⁴ Pour une analyse détaillée et une proposition de traitement lexicographique approprié des noms d'individus, voir Delaite et Polguère (2013).

3.3 NARÓD et son dérivé sémantique **Sing & S₁**

Le lien **Sing & S₁** qui part de NARÓD nous conduit, en polonais, vers les lexèmes RODAK (nom masculin) et RODACZKA (nom féminin). Il s'agit d'unités lexicales de haute fréquence d'emploi qui s'incrivent dans un registre courant plutôt que juridique.

RODAK, nom masculin

Człowiek tej samej narodowości, co i kto drugi; krajan, ziomek, współziomek. (Sobol 2003 : 847)

Individu de la même nationalité que quelqu'un d'autre ; [Quasi-Syn]
krajan, ziomek, współziomek.¹⁵

Précisons que seul le nom masculin est défini dans les dictionnaires, alors que la description de son correspondant féminin se limite à la simple mention de son existence. Au contraire, la démarche propre au RL-fr consiste à proposer un traitement lexicographique complet des deux noms. En reformulant la définition ci-dessus conformément aux exigences du RL-fr,¹⁶ on paraphrasera ainsi RODAK et RODACZKA :

RODAK, nom masculin

Individu X, qui fait partie du même NARÓD que l'individu Y

RODACZKA, nom féminin

Individu X, de sexe féminin, qui fait partie du même NARÓD que l'individu Y

Si NARÓD apparaît bien dans les deux définitions, ce qui est frappant pour RODAK et RODACZKA, c'est le caractère relationnel de ces sens. En effet, l'individu X se définit comme RODAK ou comme RODACZKA non pas directement en rapport avec NARÓD, mais dans sa relation avec un autre individu. Observons que cet aspect relationnel place RODAK et RODACZKA dans le paradigme conceptuel dont relèvent également certains noms de relations familiales. En définissant le nom KUZYN (COUSIN), on précisera tout d'abord que l'individu X est le cousin de l'individu Y (et *vice versa*) pour indiquer ensuite leur appartenance à une même famille.

¹⁵ Les quasi-synonymes KRAJAN, ZIOMEK, WSPÓLZIOMEK ont un sens plus général (*i.e.* moins spécifique, donc moins riche) que RODAK, puisqu'ils s'appliquent également à quelqu'un qui vient d'une même région, d'une même ville, d'un même village, etc.

¹⁶ Pour une présentation de la structure des définitions dans le RL-fr, voir Sikora (2012).

En suivant le lien **Sing & S₁**, on constate une asymétrie lexicale. En français, le réseau lexical construit autour du lexème NATION n'a pas lexicalisé de nom spécifique pour l'individu X appartenant à la nation. Du côté polonais, au contraire, les noms communs RODAK et RODACZKA dénotent l'individu X à travers les relations horizontales qu'il entretient avec un autre ou avec les autres individus X. L'absence de la composante structurelle que nous avons pu constater dans la section 2 s'accompagne ainsi d'une très forte charge humaine et relationnelle.

4. NATION, NAROD et leurs combinatoires

Cette section étudie deux types de collocations que peuvent former NATION et NARÓD avec des lexies verbales. Nous nous intéressons aux syntagmes collocationnels dont ces deux lexèmes constituent la base. Faute de place, nous nous limiterons, dans la sous-section 4.1, aux collocatifs **Oper₁** et **CausFunc₀**. Nous passerons ensuite, en 4.2, à la discussion de quelques énoncés libres dans lesquels NATION et NARÓD occupent une position de sujet.

4.1 Combinatoire restreinte de NATION et NAROD

L'exploitation que nous nous proposons de faire des liens syntagmatiques **Oper₁** et **CausFunc₀** tient bien compte du fait que leurs contenus sémantiques sont tributaires des lexies qu'ils permettent de relier sur le plan syntaxique. Ainsi, sans nous attarder sur les collocatifs verbaux eux-mêmes, nous les utiliserons comme des voies de navigation : nous suivrons les liens syntagmatiques que la LEC modélise sous forme de FL **Oper₁** (section 4.1.1) et **CausFunc₀** (section 4.1.2) pour vérifier quelles sont les lexies qu'ils rattachent à NATION et à NARÓD.

4.1.1 Fonction lexicale **Oper₁** ou qu'est-ce qui appartient à NATION et NARÓD ?

Dans le cas de nos lexies collectives NATION et NARÓD, la FL **Oper₁** permet de verbaliser la relation entre X et la collectivité, i.e. construire une collocation exprimant l'appartenance de X à la collectivité. La lexie dénotant X se trouve en position de sujet, alors que NATION occupe une position de complément. Le français dispose à cet effet du verbe APPARTENIR et de la locution verbale FAIRE PARTIE illustrés dans les exemples (2) et (3) ci-dessous :

- (2) Je vote parce que j'appartiens à une Nation, une communauté de gens qui ont leur avis à donner. (En ligne : <http://www.koztousjours.fr/choisir-pour-un-chretien>)
- (3) Rahm E., ancien secrétaire général de la Maison-Blanche, fait partie de la nation américaine, tout comme Bernard K. fait partie de la nation française. (En ligne : <http://www.courrierinternational.com/article/2010/10/07/l-etat-juif-peut-nuire-a-la-democratie>)

Les liens **Oper₁** permettent de repérer une différence intéressante : en polonais, seules des lexies dénotant des êtres humains apparaissent en position de sujet d'un **Oper₁** de NARÓD, alors qu'en français, X n'est pas nécessairement un individu. Il n'est pas rare, en effet, de trouver en position de sujet un nom d'entité non animée. Observons les exemples (4) et (5) :

- (4) Cette province aurait été albanisée et islamisée à l'époque du « joug turc ». En revanche, pour les Albanais, elle fait partie de la nation albanaise et a été un des fiefs du renouveau national au XIX^e siècle. (En ligne : <https://www.monde-diplomatique.fr/1997/11/CHICLET/5005>)
- (5) La monnaie est un élément qui capte ces trois dimensions : introduite et garantie par l'État, elle est utilisée par un peuple parce qu'elle fait partie de la nation. (En ligne : http://www.memoireonline.com/09/06/209/m_analyse-heterodoxe-monnaie-euro11.html)

Des lexies dénotant des entités à la fois géographiques, économiques et/ou sociales telles que PROVINCE ou MONNAIE peuvent ainsi entretenir, avec NATION, le même lien sémantico-syntaxique **Oper₁**, que les individus X. La possibilité, pour la locution FAIRE PARTIE, de sélectionner un sujet non animé corrobore l'analyse sémantique de NATION que nous avons présentée dans la section 2.2. Rappelons qu'outre la composante géographique, le trait institutionnel est très fortement présent dans son sémantisme, ce qui explique à notre avis pourquoi, en français, un **Oper₁** peut relier à NATION le nom d'une entité géographique, économique, politique ou sociale.

4.1.2 Fonction lexicale **CausFunc₀**, son **S₁** ou qu'est-ce qui construit NATION et NARÓD ?

La fonction lexicale **Func₀** relie respectivement NATION ET NARÓD, placés en position de sujet, à un verbe dont l'apport sémantique se limite à l'expression du fait d'exister, d'avoir lieu ou encore de se produire.¹⁷ Mais ce qui nous semble particulièrement révélateur pour la conceptualisation de NATION et de NARÓD dans les deux langues, c'est le rapport de causation **Caus** qui s'ajoute à la FL **Func₀** pour créer une FL complexe **CausFunc₀** qu'illustrent les verbes CONSTRUIRE dans l'exemple (6) et TWORZYĆ (CRÉER) dans (7) :

- (6) Qu'on adhère ou pas, les Français ont construit une nation souveraine et républicaine. (En ligne : <http://www.bladi.info/threads/echanges-discussions-scientifiques.334914/page-13>)
- (7) Kultura tworzy naród.
(En ligne : https://www.usosweb.ukw.edu.pl/kontroler.php?_action=actionx:kata_log2/przedmioty/pokazPrzedmiot%28kod:0000-PF-KNI-NP%29)
La culture crée naród.

Ces deux exemples permettent de comparer le premier actant **S₁** de la FL **CausFunc₀** en français et en polonais. L'exemple (6) montre que le **S₁** mène vers un ou plusieurs individus dont la démarche volontaire a pour objectif de « faire qu'une nation existe ». Or, ce type de causation agentive (Kahane et Mel'čuk 2006; Mel'čuk 2012) est difficilement concevable en polonais. Certes, il existe des verbes susceptibles d'exprimer le lien **CausFunc₀**, tels que TWORZYĆ (CRÉER) de notre exemple (7). Cependant, en polonais, le premier actant n'est pas un agent de causation, mais plutôt un élément causateur. Parmi les **S₁** de la FL **CausFunc₀**, on trouve les lexies telles que CNOTA (VERTU), WIĘŹ (LIEN), ZWIĄZEK (RELATION) dans les syntagmes *cnoty/więzi/związki, które tworzą naród* (*vertus/liens/reliations qui créent naród*). Non seulement elles sont dépourvues de tout caractère agentif, mais elles dénotent des caractéristiques que l'on attribue aux individus X qui... forment (**Oper₁**) NARÓD.

¹⁷ Il nous est impossible, dans cet article, d'énumérer toutes les formules possibles de la relation d'existence formalisée par la FL **Func₀**. Elles dépendent de la nature du sens dénoté par la base qui sélectionne le collocatif. La glose d'un **Func₀** change donc selon qu'il s'agit d'un phénomène atmosphérique (PLUIE) qui se produit ou d'un son tel KLAXON au sens de bruit, qui est émis.

4.2 De la combinatoire libre ou de ce que peuvent faire NATION et NARÓD ?

Si la combinatoire libre d'une lexie n'est pas formalisée au moyen de FL paradigmatiques et syntagmatiques, elle n'en est pas moins soumise à un système de micro-règles sémantiquement liées aux lexies que l'on combine. Elle fournit de ce fait des indices intéressants sur le sémantisme des lexèmes impliqués. Dans cette section, nous nous intéressons uniquement aux verbes dont NATION et NARÓD sont les sujets grammaticaux. Autrement dit, nous essaierons de répondre à la question de savoir ce que font NATION et NARÓD, pour comprendre comment ils sont.

Dans les deux langues, NATION et NARÓD sont souvent personnifiés, notamment par l'emploi de verbes qui dénotent des activités réservées aux humains.

- (8) La nation assure à l'individu et à la famille les conditions nécessaires à leur développement. (Frantext)
- (9) Naród sprawuje władzę przez swoich przedstawicieli lub bezpośrednio. (Konstytucja Rzeczypospolitej Polskiej, 1997)¹⁸
Naród exerce le pouvoir par ses représentants ou directement.
(Constitution de la République de Pologne, 1997)

La personnification, on le sait, consiste à présenter l'entité dénotée comme s'il s'agissait d'un être humain. Certains noms, par exemple des noms abstraits tels que JUSTICE, FRATERNITÉ, ÉGALITÉ, etc., sont soumis à ce procédé rhétorique plus souvent que d'autres.

Or, en dehors des emplois personnifiés communs à nos deux lexèmes, NARÓD apparaît souvent, voire majoritairement, comme sujet des verbes dénotant des activités quotidiennes d'un individu. Observons les exemples suivants :

- (10) 100 lat mija od urodzin (6.02.1905 r.) Władysława Gomułki. Był to facet, który uważał, że skoro nie lubi naturalnej kawy, to i naród może pić zbożówkę. (Narodowy Korpus Języka Polskiego, désormais NKJP)
Il y a 100 ans naissait (le 6/02/1905) Władysław Gomułka. C'était un type qui considérait que puisque lui n'aimait pas le vrai café, naród, lui aussi, n'avait que boire de la chicorée.

¹⁸ En ligne : <http://www.sejm.gov.pl/prawo/konst/polski/kon1.htm>.

- (11) Nasz naród wraca z pracy i podchodzi do telewizora. (NKJP)
Notre naród rentre du travail et va voir la télé.

Tout comme *assurer à qqn les conditions nécessaires*, les syntagmes verbaux, *pić zbożówkę* (*boire de la chicorée*), *wracać z pracy* (*rentrer du travail*) prennent comme argument sujet un être humain. Cependant, la démarche que nous observons dans les exemples (10) et (11) n'est pas celle de personnification. Il s'agit non pas de faire apparaître NARÓD comme une seule personne, un être humain responsable de ses gestes et de ses comportements. On lui attribue, au contraire, des actions, des activités, des comportements ou encore des attitudes qui caractérisent la majorité, sinon la totalité de ses membres. De même, quand on choisit le lexème FAMILLE (au sens de parenté) pour dire « la famille s'est mise à table », on ne la personnifie pas, mais on signifie que chacun de ses membres y a pris place. Plutôt que d'unifier les individus pour les présenter comme un seul homme, les locuteurs semblent voir NARÓD comme une multitude d'individus s'adonnant aux mêmes activités. Les phrases (10) et (11) appellent ainsi une lecture distributive.

5. Bilan de différences entre un nom collectif abstrait et un nom collectifs concret

Avant de proposer des conclusions, nous faisons dans cette section un bilan d'étape des analyses menées, afin de vérifier si les faits de langue dégagés forment une image cohérente permettant de saisir les modes de conceptualisation qui se manifestent dans le sémantisme de NATION et de NARÓD.

En français, la polysémie de NATION est fuyante, difficile à saisir. Ses sens évoluent à travers les siècles et semblent directement liés aux formes que prend la vie politique dans un contexte historique particulier. C'est la raison pour laquelle certains dictionnaires incluent des informations encyclopédiques dans les paraphrases définitives. Par ailleurs, les dictionnaires mettent en évidence l'importance des composantes structurelles.

Dans le réseau lexical qui s'organise autour de NATION, on note l'absence de lexie dénotant spécifiquement un ou plusieurs de ses membres. Sur le plan syntagmatique, la FL **Oper₁** permet de combiner NATION avec des noms d'individus, d'entités politiques ou sociales. Le lien **S₁CausFunc₀**, qui pointe vers un agent causateur, laisse penser que les locuteurs du français voient dans NATION une entité conceptuelle qui

ressemble à un artéfact, puisqu'elle peut être créée par une personne ou par une institution.

Du côté polonais, on remarque tout d'abord une polysémie stable, faite de deux acceptions que l'on retrouve dans tous les dictionnaires consultés. Le sens politique examiné ici se caractérise par l'absence de composante structurelle : il est possible de définir NARÓD en tant que pluralité d'individus, sans référence à une entité politique telle qu'un état. En revanche, le réseau que forment autour de NARÓD les lexies reliées par les liens paradigmatiques et syntagmatiques fait apparaître l'importance de la composante humaine et des aspects relationnels. Les dérivés sémantiques **S₁** sont tous des noms relationnels : ils définissent l'individu X dans son rapport avec un autre individu appartenant à la collectivité. Les liens syntagmatiques confirment l'importance des aspects humains et relationnels : seules des lexies dénotant des individus peuvent être reliées à NARÓD par le lien syntagmatique **Oper₁**. Contrairement au français, la FL **S₁CausFunc₀** conduit non pas vers un agent susceptible de gestes volontaires et contrôlés, mais vers des noms de caractéristiques que l'on peut attribuer aux membres de la collectivité. Ce sont elles qui apparaissent dans le réseau lexical comme éléments causateurs de l'existence même de l'entité collective dénotée par NARÓD.

Les indices apportés par la combinatoire libre complètent le tableau : à la différence de NATION, personnifié de manière à unifier la collectivité, NARÓD est souvent employé comme sujet des verbes dénotant des actions ou des activités humaines sans qu'il s'agisse de personnification. Plutôt que de conduire à une lecture unifiante, une interprétation distributive est de rigueur : les faits que dénotent ces verbes sont attribués individuellement à tous ou du moins à une majorité des X. De plus, ils sont, pour la plupart d'entre eux (*boire de la chicorée, rentrer du travail, regarder la télé*), directement perceptibles dans leur réalisation.

6. Nom concret versus nom abstrait

Les faits de langue mis au jour dans les sections précédentes convergent vers une hypothèse explicative du fonctionnement linguistique si différent de NATION et NARÓD, permettant de comprendre pourquoi ces deux lexies sont des faux-amis. En effet, si les deux lexèmes sont bien des noms collectifs, NATION a tout d'un nom abstrait, alors que NARÓD présente des caractéristiques d'un nom concret. En dehors des éléments que nous avons mis en avant, cette hypothèse permet d'expliquer un certain nombre de phénomènes culturels.

Commençons, dans la sous-section 6.1 par le concret (NARÓD), pour nous tourner vers l'abstrait avec NATION en 6.2.

6.1 NAROD en tant que nom collectif concret

On s'accorde généralement pour reconnaître que l'entité dénotée par un nom concret préexiste à sa dénomination,¹⁹ souvent en raison de son caractère matériel²⁰ et accessible aux sens. Ces caractéristiques la rendent – jusqu'à un certain point – indépendante de la langue. Du point de vue de la sémantique lexicale, un nom concret possède donc une extension immédiate : pour un nom collectif comme NARÓD, ce sont les individus auxquels il s'applique.

En effet, ses membres X semblent mener une existence indépendante de la lexie qui les dénote, faite d'activités et d'actions directement perceptibles. L'appartenance à la collectivité se fait sur la base des relations qui relient un X à d'autres membres de la collectivité. De ce point de vue, la structure sémantique de NARÓD ressemble fortement à celle qu'ont, respectivement en polonais et en français, les lexèmes RODZINA et FAMILLE en tant qu'ils renvoient à un ensemble d'individus apparentés, à une parentèle. Un tel ensemble, défini par les relations entre ses membres, préexiste, lui aussi, à sa dénomination et en est indépendant.

On comprend dès lors pourquoi le titre du sondage *Czy kochasz swój naród ? (Aimes-tu ton naród ?)*, contrairement à sa traduction française, est dépourvu de charge nationaliste pour un polonophone. Si, comme nous le pensons, NARÓD est un nom collectif concret, conceptualisé – comme FAMILLE – à partir des relations entre les individus X, il présente peu d'intérêt pour un discours nationaliste,²¹ car il n'est pas discriminatoire. En effet, on peut difficilement penser que le fait de ne pas être née dans la famille Kowalski, mais d'appartenir à la famille Sikora, entraîne une discrimination ou une exclusion. De même, un individu qui n'appartient pas à un NARÓD, en a forcément un autre, le sien. De plus, on peut supposer qu'un nom collectif concret peut difficilement servir d'élément fédérateur dans des propos idéologiques, *a fortiori* lorsque son usage quotidien montre clairement que NARÓD s'avère porteur des mêmes qualités et des mêmes défauts que les individus qui le forment (il n'est pas

¹⁹ Pour une discussion de la notion de *dénomination*, voir Kleiber (1984; 2012).

²⁰ Bien que la pertinence du critère de matérialité soit fort contestable. Pour une discussion, voir, entre autres, Galmiche et Kleiber (1996), Wilmet (1996).

²¹ Une recherche rapide effectuée sur un portail d'orientation nationaliste polonais nous a permis de repérer à peine trois occurrences de NARÓD. Remarquons cependant que l'adjectif NARODOWY y est abondamment employé.

rare de trouver NARÓD qualifié par les adjectifs signifiant *bête*, *sympa*, *déconneur*, etc.).

A. Wierzbicka (communication personnelle) insiste sur l'importance de la composante affective dans la structuration du concept de NARÓD en polonais, bien plus forte que celle que l'on peut trouver en anglais et en français. Notre hypothèse en fournit une explication plausible : il semble naturel qu'un ou plusieurs individus auxquels l'on se sent lié suscitent des sentiments plus intenses – positifs ou négatifs – qu'un concept dénoté par un nom abstrait.

Remarquons qu'une telle structuration du concept lexicalisée dans la langue trouve des répercussions jusque dans l'organisation de la vie politique, puisque le système juridique polonais sépare clairement la question de la nationalité et celle de la citoyenneté.

6.2 NATION en tant que nom collectif abstrait

Face à l'extrême diversité des noms abstraits, une grande prudence s'impose pour toute généralisation. En effet, on trouve parmi eux des noms de qualités (COURAGE, GÉNÉROSITÉ), de sentiments (MÉLANCOLIE, TRISTESSE), de figures géométriques (CARRÉ, TRIANGLE), voire des noms dits syncatégorématiques tels que BLANCHEUR [de X] (Kleiber 1996), auxquels nous nous apprêtons à ajouter NATION.²² Néanmoins, tous ces noms partagent une caractéristique essentielle à nos yeux : l'entité qu'ils dénotent est ontologiquement dépendante du lexème. Pour l'identifier, il faut passer par l'intension du nom abstrait, c'est-à-dire circonscrire la combinaison d'attributs qu'on lui accorde. C'est la raison pour laquelle Goddard et Wierzbicka (2014) parlent de réification qu'opèrent les noms abstraits : ils permettent aux locuteurs d'isoler une portion de réalité et de la constituer en « quelque chose », en la nommant. Sur le plan lexical, on retrouve cette combinaison d'attributs dans l'ensemble des composantes sémantiques qui forment la définition du nom abstrait. On pourra ainsi aller jusqu'à dire que l'entité dénotée doit son existence, toute relative d'ailleurs, au système lexical d'une langue donnée, au nom qu'on définit pour elle.

Les faits de langue que nous avons dégagés permettent de constater qu'en français le lexème NATION présente toutes les caractéristiques d'un nom abstrait. Ainsi, sa polysémie quelque peu « fuyante » s'explique par les modifications d'attributs constitutifs du concept sous-jacent au lexème

²² Pour une analyse du lexème anglais NATION en tant que nom abstrait, voir Stecconi (2011).

NATION dans des conditions historiques et politiques particulières : sous l'Ancien Régime, ces attributs ne sont pas ceux que définit la Constitution de la V^e République. C'est parce que le concept évolue qu'il est tellement difficile de le circonscrire dans une définition lexicographique. La tentation est forte dès lors de décrire non pas le sens lexical, mais les formes successives du concept. Or, comme le notent Goddard et Wierzbicka (2014), le sens d'un nom abstrait comprend, certes, un invariant sémantique, mais il inclut inévitablement une part de vague.

Le français dispose de moyens lexicaux permettant d'exprimer l'appartenance à la nation, i. e. le fait, pour un individu X, de posséder les attributs requis par la définition, mais contrairement au polonais, la langue a peu, voire pas du tout, lexicalisé les relations horizontales, c'est-à-dire celles qui pourraient relier les individus X entre eux. La collectivité ne préexiste pas à sa dénomination, elle est au contraire identifiée et délimitée en fonction des attributs qui structurent le concept. Ce caractère construit par assemblage de traits constitutifs se reflète dans les liens lexicaux tissés par la lexie NATION (voir notre analyse de la FL **S₁CausFunc₀** dans la section 4.2).

Le caractère abstrait de NATION en français est une explication possible de certaines réactions négatives que nous avons perçues lors des échanges avec des locuteurs du français. Elles se manifestent souvent par une prise de distance, voire par un jugement négatif à l'égard du lexème NATION, ressenti comme idéologiquement chargé et potentiellement discriminatoire. Deux raisons motivent ces attitudes. Tout d'abord, si NATION est un nom abstrait défini en fonction des attributs qu'on lui assigne, chacun des membres de la collectivité est supposé les posséder individuellement. Cela signifie *a contrario* que ceux qui ne les vérifient pas en sont exclus. On retrouve ici l'une des caractéristiques des noms abstraits discutées dans Goddard et Wierzbicka (2014). Les auteurs soulignent qu'un nom abstrait fonctionne comme un prédicat au sens logique du terme : le locuteur dit quelque chose de l'entité qu'il désigne, en l'occurrence des individus appartenant à la nation.

Deuxièmement, les fluctuations du concept qui se manifestent dans le vague définitionnel²³ des noms abstraits laissent une certaine marge de liberté dans le choix des attributs retenus comme critères d'appartenance. Le contenu du concept est modulable, nous l'avons dit, en fonction de conditions politiques et sociales, mais il peut l'être également selon les

²³ Rappelons ce qu'est le vague d'une expression linguistique : « Une expression lexicale est vague si et seulement si son sens correspond alternativement à plus d'un référent extralinguistique, alors qu'elle-même correspond à une seule lexie » (Mel'čuk, Clas et Polguère 1995 : 60).

convictions de celui ou de celle qui l'utilise sans que toute modification soit lexicalisée. Le vague définitionnel laisse au locuteur la possibilité de manipuler les attributs que doivent avoir les individus X pour appartenir, et *a contrario* pour ne pas appartenir, à la nation. Ce qui explique à nos yeux la méfiance manifestée à l'égard de cette lexie, dont l'usage peut facilement devenir discriminatoire.

7. Définir pour conclure

Une fois que nous avons pu cerner les différences conceptuelles qui font de NATION et de NARÓD de faux-amis, il nous faudra les traduire en composantes définitionnelles pour construire des définitions lexicographiques. Le tableau 2 ci-dessous permet de comparer les paraphrases que nous proposons pour les deux lexèmes.

| lexème | NATION, nom féminin | NARÓD I, nom masculin |
|--------------------------------|---|--|
| composante centrale (CC) | Ensemble d'individus | Individus |
| composantes périphériques (CP) | qui possèdent des attributs définis et requis par une entité sociale. | qui se sentent liés entre eux par des relations linguistiques et culturelles et_ou géographiques |

Tableau 2. Paraphrases définitives de NATION et de NARÓD I

On remarquera que seule la lexie NARÓD est numérotée : en effet, il convient de la distinguer de son copolysème NARÓD II ('grand nombre d'individus', 'foule', voir section 2.1). En français, au contraire, le vocable NATION est monosémique : il compte une seule acception qui doit être suffisamment vague pour inclure toutes les formes que peut prendre une nation.

Les CC des deux lexèmes se différencient par la présence du terme métacollectif ENSEMBLE II que l'on trouve du côté français. Or, on sait qu'un tel regroupement se fait sur la base de propriétés communes à tous les X. Du côté polonais, la CC marque seulement la pluralité du *denotatum*. L'adjonction d'une expression quantifiante telle que 'grand nombre de X', 'plusieurs X') devrait être envisagée, mais à ce stade, nous avons voulu cette CC aussi simple que possible.

Voyons à présent les composantes périphériques. Concernant NATION, la CP 'qui possèdent des attributs définis et requis par une entité sociale' est destinée à rendre compte du caractère abstrait et construit de ce

sens, tout en gardant la part du vague en ce qui concerne les attributs assignés aux individus X. La formule ‘définis et requis par une entité sociale’ prend en charge la composante structurelle. Une fois de plus, nous tenons à rester vague en ce qui concerne les aspects institutionnels pour que notre définition s’applique non seulement aux emplois repérés dans la Constitution de la V^e République, mais également à ceux que l’on peut faire de ce lexème en rapport avec d’autres réalités politiques et sociales. Le vague structurel marque ainsi la relation de dépendance entre une entité sociale, qui peut être aussi bien un état qu’un parti politique, et les attributs sélectionnés pour appartenir à la nation. Enfin, la définition que nous proposons fait l’économie de la composante géographique, car l’attachement d’une nation à un territoire peut soit faire partie d’attributs choisis (l’appartenance de X à une nation tient alors au fait d’habiter un territoire), soit être lié à l’entité sociale et/ou politique, qui s’étend généralement dans un périmètre bien précis.

Du côté polonais, la CP a une construction disjonctive pour indiquer le caractère non nécessaire des liens d’un type particulier. En effet, il se peut que le sentiment d’appartenance ne tienne qu’à des relations culturelles ou familiales, mais pas géographiques, notamment pour les membres de la diaspora. Cette formulation de la CP doit rendre compte de la structuration horizontale des relations entre les membres de la collectivité.

Il va de soi qu’aussi bien pour NATION que pour NARÓD, il ne s’agit que de premières ébauches, destinées à traduire en composantes définitionnelles les différences conceptuelles que nous avons cherché à cerner, et à expliciter le faux air de famille entre ces deux lexies si proches en apparence et si éloignées par le sens qu’elles lexicalisent.

Remerciements

Je tiens à remercier Cliff Goddard et Anna Wierzbicka de m’avoir communiqué leurs analyses à paraître. Je suis reconnaissante à Veronika Lux-Pogodalla, à Sandrine Ollinger, à Alain Polguère et à Roger Pouivet pour toutes les discussions sur le fonctionnement de NATION en français. Je remercie chaleureusement Paolo Frassi pour sa lecture attentive et pour ses commentaires. Toutes les incompréhensions sont entièrement de mon fait.

Références

- Assemblée Nationale. (En ligne : <http://assemblee-nationale.fr>)
- Centre National des Ressources Textuelles et Lexicales. (En ligne : <http://www.cnrtl.fr>)
- Delaite, C. et A. Polguère (2013). « Sex-Based Nominal Pairs in the French Lexical Network : It's Not What You Think ». In V. Apresjan, B. Iomdin et E. Ageeva (éds.), *Proceedings of the Sixth International Conference on Meaning-Text Theory (MTT'13)*, 29-40.
- Elnitsky, L. et I. A. Mel'čuk (1984). « Toward the Lexicographic Description of the Cooccurrence of « Parametric » Lexemes in French and Russian ». *Linguisticae Investigationes*, 8, 2, 269-284.
- Galmiche, M. et G. Kleiber (1996). « Sur les noms abstraits ». In N. Flaux, M. Glatigny et D. Samain (éds.), *Les noms abstraits. Histoire et théories*. Villeneuve-d'Ascq : Presses universitaires du Septentrion, 23-40.
- Goddard, C. et A. Wierzbicka (2014). *Words and Meanings : Lexical Semantics Across Domains, Languages, and Cultures*. Oxford : Oxford University Press.
- Iordanskaja, L. et I. Mel'čuk (1997). « Le corps humain en russe et en français. Vers un Dictionnaire explicatif et combinatoire bilingue ». *Cahiers de lexicologie*, 70, 1, 103-135.
- Kahane, S. et I. A. Mel'čuk (2006). « Les sémantèmes de causation en français ». In S. Hamon et M. Amy (éds.), *Linx*, 54, 247-292.
- Kleiber, G. (1984). « Dénomination et relations dénominatives ». *Langages*, 76, 77-94.
- . (2012). « De la dénomination à la désignation : le paradoxe ontologico-dénominateur des odeurs ». *Langue française*, 174, 45-58.
- Lammert, M. (2008). « Ensemble, groupe et collection : des noms collectifs bien particuliers ». *Scolia*, 23, 85-107
- Lux-Pogodalla, V. et A. Polguère (2011). « Construction of a French Lexical Network. Methodological Issues ». *Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011*. An ESSLLI 2011 Workshop, Ljubljana, 54-61. (En ligne : http://alpage.inria.fr/~sagot/woler2011/WoLeR2011/Program_%26_Proceedings.html)
- Mel'čuk, I. (1993). *Cours de morphologie générale (théorique et descriptive)*. Vol. 1. *Introduction et Première partie : Le mot*. Montréal – Paris : Les Presses de l'Université de Montréal – CNRS.
- . (2007). « Lexical Functions ». In H. Burger, D. Dobrovol'skij, P. Kühn et N. Norrick (éds.), *Phraseology. An International Handbook*

- of Contemporary Research*. Berlin/New York : W. de Gruyter, 119-131.
- (2012). *Semantics. From meaning to text*. Volume 1. Amsterdam, Philadelphia : John Benjamins Publishing Company.
- (2013). « Tout ce que nous voulions savoir sur les phrasèmes, mais... ». *Cahiers de lexicologie*, 1, 129-149.
- Mel'čuk, I., A. Clas et A. Polguère (1995). *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve : Duculot.
- Polguère, A. (2007). « Lexical Functions Standardness ». In L. Wanner (éd.), *Festschrift in Honour of Igor Mel'čuk*, Amsterdam-Philadelphia : John Benjamins, 43-95.
- Sejm Rzeczypospolitej Polskiej. (En ligne : <http://www.sejm.gov.pl>)
- Sikora, D. (2012). « Définir le sens dans un réseau lexical ». *Studia Romanica Posnaniensia*, 39, 3, 63-79.
- Sobol, E. (éd.) (2003). *Nowy słownik języka polskiego*. Warszawa : Państwowe Wydawnictwo Naukowe.
- Stecconi, S. (à paraître/2011). « Per un'analisi di concetti chiave dell'ambito politico secondo il Natural Semantic Metalanguage : un confronto italiano inglese ». MA thesis. Catholic University of Milan.
- Wierzbicka, A. (1999). *Semantics, Culture and Cognition. Universal Human Concepts in Culture-Specific Configurations*. Oxford : Oxford University Press.
- (2011) « What's wrong with « happiness studies ». The cultural semantics of happiness, bonheur, Glück and scast'e ». *Slovo i Jazyk*, 155-171.
- Wilmet, M. (1996). « À la recherche du *nom abstrait* ». In N. Flaux, M. Glatigny et D. Samain (éds.), *Les noms abstraits. Histoire et théories*. Villeneuve-d'Ascq : Presses universitaires du Septentrion, 67-76.

Peut-on formaliser l'excentricité ? Quand une description sémantique s'organise en réseau

1. Introduction

En ouverture de son étude consacrée à l'excentricité dans le Paris du XIX^e siècle, Gill lance comme un avertissement :

Contrary to what is often assumed, the meaning of eccentricity is neither timeless nor self-evident. The very synonyms used to define it—bizarre, singular, original, peculiar, odd—suggest the frustration of rationality and the failure of the codes by which social and mental life is interpreted. Indeed, the concept owed its success precisely to its ability to adapt to new contexts.¹

L'objectif de cet article n'est pas de réfléchir sur d'éventuelles reconceptualisations de la notion même d'excentricité : depuis son entrée fracassante dans le lexique français au début du XIX^e siècle, son, ou plutôt ses sens restent étonnamment stables, même si les normes sociales et culturelles qu'un excentrique est censé transgresser ont été considérablement modifiées. Une analyse linguistique va ainsi à l'encontre du constat de Gill : ce qui change, ce n'est pas le sens lexical, c'est la perception et les réactions face à l'excentricité et à l'excentrique.

La présente étude propose une analyse purement linguistique du vocable EXCENTRICITÉ dans une perspective lexicographique en synchronie, c'est-à-dire dans l'état actuel de l'évolution du français. Notre objectif est de convaincre le lecteur qu'une description formelle de ce terme est non seulement applicable à EXCENTRICITÉ dans toutes ses acceptions, mais indispensable : loin de gommer les différences, elle saisit au contraire ce que chacune d'elles a de particulier, et ce qui la distingue des unités lexicales sémantiquement proches.

Les conventions d'écriture sont un premier élément d'une description formelle ; elles permettent de distinguer avec précision l'objet d'étude. Dans nos analyses, qui suivent les principes de la Lexicologie Explicative et Combinatoire², nous utiliserons les petites majuscules pour transcrire les unités lexicales (ou lexies). Une unité lexicale est un signe linguistique qui réunit un signifiant (une forme phonétique ou écrite), un signifié (un sens) et des caractéristiques de combinatoire syntaxiques. Ce ne sont donc pas des 'mots', terme dont l'imprécision a maintes fois été démontrée, qui forment le lexique d'une langue, mais des unités lexicales. En parlant de vocable, nous renvoyons à l'ensemble d'unités lexicales reliées par des relations de polysémie.

Un vocable peut être, et le plus souvent est, polysémique : il réunit alors plusieurs unités lexicales – ses acceptions. C'est le cas du vocable EXCENTRICITÉ, qui regroupe toute un ensemble de lexies. L'objectif de cette présentation peut désormais être exposé d'une manière

1 Miranda Gill, *Eccentricity and the Cultural Imagination in the Nineteenth-Century Paris*, Oxford, Oxford University Press, 2009, p.1.

2 Cf. Igor Mel'čuk, André Clas, Alain Polguère, *Introduction à la Lexicologie Explicative et Combinatoire*, Louvain-la-Neuve, Duculot, 2009.

plus précise : il s'agit de saisir ce qu'une description linguistique établie selon un certain nombre de principes formelles révèle de la place et du fonctionnement des différentes lexies que recouvre le signifiant EXCENTRICITÉ.

Nous procéderons en trois étapes. La section 2 présente le cadre dans lequel cette recherche a été menée. Nous discutons, dans la section 3, deux hypothèses relatives à la polysémie du (des ?) vocable(s) EXCENTRICITÉ. L'étape suivante de notre réflexion s'organisera autour du sens 'manière d'être' ; elle nous permettra en effet de montrer les spécificités de ces lexies et nous amènera vers quelques conclusions réunies dans la cinquième et dernière section.

2. Réseau Lexical du Français : EXCENTRICITÉ dans une ressource dictionnaire

2.1 Présentation du Réseau Lexical du Français

Nos analyses du vocable EXCENTRICITÉ ont été réalisées dans le cadre du projet lexicographique 'Relief' développé depuis 2011 à l'ATILF – CNRS (UMR 7118)³ sous la direction d'Alain Polguère. L'objectif en est de construire le Réseau Lexical du Français (désormais RL-fr), base de données lexicales à large couverture pour le français dans un premier temps, pour d'autres langues par la suite⁴. Destiné à prendre le relai du Trésor de la Langue Française⁵, LA référence parmi les dictionnaires du français, le RL-fr, est pourtant très différent de son prédécesseur. Il appartient à la famille de ressources de type dictionnaires virtuels⁶. Disponible uniquement au format électronique, cette base de données regroupe toutes les informations lexicales indispensables pour générer automatiquement des dictionnaires du français général ou de spécialités, quelles qu'elles soient, allant potentiellement de la gastronomie à la physique quantique.

Au-delà de ces applications pratiques, le RL-fr est l'implémentation d'un modèle du lexique. Depuis les travaux de Kiss⁷, de nombreux chercheurs⁸ ont mis en évidence son caractère relationnel : les unités lexicales forment non pas une liste alphabétique, comme pourraient le laisser penser les dictionnaires traditionnels, mais un immense réseau. Selon Aitchison, "[...] we should imagine them as linked together in a gigantic multidimensional cobweb, in which every item is attached to scores of others."⁹

Le RL-fr fait ainsi partie des ressources dictionnaires dont l'objectif n'est pas seulement de décrire un nombre élevé de lexies d'une langue. Elles sont construites de manière à rendre compte de la structure en réseau du lexique. L'ensemble d'informations encodées dans le RL-fr peut être visualisé sous forme d'un immense graphe, une sorte de photo d'identité du lexique français reproduite ci-dessous dans la Figure 1 :

3 www.atilf.fr

4 Des travaux sur des ressources similaires, soumises aux mêmes principes théoriques et méthodologiques, ont été entrepris pour l'anglais, l'arabe, le coréen, l'espagnol et le russe.

5 Cf. <http://atilf.atilf.fr/tlf.htm>

6 Cf. Thierry Selva, Serge Verlinde, Jean Binon, « Vers une deuxième génération de dictionnaires électroniques », in *Traitement Automatique des Langues*, 44 (2), 2003, pp. 77–197.

7 George R. Kiss, « Words, associations, and networks », in *Journal of Verbal Learning and Verbal Behavior*, 7, 1968, pp. 707-713.

8 Cf. entre autres, Jean Aitchison, *Words in the Mind: An Introduction to the Mental Lexicon*, Oxford, Blackwell Publishing, [1987]2003, Christiane Fellbaum (ed.) *WordNet: An electronic lexical database*, Cambridge (Mass.), MIT Press, 1998, Mariano Sigman, Guillermo Cecchi, « Global organization of the wordnet lexicon », in: *Proceedings of the National Academy of Sciences*, 99(3), 2002, pp. 1742–1747.

9 Jean Aitchison, 2003, p. 72.

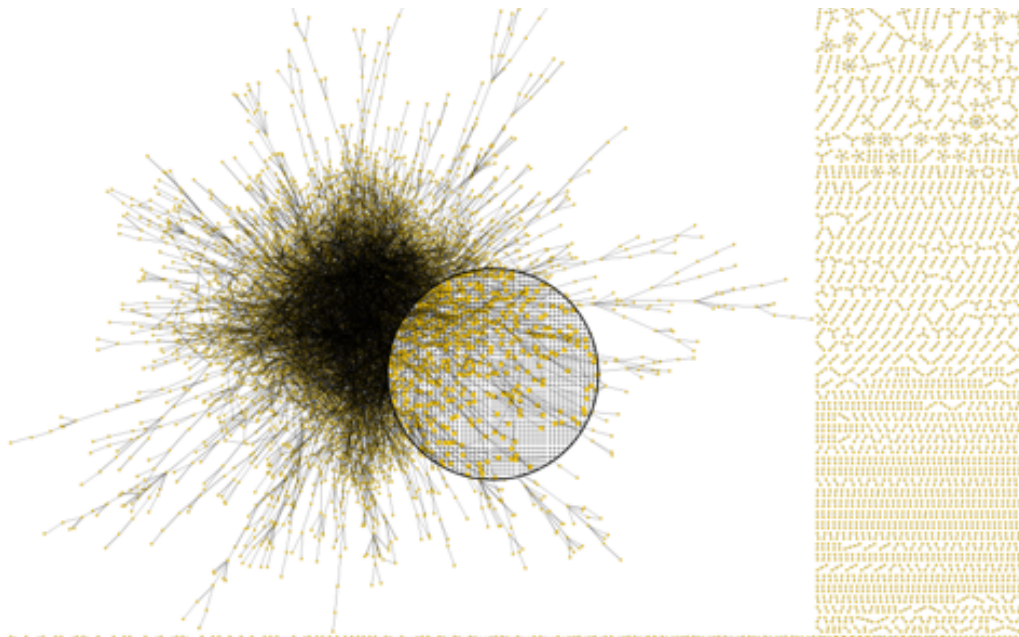


Figure 1 : Graphe lexical *Réseau Lexical du Français* (juin 2015)

Le graphe RL-fr est constitué de nœuds – environ 27 400 unités lexicales à l’heure actuelle¹⁰ – reliés par des arcs qui représentent des relations lexicales de nature diverse et variée : liens paradigmatisés (c’est-à-dire ceux qui se tissent à travers le sens : polysémie, quasi-synonymie, antonymie, généricité, etc.) et syntagmatiques qui se manifestent lorsque la lexie doit se combiner avec d’autres dans la phrase. Par exemple, pour parler de EXCENTRICITÉ I.1b, celle qui caractérise une orbite, le locuteur emploie les adjectifs FORT II.2 ou FAIBLE II.a, plutôt que FOU, réservé à EXCENTRICITÉ III.1a ‘manière d’être’.

Le RL-fr et les principes de sa structuration ont été exposés et discutés dans de nombreuses publications ; nous nous limitons donc ici à une présentation très rapide¹¹. Avant de quitter cette section, il nous faut cependant insister sur ce qui distingue le RL-fr des autres modèles relationnels du lexique, notamment ceux de type WordNet¹² : il s’agit de système lexical¹³, modèle non-ontologique, qui sans privilégier les relations entre concepts, intègre tous (ou presque) les liens – quelle que soit leur nature – qui se tissent entre les unités lexicales.

2.2 Modélisation de la polysémie dans le Réseau Lexical du Français

Les nœuds lexicaux du graphe RL-fr sont – nous l’avons dit – des unités lexicales. Pour situer EXCENTRICITÉ dans le lexique, le premier pas consiste donc à déterminer combien de sens différents s’associent avec ce signifiant. Dégager la polysémie d’un vocable revient donc à créer un premier mini-réseau, celui que constituent ses différentes acceptions¹⁴ en tant que nœuds du graphe lexical.

10 Novembre 2016. Notons que compte tenu de sa croissance, aussi bien pour ce qui est du nombre de nœuds lexicaux que de la densité des liens, une visualisation devient difficilement lisible. C’est la raison pour laquelle nous proposons ci-dessus le graphe tel qu’il se présentait en 2015.

11 On pourra se reporter par exemple à Veronika Lux-Pogodalla, Alain Polguère, « Construction of a French Lexical Network. Methodological Issues », in *Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011. An ESSLLI 2011 Workshop*, Ljubljana, 2011, pp. 54–61, Alain Polguère, Dorota Sikora, « Modèle lexicographique de croissance du vocabulaire fondé sur un processus aléatoire, mais systématique », in Caroline Masseron, Christiane Garcia-Debanç, Christophe Ronveaux (éds) *Enseigner le lexique. Pratiques sociales, objets à enseigner et pratiques d’enseignement*, Namur, Presses universitaires de Namur, 2013, pp. 35-61, Alain Polguère, « From Writing Dictionaries to Weaving Lexical Networks », in: *International Journal of Lexicography*, 27/4, 2002, pp. 396-418.

12 Ressource développée, pour l’anglais, à l’Université de Princeton (<http://wordnet.princeton.edu/>) depuis les années 1990 par l’équipe de George A. Miller et Christiane Fellbaum et pour des centaines d’autres langues, réunies au sein de Global Wordnet Association (<http://globalwordnet.org/wordnets-in-the-world/>).

13 Cf. Alain Polguère, 2002.

14 Notons par souci de précision que chacune de ces acceptions est une unité lexicale.

Si la polysémie est un phénomène omniprésent dans la langue, il existe plusieurs façons de la décrire. Dans la mesure où le RL-fr modélise le lexique du français contemporain tel qu'il se présente pour un locuteur non spécialiste, les aléas étymologiques et les évolutions du sens en diachronie ne sont pas pris en compte. On admet au contraire que toute polysémie dans l'état actuel de la langue s'organise à partir d'une unité lexicale de base (désormais ULB), permettant de tisser des liens sémantiques entre les acceptions, c'est-à-dire d'identifier les composantes sémantiques communes.

Notons qu'une telle approche conduit parfois à des résultats qui se trouvent à l'opposé d'une description historique. Il se peut qu'en français contemporain, un sens historiquement premier ait perdu sa position centrale d'ULB au profit d'une acception plus tardive. C'est le cas de VOITURE : pour un locuteur contemporain, l'ULB n'est plus 'véhicule monté sur roues, tiré ou poussé par un animal, un homme', mais bien 'véhicule automobile'.

Le cas de BLAGUE et de ses deux sens 'petit sac à tabac' et 'histoire drôle' est encore plus frappant : si l'on adopte une perspective diachronique, on sera amené à conclure à la polysémie, puisque ces deux acceptions ont évolué à partir d'une même source empruntée au néerlandais. En synchronie cependant, les éléments sémantiques communs ne sont plus perceptibles pour le locuteur, qui y verra dès lors un cas d'homonymie.

3. EXENTRICITÉ : polysémie ou homonymie ?

L'analyse des sens associés au signifiant EXCENTRICITÉ en français contemporain a eu pour objectif premier de dégager leur nombre pour, dans un second temps, identifier l'ULB, c'est-à-dire l'acception à partir de laquelle s'organise cette polysémie.

3.1 Polysémie du vocable excentricité : une première analyse

Une première hypothèse a ainsi consisté à organiser huit unités lexicales EXCENTRICITÉ à partir du sens astronomique. Attesté pour la première fois au XVI^e siècle¹⁵, il est plus « jeune » que son copolysème géométrique, mais il est incontestablement le plus connu de toutes les acceptions terminologiques de ce vocable. Il pourrait de ce fait être considéré comme source de métaphore pour EXCENTRICITÉ 'manière d'être'.

La figure 2 ci-dessous visualise la polysémie d'EXCENTRICITÉ sous forme d'un mini-graphe lexical. Nous avons ensuite reproduit manuellement ce graphe en figure 3, en y introduisant des exemples qui illustrent chaque acception.

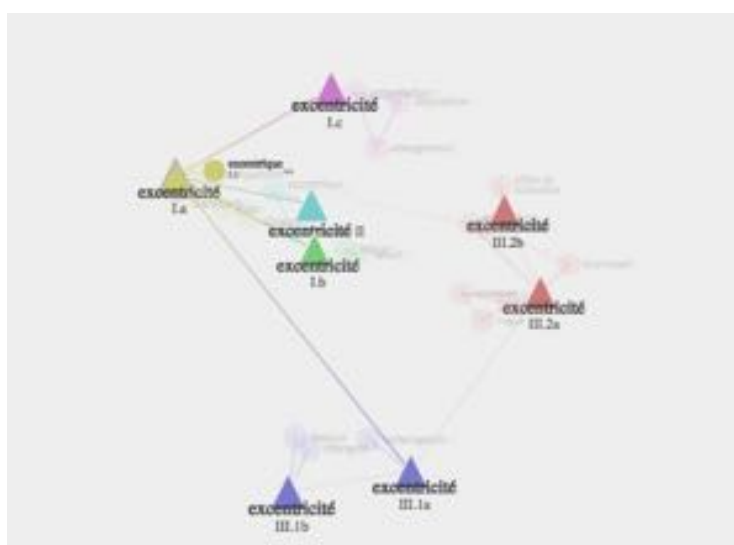


Figure 2 : Graphe de polysémie du vocable EXCENTRICITÉ issu d'une première analyse.¹⁶

15 Selon le Trésor de la Langue Française, cette attestation daterait de 1562.

16 Visualisation réalisée via www.spiderlex.net.

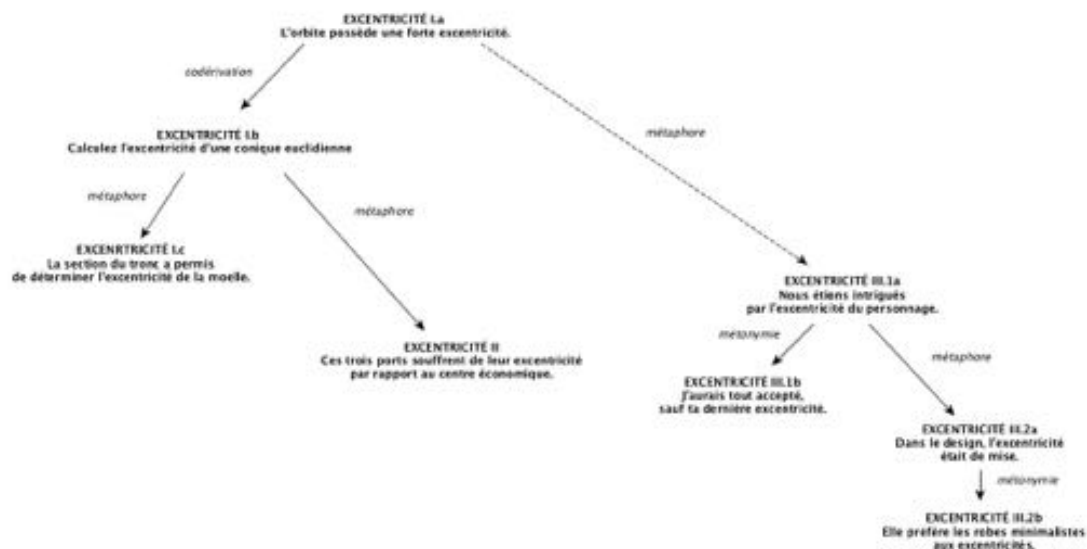


Figure 3 : Graphe de polysémie du vocable EXCENTRICITÉ issu d'une première analyse avec des exemples.

Le graphe de la figure 2, généré automatiquement à partir des données encodées dans le RL-fr révèle une topologie plutôt curieuse (nous nous sommes efforcées de la préserver dans la figure 3). En effet, on y aperçoit deux familles de sens relativement unies, mais éloignées l'une de l'autre. Les acceptions terminologiques (en géométrie, en astrologie, en botanique) ne sont pas décrites exhaustivement¹⁷ dans l'état actuel du développement de la ressource, mais elles restent liées au sens métaphorique EXCENTRICITÉ II, qui semble sémantiquement construit sur celui qu'on utilise en astronomie. Celui-ci est également relié à EXCENTRICITÉ III.1a, 'manière d'être', une autre métaphore potentielle qui se trouve au cœur d'une autre famille de sens.

Dans la mesure où la longueur des arcs est fonction de la distance sémantique (proxémie), la relation de sens entre EXCENTRICITÉ I.a et EXCENTRICITÉ III.1a, qui se voit en même temps assurer le lien entre les deux familles de sens copolysèmes, doit être attentivement observée. La question qui se pose est celle de la validité de ce lien métaphorique, marqué en pointillés dans la figure 3. Peut-on considérer que l'acception astronomique EXCENTRICITÉ I.a est réellement source de métaphore d'EXCENTRICITÉ III.1a 'manière d'être' ? Selon la réponse apportée à cette question, soit il nous sera possible de confirmer l'hypothèse de la polysémie, soit nous serons amenés à conclure à l'homonymie, qui se résume à une identité de forme dont les différentes acceptions n'ont aucun lien de sens¹⁸. Si cette seconde hypothèse venait à se confirmer, il nous faudrait scinder le graphe représenté dans les figures 2 et 3 en deux réseaux correspondant respectivement aux vocables EXCENTRICITÉ¹ et EXCENTRICITÉ², dont chacun aura développé sa propre polysémie.

3.2 Hypothèse de polysémie en question(s)

Afin de vérifier si EXCENTRICITÉ III.1a est bien une extension métaphorique du sens astronomique, nous avons adopté d'abord adopté une démarche synchronique (3.2.1), pour examiner ensuite des données en corpus en quête de renseignements diachroniques (3.2.2).

3.2.1 Vérification en synchronie

¹⁷ Notons que contrairement aux dictionnaires traditionnels, l'architecture des données dans le RL-fr permet leur réajustement à tout moment. Ainsi, d'autres acceptions terminologiques (par exemple celle employée en artillerie) pourront être intégrées dans le réseau.

¹⁸ L'absence de lien sémantique entre deux acceptions en cas d'homonymie a des causes diverses. AVOCAT¹ ('juriste') et AVOCAT² ('fruit') sont homonymes en raison d'étymologies différentes. BLAGUE¹ ('petit sac à tabac') et BLAGUE² ('histoire drôle') ont au contraire une racine étymologique commune, mais le lien de sens n'est plus perceptible pour un locuteur contemporain du français, qui y verra par conséquent un cas d'homonymie.

Peut-on admettre que le locuteur contemporain perçoit un lien de métaphore dont le sens astronomique EXCENTRICITÉ I.a serait la source pour EXCENTRICITÉ III.1a ? En quête d'indices, nous avons proposé à 67 personnes le questionnaire reproduit ci-dessous dans la figure 4.

Vous connaissez certainement le vocable EXCENTRICITÉ, qui a plusieurs sens. Lisez les définitions dans le tableau et choisissez celles qui correspondent à EXCENTRICITÉ. En utilisant les chiffres 1 (sens de base), 2 (sens proche du sens de base) et 3 (sens le plus figuré), indiquez vos préférences.
Pour les sens qui ne correspondent pas à EXCENTRICITÉ, inscrivez simplement NON dans la cas de droite.

Est-ce que, selon vous, EXCENTRICITÉ veut dire

| Définition | |
|--|--|
| 'Manière d'être qui s'éloigne des habitudes reçues' | |
| 'Action inhabituelle et bousculant les habitudes.' | |
| 'Écart de forme qui caractérise l'orbite d'un corps céleste par rapport à un cercle parfait' | |
| 'Déformation d'un pneu causée par un choc' | |
| 'Force qui repousse les corps en direction opposée au centre de la Terre' | |
| 'Vêtement inhabituel et étonnant dans une situation donnée' | |

Figure 4 : Questionnaire destiné à vérifier l'organisation de la polysémie du vocable EXCENTRICITÉ

Parmi les définitions relativement rudimentaires que nous avons soumises à un public étudiant, quatre correspondent aux acceptions existantes. Si nous y avons entremêlé des paraphrases de sens inexistantes, c'est dans le but d'éviter des mises en relation coûte que coûte de tous les éléments proposés. Nous espérons ainsi encourager les personnes interrogées à s'en remettre entièrement à leur compétence de locuteur et à leur intuition linguistique.

Parmi les 67 réponses obtenues, six seulement (soit environ 9 %) positionnent EXCENTRICITÉ I.a comme ULB (sens de base) et de ce fait, source de la métaphore qui nous intéresse. Pour 26 informateurs (39 %), ce sens de terminologie astronomique est connu, mais non central. 35 d'entre eux (52 %) ne le connaissent pas du tout. Ces résultats nous amènent à conclure que pour une majorité écrasante de personnes interrogées (91 % = 39 % + 52 %), le lien de métaphore dont l'acception EXCENTRICITÉ I.a serait la source pour EXCENTRICITÉ III.1a n'est pas établi. A contrario, 52 réponses (soit 77,6 % de nos informateurs) placent cette dernière en position d'ULB, en reconnaissant ainsi son rôle central dans l'organisation de la polysémie du vocable.

Cette première investigation montre clairement l'extrême faiblesse de ce lien de métaphore dans la polysémie du vocable EXCENTRICITÉ et ne permet pas de le valider.

3.2.2 Vérification en diachronie

Le modèle descriptif adopté dans le RL-fr ne prend pas en compte les aspects historiques relatifs à l'évolution diachronique de la polysémie d'un vocable, c'est-à-dire notamment à l'ordre dans lequel ses différentes acceptions ont vu le jour. Cependant, dans le cas d'EXCENTRICITÉ, une entorse aux préceptes méthodologiques du RL-fr nous a paru indispensable, car elle pouvait apporter des informations susceptibles d'expliquer la fragilité du lien de sens entre EXCENTRICITÉ I.a et EXCENTRICITÉ III.1a.

Dans cette optique, nous avons exploré la base textuelle Frantext¹⁹, en effectuant un retour vers le passé. Nous avons sélectionné des sous-corpus de travail par tranches de 50 ans, en commençant à partir de 1550, puisque la première attestation d'EXCENTRICITÉ semble dater de 1562. Le tableau 1 ci-dessous présente les résultats obtenus : nos requêtes successives n'ont ramené aucune occurrence dans les sous-corpus antérieurs à 1750²⁰.

19 Frantext (www.frantext.fr) réunit 4746 textes français allant du X^e au XXI^e, qui constituent un corpus d'environ 286 millions de mots. Il est à noter que la base est mise à jour à intervalles réguliers : ces informations correspondent à la période (décembre 2016) de notre consultation pour les besoins de la présente étude.

| sous-corpus | taille du sous-corpus | nombre d'occurrences |
|-------------|-------------------------------|---|
| 1750-1799 | 349 textes 13 950 560 mots | 55 (sens astronomique) |
| 1800-1849 | 468 textes 24 786 653 mots | 34 (dont 30 de la famille EXCENTRICITÉ III) |
| 1850-1899 | 561 textes 31 309 426 mots | 95 |
| 1900-1949 | 940 textes 42 789 464 mots | 160 |
| 1950-1999 | 903 textes 61 222 159 mots | 97 |
| 2000-2016 | 188 textes 14 334 553 mots | 32 |

Tableau 1 : Attestations de la forme EXCENTRICITÉ par sous-corpus de la base textuelle Frantext.

Après une première période (1750-1799) où seule l'acception EXCENTRICITÉ I.a est attestée, essentiellement dans des traités scientifiques, le sens 'manière d'être', considéré aujourd'hui comme central, fait son entrée triomphale dans le lexique français. Le premier texte qui témoigne de sa présence est sorti de sous la plume de Madame de Staël, qui nous livre ainsi une définition d'EXCENTRICITÉ III.1a :

Chez un peuple où tout est prononcé, comme en Angleterre, les contrastes sont d'autant plus frappants. La mode a un singulier empire sur les habitudes de la vie, et cependant il n'est point de nation où l'on trouve autant d'exemples de ce qu'on appelle l'excentricité, c'est-à-dire une manière d'être tout à fait originale, et qui ne compte pour rien l'opinion d'autrui.²¹

Notons cependant que parmi les 34 occurrences repérées entre 1800 et 1849, figurent également les autres acceptions de la famille polysémique référencée par le chiffre romain III. Il est frappant de constater qu'elles apparaissent systématiquement dans des contextes où il est question d'attitudes et comportements observés chez certains Britanniques. Il s'agit en effet d'une période où – comme l'observe Gill – un vif intérêt pour le mode de vie anglais gagne la France (ou du moins ses milieux artistiques) : « After Napoleon's defeat in 1815 (...) French writers began to appropriate the concept of eccentricity together with a range of English artefacts and myths, as a wave of Anglomania swept through Paris. »²²

Les résultats de nos recherches sur corpus et les conclusions d'historienne des cultures convergent ainsi, en précisant le cheminement sémantique du signifiant EXCENTRICITÉ : tout porte à croire que l'acception III.1a 'manière d'être', centrale pour la famille III, est entrée dans le lexique français au début du XIX^e siècle en tant qu'emprunt d'anglais, sur la vague d'anglophilie mentionnée par Gill.

Les emprunts relèvent d'un procédé d'enrichissement lexical on ne peut plus courant : TOUNDRA, TSUNAMI, KRACH, CHALLENGE et bien d'autres ont été intégrés en français en provenance respectivement du russe, du japonais, de l'allemand et de l'anglais. Les deux premiers sont des xénismes : emprunts qui ont permis de nommer des phénomènes inconnus dans les territoires francophones et pour lesquels, par conséquent, le français ne disposait pas de dénomination.

L'unité lexicale EXCENTRICITÉ III.1a serait ainsi un xénisme bien particulier, car limité à la seule dimension sémantique (i. e. au signifié). En effet, il existait, bien avant le XIX^e siècle,

20 L'absence d'attestation en corpus n'exclut pas cependant des usages antérieurs : le Dictionnaire du Moyen Français permet d'antédater les emplois astronomiques : *eccentricité* apparaît en effet au XIV^e siècle. Merci à Béatrice Stumpf pour cette indication.

21 Frantext, Germaine de Staël, *Considérations sur les principaux évènements de la Révolution française*, 1817, p. 371.

22 Miranda Gill, 2009, p.2.

un même signifiant associé notamment aux sens astronomique et géométrique (EXCENTRICITÉ I.a et EXCENTRICITÉ I.b). Or, selon notre hypothèse, ce signifiant s'est vu doté du signifié 'manière d'être' emprunté à l'anglais, afin de nommer une réalité britannique, avant que celle-ci ne soit à son tour rapatriée (ou tout simplement repérée) en France : « Though eccentricity was identified with Englishness during the eighteenth century, by both continental observers and English themselves, it had been 'frenchified' by the *fin de siècle* »²³. Ainsi, nos investigations en diachronie expliquent la faiblesse, voire l'absence de lien de métaphore entre les unités lexicales EXCENTRICITÉ I.a et EXCENTRICITÉ III.1a : cette dernière n'est pas le résultat d'une extension de sens qui se serait produite en français. Et si certains de nos informateurs ont pu l'établir, on peut légitimement s'interroger, si cette relation sémantique n'est pas le résultat d'une tendance à l'hypercorrection s'appuyant sur des connaissances encyclopédiques acquises tout au long du parcours scolaire, plutôt que sur les compétences purement lexicales de locuteurs natifs.

3.3 Homonymie. What else?

Les vérifications en synchronie et en diachronie apportent ainsi des arguments forts en faveur de la seconde hypothèse formulée à la fin de la section 2.2 : en l'absence de lien de sens susceptible de relier la famille d'acceptions terminologiques I et celles regroupées en III, force est de conclure à l'homonymie. Le graphe présenté dans les figures 3 et 4 doit donc être réorganisé, divisé en deux réseaux. Le premier possède quatre nœuds lexicaux, avec l'acception 'manière d'être' en position d'ULB. Notons que compte tenu de la redistribution des sommets lexicaux, les distances sémantiques doivent, elles aussi, être reconsidérées : raison pour laquelle nous avons renuméroté les unités lexicales qui forment les deux vocables EXCENTRICITÉ¹ et EXCENTRICITÉ². Leurs réseaux de polysémie sont présentés ci-dessous dans les figures 5 et 6.



Figure 5 : Polysémie d'EXCENTRICITÉ¹

²³ Ibidem, pp. 2-3.

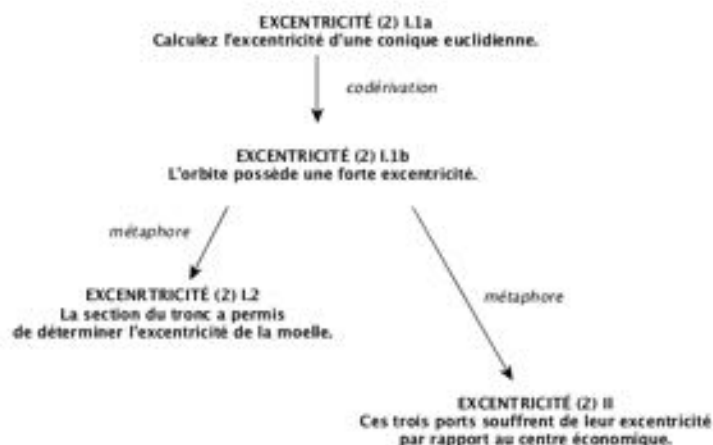


Figure 6 : Polysémie d'EXCENTRICITÉ²

4. Éléments de description formelle

Une fois que la polysémie d'EXCENTRICITÉ est dégagée, chacune des acceptions reçoit une description lexicographique. Le caractère strictement formel de celle-ci, impliquant le respect d'un certain nombre de contraintes n'a – comme on le verra dans cette section – rien de réducteur. Bien au contraire, appliqué consciencieusement, ce traitement positionne chaque unité lexicale par rapport à d'autres, en mettant en évidence ses idiosyncrasies. Nous les discuterons ici, au fil de ce que Gaume appelle balade aléatoire²⁴ dans le sous-réseau qui s'organise autour d'EXCENTRICITÉ¹ I.1 (cf. figure 5 ci-dessus).

4.1 Contenu informationnel d'un nœud lexical

Un nœud du RL-fr contient tout un ensemble d'informations lexicales, qui – même en l'absence de définition lexicographique à proprement parler²⁵ – constituent un article de dictionnaire. En cliquant sur le nœud EXCENTRICITÉ¹ I.1 (sommets triangulaires dans la partie gauche de la figure 7), on affiche son contenu lexicographique (volet à droite). Trois éléments de la description formelle proposée dans le RL-fr y sont pointés : l'étiquette sémantique, la forme propositionnelle et les fonctions (relations) lexicales.

24 Cf. Bruno Gaume, « Balades aléatoires dans les Petits Mondes Lexicaux », in *Information, Interaction, Intelligence*, 4.2, 2004, pp. 39-96.

25 À terme tous les nœuds lexicaux contiendront une définition. Celle-ci devra néanmoins rester cohérente avec l'ensemble des données, et même laisser prédire certaines d'entre elles, notamment les relations lexicales (cf. figure 7 : fonctions lexicales). À propos de l'évolution des contraintes pesant sur les définitions, on pourra se reporter à Igor Mel'čuk *et al.*, 1995 (voir note 3), Dorota Sikora, « Définir le sens dans un réseau lexical », in *Studia Romanica Posnaniensia*, 39/3, 2012, pp. 63-79, et surtout Igor Mel'čuk, Alain Polguère, « Définition lexicographique selon la Lexicologie Explicative et Combinatoire », in *Cahiers de Lexicologie* 109, sous presse, pp. 63-94.

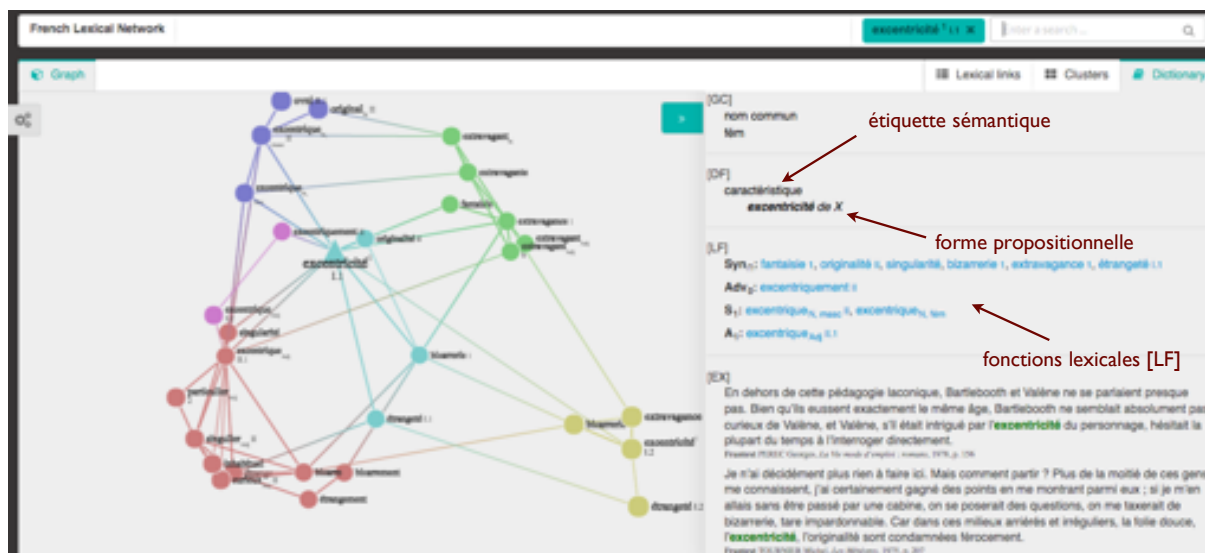


Figure 7 : Nœud lexical EXCENTRICITÉ¹ I.1 et son contenu lexicographique

L'étiquette sémantique²⁶ est une forme normalisée du genre prochain de la future définition. C'est une caractérisation minimale, dont la fonction consiste à situer la lexie vedette, en l'occurrence EXCENTRICITÉ¹ I.1, dans une classe sémantique. Le choix de l'étiquette n'est pas anodin dans la mesure où les unités lexicales d'une classe de sens contiennent des composantes communes que l'on retrouvera ensuite dans leurs définitions. Pour EXCENTRICITÉ¹ I.1, nous avons choisi 'caractéristique', en proposant de l'affiner par une étiquette-fille 'manière d'être'.

La forme propositionnelle rend compte du caractère prédicatif d'une lexie telle que EXCENTRICITÉ¹ I.1 : une manière d'être (ou une caractéristique) a toujours besoin d'un porteur. Il n'existe pas d'excentricité en soi – les articles réunis dans ce volume le montrent clairement – on ne peut parler que d'excentricité de X. Notre lexie vedette est donc un prédicat nominal avec un argument sémantique : le porteur du trait EXCENTRICITÉ¹ I.1.

La forme propositionnelle permet également de tisser un certain nombre de relations (fonctions) lexicales dont EXCENTRICITÉ I.1 est la source. Ainsi, le lexique français dispose de lexies permettant de nommer (EXCENTRIQUE_{N,masc}, EXCENTRIQUE_{N,fém}) et de qualifier (EXCENTRIQUE_{Adj}) l'argument X.

Avant de proposer une lecture plus approfondie de ces données, deux points méritent d'être soulignés. Premièrement, si l'aspect formel de ce travail lexicographique peut paraître quelque peu rébarbative, la cohérence entre les différents éléments de la description à l'intérieur d'un nœud constitue un apport informationnel considérable. En effet, l'étiquette sémantique et la forme propositionnelle doivent être compatibles entre elles pour permettre ensuite de tisser les liens avec d'autres sommets du réseau que forme le lexique français. Deuxièmement, considérées ensemble, ces éléments descriptifs permettent de brosser un portrait en demi-teintes de toutes les EXCENTRICITÉ-s et de leur fonctionnement dans le lexique, comme nous le montrerons dans la sous-section suivante.

4.2 Ce que révèlent les éléments d'une description formelle

Cette section propose un regard plus attentif sur ce que les unités lexicales EXCENTRICITÉ¹ I.1 et I.2 (cf. section 3, figure 5), c'est-à-dire sur les deux sens étiquetés respectivement 'manière d'être' et 'action d'un certain type'.

La figure 8 ci-dessous montre le mini-graphe lexical (à gauche) et le contenu informationnel du nœud lexical (à droite) qui se forme autour de EXCENTRICITÉ I.2 ('action d'un certain type').

26 Alain Polguère, « Classification sémantique des lexies fondée sur le paraphrasage », in *Cahiers de Lexicologie*, 98-1, 2011, pp. 197-211.

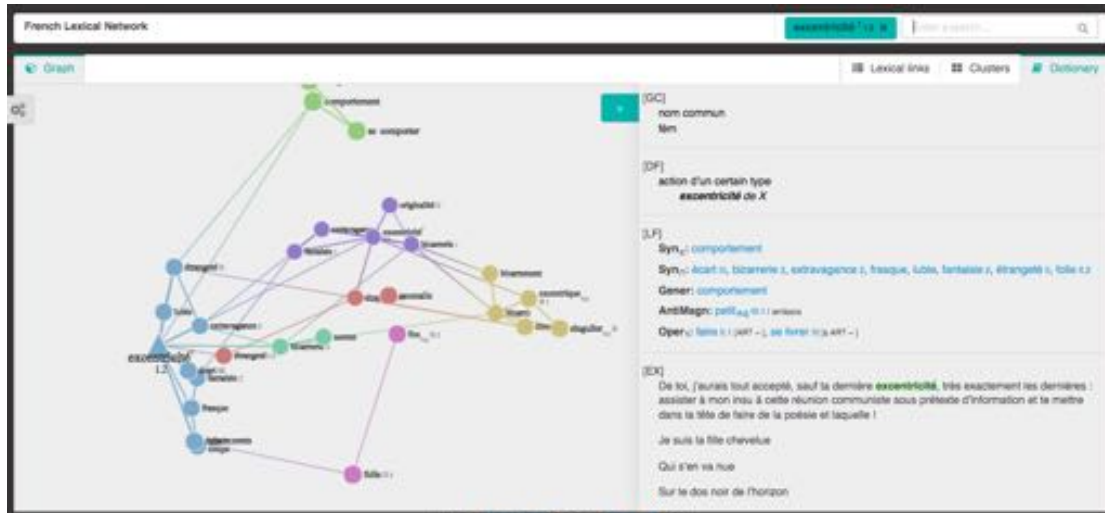


Figure 8 : Nœud lexical EXCENTRICITÉ¹ I.2 et son contenu lexicographique

Une comparaison avec le contenu du nœud lexical représenté dans la figure 7 ci-dessus permet de repérer des différences notables. Contrairement à EXCENTRICITÉ¹ I.1, qui en tant que ‘manière d’être’ (étiquette-fille de ‘caractéristique’) est un nom massif et de ce fait, résistant aux emplois pluriels, son copolysème I.2 est une action et, par conséquent, peut avoir plusieurs instances. Linguistiquement, il s’agit donc d’un nom comptable : il est possible de faire deux, trois ou plusieurs excentricités à suivre.

Les unités lexicales vers lesquelles pointent les liens de EXCENTRICITÉ¹ I.2 sont bien différentes de celles liées à son ULB. Au-delà d’un ensemble de quasi-synonymes et d’un terme générique (COMPORTEMENT) qui lui sont propres, on notera l’absence de noms et d’adjectif relatifs au premier actant X. En effet, celle ou celui qui ‘fait’ ou ‘se livre’ à une ou plusieurs excentricités (I.2) successives n’est pas nécessairement une ou un excentrique, c’est-à-dire porteur de la caractéristique EXCENTRICITÉ I.1.

Parmi les cibles de liens lexicaux, il est parfois intéressant de repérer celles qui brillent par leur absence. Ainsi, si notre X peut faire ou se livrer à des excentricités (I.2), en général, il ne les commet pas. Autrement dit, le verbe COMMETTRE, qui se combine régulièrement avec des actions étiquetées comme ‘méfaits’, ne figure pas parmi les relations syntagmatiques de EXCENTRICITÉ I.2. Il en ressort que, du point de vue sémantique, cette unité lexicale ne véhicule aucune composante négative ou répréhensible²⁷. Un corpus allant de 1800 à nos jours (soit 3390 textes d’un total de 215 893 750 mots) permet de repérer à peine trois occurrences de la collocation ‘commettre une excentricité’, dont celle dans *L’incognito* de Hervé Guibert :

Le roi Juan Carlos est en vacances sur l’île d’Elbe. La gazette locale, le Tirreno, lui consacre chaque jour sa première page, avec le détail de son emploi du temps, ses promenades en yacht, ses soirées chez le couturier Trussardi, les gens du village s’arrêtent devant les panneaux publicitaires de la librairie pour apprendre quelles excentricités le roi aura encore commises. Il ne peut plus faire de croisières en mer, la tempête s’est levée.²⁸

S’il n’est donc pas impossible d’employer la lexie EXCENTRICITÉ I.2 avec COMMETTRE, ce dernier ne fait pas pour autant partie de sa combinatoire restreinte. Il suffit à ce titre de la comparer avec FAUX-PAS, MALADRESSE et bien d’autres qui appellent quasi-automatiquement ce verbe dans une collocation. Bien au contraire, dans les trois emplois avec COMMETTRE attestés en corpus, le choix de ce collocatif s’explique par la volonté de conférer à

27 Notre discussion s’appuie ici sur des données de corpus et elle a un caractère purement sémantique. Nous ne nous intéressons d’aucune façon à des perceptions collectives ou individuelles que telle ou telle excentricité et son auteur peuvent avoir dans une situation particulière. On peut supposer cependant qu’en cas de comportement réellement répréhensible, on ne parlera plus d’excentricité.

28 Frantext, Hervé Guibert, *L’incognito*, 1989, p. 224.

EXCENTRICITÉ I.2 un caractère négatif de méfait que cette unité lexicale n'a pas intrinsèquement.

5. Conclusions

Notre balade dans le graphe lexical s'arrête là, mais plusieurs chemins d'exploration restent ouverts. On pourrait par exemple suivre les liens de quasi-synonymie pour identifier ce que EXCENTRICITÉ I.1 et EXTRAVAGANCE ont sémantiquement en commun et ce qui les distingue.

Dans cet article, nous avons porté un regard purement linguistique sur le vocable EXCENTRICITÉ et sur ses acceptions. Notre objectif était de montrer comment une analyse et description formelles, respectant des principes clairement définis, conduit à positionner des unités lexicales telles que EXCENTRICITÉ I.1 et I.2 dans le lexique, pour ensuite explorer les réseaux autour d'elles en quête d'informations sur ce qui leur est propre.

La réponse à la question posée dans le titre est donc affirmative : oui, comme toute lexie de la langue, les différentes EXCENTRICITÉ-s non seulement peuvent, mais doivent faire objet d'un rigoureux traitement formel.

Combinatoire lexicale et polysémie : le cas de MANQUER

Dorota Sikora (Université du Littoral-Côte d'Opale)

I. INTRODUCTION

1. Question de recherche

En analysant deux acceptions du verbe MANQUER¹, cet article propose une réflexion sur l'apport de la combinatoire lexicale à la discrimination des sens copolysèmes² et à leur description, notamment au sein d'une ressource dictionnaire. Sauf en cas de monosémie, relativement rare dans le lexique d'une langue, toute entrée de dictionnaire est construite selon une hypothèse sur la distinction des acceptions dont il faut y rendre compte. Or, l'un des principaux critères permettant d'étayer les intuitions des sens non-unifiables repose sur la combinatoire lexicale de chacun d'entre eux. Autrement dit, on admet que des acceptions copolysèmes se repèrent entre autres par leur autonomie compositionnelle (Croft et Cruse)³, qui se manifeste dans les collocations qu'elles sont capables de former. De ce point de vue, l'adjectif MALADE peut être considéré comme un cas d'école, puisque ses acceptions MALADE I dans les exemples (1) et l'extension métaphorique MALADE II en (2) se combinent avec des modificateurs adverbiaux différents. Certes, on pourrait accepter les énoncés (2a) et (2b), à condition néanmoins d'y interpréter MALADE au sens I, en confirmant ainsi l'utilité de l'autonomie compositionnelle dans l'analyse de la polysémie d'un vocable.

MALADE I

1. Mon voisin a été malade, il a passé un mois à l'hôpital.
 - a) Mon voisin a été gravement malade, il a passé un mois à l'hôpital.
 - b) Mon voisin a été malade comme un chien / à crever, il a passé un mois à l'hôpital.
 - c) ? Mon voisin a été absolument / complètement / littéralement malade, il a passé un mois à l'hôpital.

MALADE II

2. Mon voisin a été malade d'avoir rayé sa voiture.
 - a) ? Mon voisin a été gravement malade d'avoir rayé sa voiture.
 - b) ? Mon voisin a été malade comme un chien / à crever d'avoir rayé sa voiture.
 - c) Mon voisin a été absolument / complètement / littéralement malade, d'avoir rayé sa voiture.

Cependant, dans de nombreux cas, la situation ne se présente pas de manière aussi limpide. Il n'est pas rare en effet que des lexèmes verbaux fonctionnent avec peu de collocatifs-modificateurs et que des acceptions différentes apparaissent avec les mêmes adverbiaux, par exemple pour ce qui est de leur intensification. En témoigne les deux copolysèmes de MANQUER dans les phrases (3) et (4).

3. La nourriture manque cruellement / un peu / terriblement / affreusement dans la région.

1 Tout au long de ce texte, nous désignons les formes de nommage de vocables et de leurs acceptions par de petites capitales. Leurs occurrences dans des syntagmes sont indiquées en italiques.

2 Nous empruntons les termes de *copolysémie* et de *copolysème* à Polguère (2018). Le premier désigne l'ensemble des relations sémantiques entre les acceptions appartenant à un même vocable polysémique. Il s'agit ainsi de mettre en avant la nature relationnelle de la polysémie, qui la distingue notamment de l'homonymie. On qualifie de copolysème chacune des acceptions qui forment la polysémie associée à un même signifiant.

3 Nous adoptons le terme proposé par Croft et Cruse (2004), mais bien sûr, les différences de combinatoire lexicale entre des acceptions copolysèmes d'un vocable sont discutées dans la littérature, sous différents noms, depuis bien plus longtemps (cf. Apresjan 1974, Mel'čuk, Clas et Polguère 1995, puis Kleiber 2008 à titre d'exemple).

4. Mes filles me manquent cruellement / un peu / terriblement / affreusement.

Le phénomène pourrait certes s'expliquer par le caractère syncatégorématique de ces collocatifs adverbiaux intensifiants. Notre objectif dans cet article est d'examiner néanmoins les collocations qui forment les deux acceptions copolysèmes du verbe MANQUER, illustrées dans les exemples (3) et (4) pour montrer qu'une combinatoire lexicale, même lorsqu'elle semble peu discriminante, mérite d'être explorée en quête d'informations pertinentes pour construire une paraphrase définitoire. Si cette étude est menée sur des cooccurrences de forme V + Adv attestées en corpus, les collocations n'y sont pas définies sur des bases fréquentielles. Nous en adoptons la définition proposée par Mel'čuk (2013 : 138) : une collocation est un phrasème compositionnel, dont l'un des éléments – la base – est librement choisi par le locuteur, alors que l'autre – le collocatif – est contraint d'une part par le sens à exprimer, et de l'autre, sur le plan lexical, par les contraintes sémantiques exercées par la base. Ainsi, le recours à l'adverbe GRAVEMENT (collocatif) dans l'exemple 1 est imposé en quelque sorte par le lexème MALADE I. On voit bien dès lors que la recherche des congruences sémantiques (cf. Mejri 2017) qui sous-tendent la relation syntagmatique entre une base et son collocatif est un va-et-vient. Les collocatifs sont susceptibles de conduire vers la composante du sens de la base qu'ils modifient, et de ce fait, identifier la contrainte en œuvre pour leur sélection.

2. Étude de cas : deux copolysèmes du vocable MANQUER

Les réflexions développées dans cet article ont été largement inspirées par le travail lexicographique d'implémentation de données lexicales dans le *Réseau Lexical du Français*, ressource de grande envergure dont la nomenclature couvre à ce jour⁴ plus de 28 000 unités lexicales. Cette base de données a une structure relationnelle, modélisant le lexique du français sous forme d'un réseau dont les unités lexicales forment les nœuds, reliés par des relations paradigmatiques et syntagmatiques. Le cadre théorique et méthodologique de cette ressource, son architecture, ainsi que les solutions adoptées pour résoudre de nombreuses difficultés ont fait objet de plusieurs publications (Polguère 2014, 2016, Sikora 2017), qui donneront au lecteur un panorama des défis de cette entreprise lexicographique.

Nous y avons modélisé la polysémie de MANQUER, en en identifiant douze acceptions et en implémentant les nœuds lexicaux correspondants. Notre objectif ici est d'aller au-delà de cette expérience lexicographique, en proposant une étude de données extraites de corpus pour explorer l'apport de la combinatoire lexicale à l'analyse et à la description du sens. Faute de place, nous n'essaierons pas de justifier notre hypothèse de copolysémie de ce vocable, ni d'appliquer nos analyses à tous ses copolysèmes. Nous nous limiterons à MANQUER I et II⁵, qui dénotent respectivement un état physique d'absence d'une entité et un état affectif ressenti par un expérimenteur (exprimé par un complément prépositionnel datif) en raison de son éloignement de l'actant-sujet. Nous analyserons les collocations dont chacune de ces deux acceptions forme la base et les composantes sémantiques qui sous-tendent sa relation, en apparence purement syntagmatique, avec le collocatif adverbial intensificateur. Notre objectif est de montrer que même lorsque la combinatoire lexicale, notamment celle d'un verbe, se limite à un nombre réduit de modificateurs adverbiaux, elle mérite un examen

4 Nous faisons référence à son état de développement du 15 décembre 2019.

5 La numérotation de ces acceptions a été simplifiée pour les besoins de cette étude. Dans la mesure où les autres copolysèmes ne sont pas discutés ici, il nous a paru utile de marquer l'écart sémantique par deux chiffres successifs. Dans le réseau de copolysémie encodé dans le *RL-fr*, notre acception MANQUER I correspond au sens de base MANQUER I.1, et le verbe psychologique MANQUER II de notre exemple 4 au nœud lexical MANQUER IV.

attentif. Des collocatifs identiques reposent en effet sur des relations de sens, qui fournissent des indices à prendre en compte dans la démarche descriptive en vue d'une définition lexicographique.

Dans la section suivante, nous observerons des collocations, extraites d'un sous-corpus contemporain de Frantext⁶, que les deux acceptions de MANQUER forment avec des modificateurs adverbiaux. On analysera la nature de l'intensification qu'ils opèrent, en identifiant la composante sémantique de la base verbale qui en est l'objet.

Une seconde étape consistera, pour chacun des deux copolysèmes MANQUER, à vérifier s'ils acceptent tout type d'intensificateurs (massifs ? scalaires ?), en admettant qu'une combinatoire négative, c'est-à-dire des collocatifs impossibles sont souvent susceptibles d'apporter, eux aussi, des informations pertinentes sur le sens de la base. Nous espérons ainsi proposer des paraphrases définitives pour les deux acceptions de MANQUER, en nous appuyant largement sur l'étude des phénomènes de congruence, sémantique dans le cas des collocations, qui conditionnent les relations syntagmatiques entre les items lexicaux.

II. COMBINATOIRE LEXICALE DE *MANQUER* EN CORPUS

Pour repérer les collocatifs adverbiaux sélectionnés par le verbe MANQUER, nous nous sommes tournée vers la base textuelle Frantext, en limitant nos recherches à un sous-corpus contemporain allant de 1980 à 2019, soit à 550 documents d'une teneur de 35 939 491 mots. Le verbe MANQUER s'y manifeste par 7497 occurrences.

Notre requête a retourné 248 syntagmes verbaux *manquer* + *Adv*, soit 3,3 % à peine des emplois attestés dans notre corpus de travail, ce qui signifie que quel que soit son sens, ce verbe montre une tendance très faible à se combiner avec ce type de modificateurs. Ce nombre peu élevé a encore été réduit suite à un élagage. Premièrement, nous avons retenu uniquement des SV dont le modificateur adverbial a un caractère intraprédicatif, c'est-à-dire lorsqu'il porte sur le noyau sémantique du verbe. Du point de vue énonciatif, ce sont des adverbes endophrastiques, « constituants internes de la phrase, qui affectent le contenu même de l'élément sur lequel ils portent et, ce faisant, participent à la construction du sens référentiel de la phrase » (Guimier 1996 : 6). Dans un second temps, il a fallu désambiguïser les extraits pour garder uniquement ceux qui contiennent l'une des deux acceptions de MANQUER étudiées dans cet article. Ces procédures d'élimination ont conduit à retenir 207 SV, soit autant de collocations formées par nos deux copolysèmes de MANQUER avec 22 adverbes différents. Ils sont réunis dans le tableau 1 ci-dessous, avec leur fréquence absolue dans notre corpus de travail.

6 www.frantext.fr

| Rang | Adverbe | Fréq. | Rang | Adverbe | Fréq. |
|------|---------------------------|-------|------|-----------------|-------|
| 1. | cruellement | 29 | 12. | physiquement | 3 |
| 2. | peu (un peu, quelque peu) | 29 | 13. | horriblement | 3 |
| 3. | beaucoup | 29 | 14. | gravement | 3 |
| 4. | tellement | 21 | 15. | désespérément | 3 |
| 5. | terriblement | 20 | 16. | douleureusement | 3 |
| 6. | totalement | 15 | 17. | charnellement | 2 |
| 7. | absolument | 13 | 18. | sévèrement | 2 |
| 8. | autant | 12 | 19. | furieusement | 2 |
| 9. | affreusement | 7 | 20. | effroyablement | 1 |
| 10. | énormément | 4 | 21. | sensiblement | 1 |
| 11. | complètement | 4 | 22. | atrocement | 1 |

Tableau 1: Collocatifs adverbiaux de MANQUER I et de MANQUER II attestés en corpus.

Non seulement MANQUER semble peu porté sur la modification adverbiale, mais dans le peu de cas où il l'accepte, huit adverbes seulement sont attestés plus de dix fois dans notre corpus. De plus, les deux acceptions MANQUER I et MANQUER II affichent clairement les mêmes préférences adverbiales, comme le montrent les exemples (5) et (6) ci-dessous.

5. La nourriture manque cruellement / un peu / terriblement / affreusement dans la région.
6. Ma mère me manque cruellement / un peu / terriblement / affreusement.

Faut-il en conclure qu'il s'agit de polysémie sans autonomie compositionnelle ? Si l'on s'en tient aux relations purement syntagmatiques, il faudrait y répondre par l'affirmative. Mais en explorant les congruences sémantiques à la recherche des éléments qui se trouvent intensifiés (ou non) par les modificateurs adverbiaux, des différences importantes se dégagent.

III. *MANQUER I* ET SES COLLOCATIFS ADVERBIAUX

Si tous les adverbiaux présents dans la phrase (5) sont également compatibles avec les deux copolysèmes (et avec d'autres encore) de MANQUER, un examen plus attentif permet de voir que l'intensification qu'ils produisent n'est pas de même nature. Dans cette section, nous nous proposons d'examiner les rapports sémantiques qui consolident les collocations de MANQUER I avec des modificateurs adverbiaux d'évaluation quantitative et des quantifieurs intensifs-appréciatifs. Nous verrons que les deux types de collocatifs réalisent une intensification sur les composantes différentes du sens verbal.

1. Intensificateurs d'évaluation quantitative

MANQUER I se combine avec des adverbiaux scalaires – telle est en effet la lecture qu'il convient d'accorder à UN PEU dans (5) –, qui permettent d'évaluer une propriété, et plus précisément le degré de

sa possession, selon une échelle. Celle-ci est d'ailleurs fermée, puisque UN PEU s'inscrit ici dans une série comprenant des adverbes de complétude COMPLÈTEMENT, TOTALEMENT. La phrase (7) a été construite pour les besoins de notre démonstration, mais les énoncés (8) et (9) ont été extraits du web – il s'agit d'exemples authentiques. Il est intéressant de remarquer que, par ailleurs, une évaluation par quantification massive, comme celle que nous avons essayé d'exprimer en (10), n'est pas acceptable.

7. La nourriture manque un peu / quelque peu / complètement / totalement dans la région.

8. Malheureusement pour lui, il n'y a pas grand chose à trouver dans ma cuisine. La nourriture manque quelque peu, mais je ne vais pas commencer à me plaindre.

9. J'ai pensé surtout aux personnes du Soudan, d'Afrique et d'Éthiopie qui meurent de faim, car la nourriture manque totalement !

10. ? J'ai pensé surtout aux personnes du Soudan, d'Afrique et d'Éthiopie qui meurent de faim, car la nourriture manque tant / tellement / beaucoup !

La combinatoire lexicale attestée, ainsi que les incompatibilités constatées en (10), s'expliquent par la composante sémantique du sens verbal sur laquelle portent les modificateurs collocatifs, et qui se trouve intensifiée, à savoir l'état que le prédicat attribue à son premier actant X (*la nourriture* dans 7, 8 et 9). Observons que l'entité susceptible de l'instancier a une structure massive homogène. Par conséquent, X peut être soit totalement absent, soit présent en quantité toujours insuffisante, mais susceptible d'être évalué et quantifié selon une échelle. Il s'agit de ce fait d'un premier type d'intensité distinguée et discutée par Kleiber (2013) : elle se présente de manière scalaire, car elle repose sur le degré de possession de la propriété attribuée.

2. Adverbes quantificateurs intensifs-appréciatifs

Les adverbes CRUELLEMENT, TERRIBLEMENT, AFFREUSEMENT de l'exemple (5), ainsi que HORRIBLEMENT, GRAVEMENT, DÉSESPÉRÉMENT, DOULOUREUSEMENT, SÉVÈREMENT, FURIEUSEMENT, EFFROYABLEMENT, SENSIBLEMENT, ATROCEMENT (cf. le tableau 1 ci-dessus) sont classés par Molinier (1990) et Molinier et Levrier (2000) parmi les quantificateurs intensifs-appréciatifs. Ils auraient ainsi un effet intensifiant, lorsqu'ils sont collocatifs des bases adjectivales, alors qu'en modifiant les verbes, ils exprimeraient une appréciation. Si tel est le cas, qu'est-ce qui est soumis à l'évaluation appréciative ? Quelle est la composante du sens de MANQUER I susceptible de ce type de modification, et par là même, responsable de la sélection des adverbes de quantification intensive et appréciative en (11) ? L'exemple (12), prélevé dans un blog internet, emploie les deux copolysèmes de MANQUER qui nous intéressent ici. Dans cette section, nous nous pencherons sur la collocation *manquer terriblement* et reprendrons le cas de *manquer cruellement* dans celle qui suit.

11. La nourriture manque (I) gravement / sévèrement / cruellement / terriblement / affreusement / horriblement dans la région.

12. La nourriture manque (I) terriblement, on ne se rend pas compte à quel point nous mangeons exceptionnellement bien en France, le vin, le fromage, cela manque (II) cruellement au bout d'un moment.

Même si MANQUER I est incontestablement un support verbal, dans les exemples (11) et (12), la modification adverbiale apporte un sens d'intensification, du moins autant que celui d'appréciation. Par ailleurs, cette dernière s'oriente clairement vers le pôle négatif. Il est intéressant de remarquer que pour en identifier l'objet dans le contenu sémantique du prédicat, il s'avère utile de le nominaliser. Ainsi, pour reformuler (11), pourra-t-on dire en (13) :

13. Le manque de nourriture est grave / sévère / cruel / terrible / affreux / horrible pour la population / dans la région.

De fait, en évaluant les conséquences négatives que le manque de nourriture entraîne pour l'argument Y, les attributs de la série introduite dans la phrase (13) indiquent que le phénomène atteint un haut degré d'intensité. On l'observe également dans les phrases (14) à (16), qui rendent la signification de (13) de manière plus explicite encore.

14. La nourriture manque et les conséquences en sont graves / sévères / cruelles / terribles / affreuses / horribles pour la population / dans la région.

15. Les conséquences du manque de nourriture sont graves / sévères / cruelles / terribles / affreuses / horribles pour la population / dans la région.

16. La nourriture manque et la population / la région le ressent gravement / sévèrement / cruellement / terriblement / affreusement / horriblement.

Ainsi, l'effet intensifiant des collocatifs adverbiaux dans (11) est directement lié à l'appréciation négative des conséquences ressenties par Y. Par conséquent, plutôt que parler d'adverbes intensifs-appréciatifs, nous proposons de les considérer comme ceux d'intensification appréciative.

En même temps, on remarque que MANQUER I n'accepte pas de collocatifs adverbiaux de quantification massive tels que BEAUCOUP, TANT, TELLEMENT et UN PEU (dans cette interprétation). On comprend pourquoi : rien, dans le sens lexical de la base verbale, n'est susceptible d'être quantifié de la sorte. Une propriété scalaire telle qu'un état d'absence ne s'y prête évidemment pas et les conséquences pour Y autorisent plutôt une quantification comptable, comme celle en (17).

17. Le manque de nourriture a beaucoup de conséquences graves pour la population.

Si une définition lexicographique doit, comme le veut la *LEC*, permettre d'anticiper la combinatoire d'une unité lexicale, l'analyse de ses collocatifs permet *a contrario* de mettre le doigt sur les composantes à intégrer dans la paraphrase définitoire. Les modificateurs adverbiaux avec lesquels MANQUER I forme des collocations nous ont menée vers deux éléments sémantiques que le lexicographe doit expliciter dans une définition. D'une part, il y a l'état d'absence ou d'insuffisance de X, susceptible d'être évalué de manière scalaire avec UN PEU, QUELQUE PEU, COMPLÈTEMENT, TOTALEMENT. L'intensification porte alors sur le degré que présente une propriété particulière, en l'occurrence celle que le prédicat attribue à son actant X. D'autre part, les effets produits pour Y subissent une évaluation appréciative, ou plutôt dépréciative, puisqu'elle s'oriente systématiquement vers le pôle négatif. Avec ces informations, nous pouvons esquisser une paraphrase définitoire du lexème verbal MANQUER I.

MANQUER I : $X \sim à Y$

'X est présent en quantité insuffisante ou est absent de Y, qui en éprouve des effets négatifs'

IV. MANQUER II – VERBE DE SENTIMENT ET SES COLLOCATIFS

1. Adverbes de quantification massive

Comme son copolysème, le verbe de sentiment MANQUER II a deux actants sémantiques : X (*ma mère*) et Y réalisé par le clitique *me* dans l'exemple (6). Contrairement à son unité lexicale de base, il s'accorde cependant avec des modificateurs de quantification massive TANT, TELLEMENT, ÉNORMÉMENT, ainsi que BEAUCOUP et UN PEU qui reçoivent dès lors cette lecture justement. Les collocatifs exprimant une évaluation scalaire, y compris les adverbiaux de complétude, sont au contraire impossibles. Dans l'extrait (18), on pourra en effet substituer TANT, TELLEMENT, ÉNORMÉMENT ou UN PEU (dans une lecture massive) à BEAUCOUP, alors que QUELQUE PEU, COMPLÈTEMENT et TOTALEMENT sont exclus, comme le montre la manipulation effectuée en (19).

18. Bien avant Stalingrad il s'aperçoit que ces trois garçons lui manquent beaucoup. (Frantext)

19. ? Bien avant Stalingrad il s'aperçoit que ces trois garçons lui manquent quelque peu / complètement / totalement.

La possibilité de quantification massive, qui se manifeste dans l'exemple (18) entre autres, est caractéristique, selon Kleiber (2013), d'un second type d'intensité, différente de celle que nous avons repérée pour MANQUER I : c'est celle qui est elle-même une propriété, l'un des aspects de l'état dénoté par le lexème verbal (cf. Niedenthal *et al.* 2004). Comme une dimension constitue l'une des propriétés d'un objet physique, cette intensité en est une dans le cas d'un son, d'une couleur, d'une odeur, d'un sentiment, du courant électrique, etc. Elle est différente selon le type d'entité « dont elle est une qualité, mais elle présente néanmoins un dénominateur commun plus ou moins stable, celui d'impact sensoriel » (Kleiber 2013 : 71).

En attendant de dégager d'autres composantes définitionnelles, on retient la nécessité d'en prévoir une, dans la paraphrase du sens, qui rende compte de l'intensité caractérisant l'état affectif de Y.

2. Adverbes d'intensification appréciative

Les adverbes d'intensification appréciative, nous l'avons dit, sont sélectionnés par nos deux copolysèmes, ce qui, sur le plan syntagmatique, pourrait conduire à la conclusion de l'absence d'autonomie compositionnelle. Néanmoins, une étude plus attentive de ces collocatifs en apparence peu spécifiques apporte des informations sémantiques pertinentes pour la paraphrase du sens lexical de la base. Une fois de plus, il convient dès lors de s'interroger sur la composante de son sens qui se trouve intensifiée par le biais d'une appréciation. On peut ainsi travestir l'exemple (18), en intensifiant le sentiment exprimé par des collocatifs adverbiaux de cette classe :

20. Bien avant Stalingrad il s'aperçoit que ces trois garçons lui manquent cruellement / sévèrement / terriblement / affreusement / horriblement / douloureusement / atrocement.

Or, pour le lexème MANQUER II, le recours à la nominalisation est problématique dans la mesure où le format nominal ne semble pas lexicaliser fidèlement le contenu sémantique du prédicat verbal. Une comparaison de *manque* de l'exemple (13) rappelé ci-dessous et de notre tentative de translation nominale de la phrase (20) en (21) met en avant cette difficulté.

13. Le manque de nourriture est grave / sévère / cruel / terrible / affreux / horrible pour la population / dans la région.

21. ? Le manque de ces trois garçons est cruel / sévère / terrible / affreux / horrible / douloureux / atroce.

Force est de reconnaître que le sentiment qu'exprime MANQUER II résiste à la nominalisation et que le nom correspondant fait défaut dans le lexique français. Pour identifier l'élément qui est soumis à l'appréciation, il nous faudra recourir au nom générique SENTIMENT, qui reprend anaphoriquement le contenu du prédicat verbal en (22).

22. Bien avant Stalingrad il s'aperçoit que ces trois garçons lui manquent et ce sentiment est cruel / sévère / terrible / affreux / horrible / douloureux / atroce.

Il apparaît ainsi que lorsque MANQUER II forme des collocations avec des adverbes d'intensification appréciative, ces modificateurs portent sur le ressenti de l'actant Y. Son état affectif à caractère clairement négatif est soumis à une appréciation, qui à son tour produit un effet d'intensification.

3. Quelques indices par les collocatifs adverbiaux de manière

Le tableau 1 ci-dessus fait état de quelques adverbes de manière (DÉSÉSPÉRÉMENT, FURIEUSEMENT, PHYSIQUEMENT, CHARNELLEMENT), dont on trouve des attestations en tant que modificateurs de MANQUER II. Dans la mesure où ils sont peu nombreux, dans un corpus de textes littéraires qui plus est, on peut se demander d'ailleurs s'il s'agit encore de collocations ou plutôt des cas de combinatoire libre. Ils semblent néanmoins mériter un détour sémantique.

23. C'est dans ces moments-là que la Tunisie m'a le plus manqué physiquement. (Frantext)

24. Je ressentais aussitôt combien charnellement me manquaient cette Bretagne, cette Bourgogne ou cette Normandie, si proches de mon cœur, si familières à ma mémoire, et où je n'avais jamais mis le pied. (Frantext)

25. Mon incarcération à la prison de Champrond. Faux. Jérémiesque besoin d'étendre sur nos misères l'ombre odoriférante de Stojil, voilà tout. C'est qu'elle nous manque furieusement, ces temps-ci, l'ombre de notre oncle Stojil ! (Frantext)

Le recours au nom générique en reprise anaphorique en (26) montre qu'il ne s'agit pas d'un jugement appréciatif porté sur le caractère du ressenti.

26. Bien avant Stalingrad il s'aperçoit que ces trois garçons lui manquent et ? ce sentiment est physique / charnel / désespéré / furieux.

De fait, de nombreuses études, à commencer par celle de Johnson-Laird et Oatley (1989), insistent sur les états physiques par lesquels se manifestent les affects. Ces manifestations sont souvent inscrites dans le sens lexical, ce qui signifie qu'une paraphrase définitoire peut contenir une composante qui les explicite. Contrairement à HONTE par exemple, dans le cas de MANQUER II, il ne semble pas y avoir de manifestation physique inscrite dans le sens lexical⁷. Cela n'empêche pas pour autant le locuteur de le caractériser en cas de besoin. Il est en effet possible de paraphraser les exemples (23) à (25) par une phrase évoquant la manière dont l'état affectif en question se manifeste.

27. Bien avant Stalingrad il s'aperçoit que ces trois garçons lui manquent et ce sentiment se manifeste de manière physique / charnelle / furieuse / par du désespoir.

⁷ La question mérite certainement une étude plus poussée, notamment en ce qui concerne les indices présents dans la combinatoire lexicale : si l'on peut former la collocation *rouge de honte*, aucun collocatif de ce type n'est disponible pour MANQUER II. Cela conduit à penser que la manifestation physique n'est pas codée dans le lexème et que les adverbes de manière étudiés dans cette section relèvent de la combinatoire libre.

Pour MANQUER II, les collocations formées avec les adverbes de quantification massive, ceux d'intensification appréciative et de manière ont permis de détecter la composante d'intensité qui caractérise le sentiment dénoté, sa polarité négative, ainsi que ses manifestations possibles. Il convient d'y intégrer également le présupposé que véhicule ce lexème verbal, celui d'un contact et d'une présence préalable de X auprès de Y. En effet, en cas de séparation, X peut manquer (ou ne pas manquer d'ailleurs) à Y, si et seulement si, il y avait auparavant une proximité physique. Une paraphrase définitoire pourrait dès lors prendre la forme suivante, avec entre crochets une composante présuppositionnelle.

MANQUER II : X ~ à Y

[Y ayant été dans le passé en présence de X], l'éloignement de X cause chez Y un sentiment négatif pour lui, dont l'intensité peut varier, Y souhaitant être de nouveau en présence de X

V. CONCLUSIONS

Dans cette étude, nous avons exploré la combinatoire lexicale de deux acceptions du vocable MANQUER – deux unités lexicales à part entière, appartenant à des champs sémantiques différents. Une recherche en corpus a montré qu'elles forment peu de collocations, essentiellement avec des adverbes intensificateurs, souvent compatibles avec les deux sens. Ainsi, on pourrait admettre que les lexèmes MANQUER I et MANQUER II sont en relation de copolysémie sans autonomie compositionnelle. Cependant, un examen plus approfondi de leurs collocations nous a permis de mettre en avant des restrictions de leurs combinatoires respectives. Celles-ci excluent, pour le premier de ces sens, les adverbes de quantification massive, parfaitement acceptables pour le verbe de sentiment MANQUER II. À son tour, ce dernier se montre incompatible avec les modificateurs scalaires. Ces analyses nous amènent vers deux conclusions. Premièrement, ces compatibilités montrent qu'il y a bien une autonomie compositionnelle, même si elle n'est que partielle. Deuxièmement, l'intensification réalisée par les adverbes n'est pas de même nature dans le cas de MANQUER I et de MANQUER II, car elle ne porte pas sur la même composante sémantique de leur sens lexical.

Lorsque l'autonomie compositionnelle est partielle, on a inévitablement affaire à une combinatoire qui présente des traits communs, c'est-à-dire à une quantité plus ou moins importante de collocatifs sélectionnés par les deux copolysèmes. Nous nous sommes attachée à montrer que ces modificateurs adverbiaux en apparence peu spécifiques méritent une analyse, car par leur portée, ils peuvent mettre en avant des composantes sémantiques propres à chacune des acceptions et de ce fait, servir de repères pour construire une définition lexicographique.

À l'issue de ce parcours, il nous semble intéressant d'en entreprendre un autre. Il prendrait comme point de départ non pas la base de la collocation (par exemple MANQUER I ou MANQUER II, comme nous l'avons fait dans cette étude), mais le collocatif (CRUELLEMENT, TERRIBLEMENT, etc.) pour examiner les différentes bases susceptibles de le sélectionner et vérifier s'il est possible de dégager ainsi des éléments pertinents pour construire des paraphrases définitoires de ces adverbes qui méritent bien le nom de *pervers polymorphes* (Goes 2005), notamment à cause de leur sens fuyant.

BIBLIOGRAPHIE

APRESJAN Juri D. « Regular Polysemy », *Linguistics*, 1974, 142, p. 5-32.

- CROFT William et CRUSE Alan, *Cognitive Linguistics*, Cambridge, Cambridge University Press, 2004, 374 p.
- GOES Jan (éd.) *L'adverbe: un pervers polymorphe*, Arras, Artois Presses Université, 2005, 304 p.
- GUIMIER Claude *Les adverbes du français. Le cas des adverbes en –ment*, Paris, Ophrys, 1996, 170 p.
- JOHNSON-LAIRD Philipp. N. et OATLEY Keith « The Language of Emotions: An Analysis of a Semantic Field », *Cognition and Emotion*, 1989, 3 (2), p. 81-123.
- KLEIBER Georges « À la recherche de l'intensité », *Langue Française*, 2013, 177, p. 63-76.
- MEJRI Salah « Les trois fonctions primaires. Une approche aystématique. De la congruence et de la fixité dans le langage », in CARVALHO C., PLANELLES IVÁÑEZ M. et SANDAKOVA E. *De la langue à l'expression: le parcours de l'expérience discursive: Hommage à Marina Aragón Cobo*, Alicante, Universidad de Alicante, 2017, p. 123-144.
- MEL'ČUK, Igor A. « Tout ce que nous voulions savoir sur les phrasèmes, mais... », *Cahiers de Lexicologie*, 2013, vol. 102, n° 1, p. 129-149.
- MEL'ČUK, Igor A., CLAS André et POLGUÈRE Alain *Introduction à la Lexicologie Explicative et Combinatoire*, Louvain-la-Neuve, Éditions Duculot, 1995, 256 p.
- MOLINIER, Christian « Une classification des adverbes en –ment », *Langue Française*, 1990, 88, p. 28-40.
- MOLINIER, Christian et LEVRIER Françoise *Grammaire des adverbes. Description des formes en –ment*, Genève-Paris, Droz, 2000, 258 p.
- NIEDENTHAL Paula, AUXIETTE Catherine, NUGIER Armelle, DALLE Nathalie, BONIN Patrick, FAYOL Michel « A Prototype Analysis of the French Category “Émotion” », *Cognition and Emotion*, 2004, vol. 18, n° 4, p. 289-312.
- POLGUÈRE, Alain « From Writing Dictionaries to Weaving Lexical Networks », *International Journal of Lexicography*, 2014, vol. 27, n° 4, p. 396-418.
- « La Question de La Géométrie Du Lexique », in *Congrès Mondial de Linguistique Française 2016*. SHS Web of Conferences, 2016
www.shs-conferences.org/articles/shsconf/abs/2016/05/contents/contents.html, consulté le 17 mars 2017.
- « A Lexicographic Approach to the Study of Copolysemy Relations », *Russian Journal of Linguistics*, 2018, vol. 22, n° 4, p. 788-820.
- SIKORA, Dorota « Modélisation de connaissances lexicales dans une ressource », *Revue Française de Linguistique Appliquée*, 2017, vol. XXII, n°1, p. 61-74.

IDENTIFICATION D'UNITES PHRASEOLOGIQUES ET EQUIVALENCE SEMANTIQUE DANS LA TRADUCTION

Dorota Sikora

Université du Littoral – Côte d'Opale

& HLLI (EA 4030)

Dorota.Sikora@univ-littoral.fr

Abstract

Although lexical resources tend to offer more and more extended coverage of multi-word expressions, the available machine translation programs still encounter difficulties in dealing with phraseological units. In some contexts, sentences they generate are infelicitous, or even grammatically incorrect. Such translation problems rise questions of the ways online dictionaries describe multi-word expressions. Considering the case of four French and English strong idioms (fully non-compositionnal phrasemes, cf. Mel'čuk, 2013), the present paper discusses the encoding of phraseological units in two lexical databases : the *French Lexical Network* and the *English Lexical Network* built in the same theoretical and methodological frame, as implementations of a graph model named lexical system (Polguère, 2009, 2014). Since each of these resources is basically monolingual, multilingual connexions between two or more lexical systems can be implemented. Beyond the particular case of the idioms the present study focuses, we suggest that a homogeneous and coherent treatment of multi-word expressions in lexical resources underlying machine translation programs can significantly improve rendering of idiomaticity in the target language.

1. INTRODUCTION

Le présent article propose une réflexion sur ce qui – selon notre hypothèse – est l'une des principales sources d'erreurs dans la traduction automatique d'unités phraséologiques (désormais UP) : un diagnostic imprécis de leur statut linguistique. Nous présenterons le traitement des locutions fortes (ou phrasèmes complets, cf. Mel'čuk, 2013 : 134) dans deux bases de données lexicales développées actuellement à l'Atilf (CNRS UMR 7118) : *Réseau Lexical du Français (RL-fr)*, cf. Gader et al., 2012, Polguère, 2014), et *English Lexical Network (en-LN)*, Gader et al., 2014). Ces ressources sont construites en parallèle, selon les principes théoriques et méthodologiques homogènes, définis initialement dans le cadre de la Lexicologie Explicative et Combinatoire et développés récemment par Polguère (2009, 2014) dans un modèle de *système lexical* (cf. section 2).

Si l'intérêt des lexicographes pour l'intégration des UP dans les ressources dictionnaires est aujourd'hui manifeste (cf. Osherson et Fellbaum, 2010, Fellbaum,

2014), les utilisateurs n'en constatent pas moins de problèmes dans les résultats obtenus. Trois types de difficultés sont régulièrement repérés.

Premièrement, dans un contexte monolingue, la couverture que proposent les dictionnaires reste faible (cf. Vaguer, 2010). Elle est de plus hétérogène : même les locutions les plus fréquentes ne s'y retrouvent pas systématiquement. Dans une perspective bilingue, cela signifie qu'il est difficile de trouver l'équivalent d'une UP de la langue L₁ dans une ressource de la langue L₂.

Deuxièmement, il arrive souvent que tout en préservant l'équivalence sémantique entre la phrase source en L₁ et la traduction proposée en L₂, l'idiomaticité de l'original se perd, comme dans l'exemple (1b) :

- (1) a. Cynthia et Ted étaient aux anges. (Frantext)
b. Cynthia and Ted were ecstatic. (<https://translate.google.fr>)

D'un point de vue communicatif, il s'agit certes d'un moindre mal, même si comme l'observe Svensén (2009 : 156), « [i]dioms in the source language must as far as possible be paralleled in the target language by idioms with the same content ».

Un troisième cas de figure est celui où une même ressource produit des résultats corrects dans certains contextes, alors que dans d'autres, elle restitue des phrases inacceptables. L'exemple (1c) montre que Reverso établit correctement l'équivalence entre le syntagme *être aux anges* dans (1a) et le phrasème anglais *to be in seventh heaven*. Cependant, la phrase (2a) pose un problème d'analyse aux deux utilitaires.

- (2) a. Tu as une voix qui met aux anges... (FrWac)
b. You have a voice(vote) which puts ?in the angels... (traduction Reverso, 12 juin 2015)
c. You have a voice that ?makes angels... (traduction Google, 12 juin 2015)

Notre étude se concentre sur ce dernier type de problèmes. Pour identifier leur source, observons la figure (1) qui reproduit l'information fournie par Reverso à ses utilisateurs : *être aux anges* est encodé comme une locution verbale, équivalente de la collocation *be extremely happy*⁴⁷, considérée à son tour comme quasi-synonyme de plusieurs UP en L₂. Selon notre hypothèse, l'origine des erreurs constatées en (2b) et en (2c) se trouve dans cet encodage non conforme à la réalité linguistique.

Dans la section 2, nous présenterons le modèle de *système lexical* (cf. Polguère, 2009, 2014), ainsi que le mode d'encodage et la structuration d'informations dans le *RL-fr* et dans le *en-LN*, ressources lexicales qui sont des implémentations de ce modèle. La section 3, consacrée à l'analyse d'un corpus d'exemples français et anglais qui attestent des caractéristiques de quatre phrasèmes : *être aux anges*, *être au septième ciel* en français, *to be over the moon*, *to be in seventh heaven*. Notre objectif sera de bien distinguer la locution elle-même et les collocations qu'elle peut former. Les conclusions proposées en 4 porteront sur les possibilités d'interconnexion de systèmes lexicaux.

⁴⁷ L'amalgame évident entre la définition proposée (la collocation *be extremely happy about smth*) et une série de quasi-synonymes en L₂ (*be walking on air*, etc.) présentés comme des locutions verbales n'est certainement pas étranger aux problèmes de traduction constatés en (2b).



Figure 4. Traduction du syntagme *aux anges* et description lexicographique de *être aux anges* proposées par Reverso (18/03/2015)

2. TROIS RESSOURCES, DEUX SYSTEMES LEXICAUX

Malgré l'homogénéité méthodologique qui les caractérise, les deux ressources ont des histoires différentes. Le *RL-fr* a démarré en juin 2011, avec une nomenclature dite d'amorçage de 3734 vocables lexémiques, compilée à partir de ressources pédagogiques (cf. Polguère et Sikora, 2013), pour couvrir le vocabulaire de base (McCarthy, 1999) du français contemporain. Depuis, sa croissance se poursuit selon les règles décrites dans plusieurs publications (cf. Lux-Pogodalla et Polguère, 2011, Polguère et Sikora, 2013). Les UPs, notamment les locutions fortes, qui ne figuraient pas dans la nomenclature d'amorçage, sont systématiquement décrites par les lexicographes.

La base *en-LN* a été créée en 2012 par conversion des données de WordNet. Pour une présentation détaillée de cette opération, on pourra se reporter à Gader et al. (2014), mais il est important de souligner qu'elle a consisté dans une transformation d'un modèle ontologique du lexique reposant sur une hiérarchisation conceptuelle des lexies en un système non ontologique (i. e. en un système lexical, cf. *infra*), organisé par une multitude de relations de différente nature (paradigmatiques, syntagmatiques, de copolysémie entre les lexies d'un vocable polysémique et d'inclusion formelle entre les UPs et les lexies qui les composent). Techniquement, cela signifie qu'un graphe lexical dont les nœuds sont formés par des synsets, ensembles de lexies avec une composante conceptuelle commune (WordNet), devait être converti en un graphe connectant des lexies, avec des éléments descriptifs de nature purement linguistique associés à chacune d'elles. Les figures 2 et 3 représentent respectivement la locution *seventh heaven* dans WordNet et son entrée dans *en-LN* issue de l'injection.

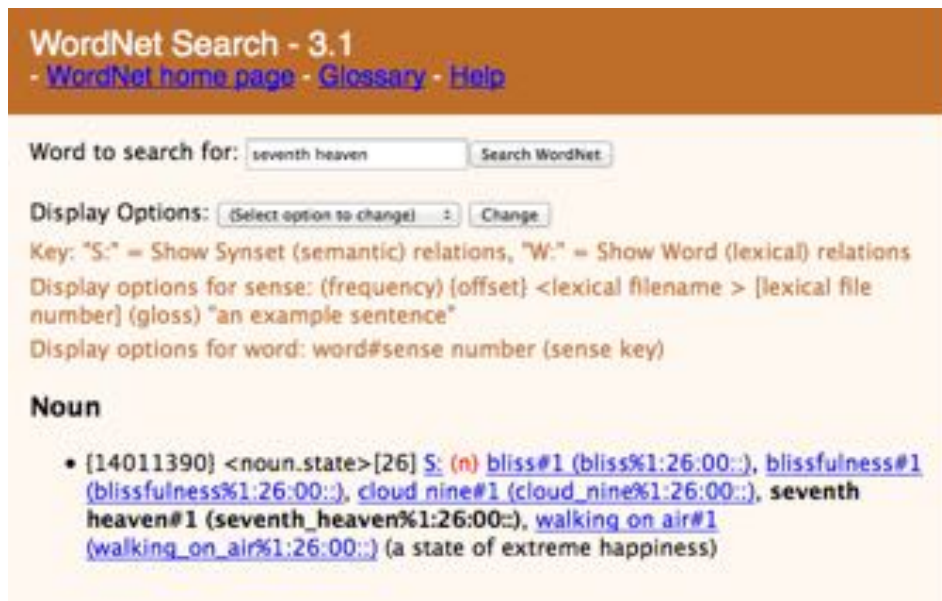


Figure 5. Locution *seventh heaven* dans WordNet.

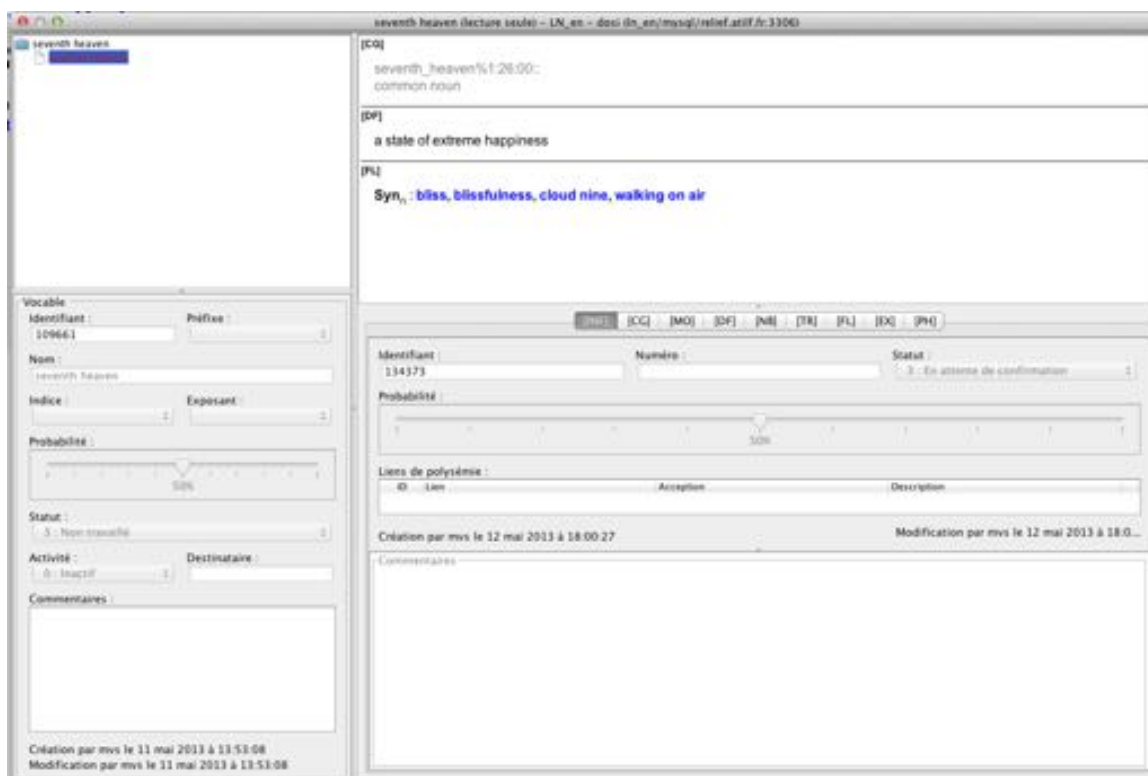


Figure 6 : Locution *seventh heaven* injectée automatiquement dans *en-LN* par conversion des données de WordNet.

Un système lexical tel qu'il est proposé par Polguère (2014) est une modélisation du lexique sous forme de réseau. C'est un graphe au sens mathématique, fait de sommets que constituent les unités lexicales (lexies) et des arcs représentant des relations lexicales. Cette structure est une première des quatre caractéristiques des systèmes lexicaux. Deuxièmement, il s'agit – rappelons-le – d'un modèle non ontologique, qui ne repose sur aucune classification hiérarchique des concepts, ni de connaissances. Le réseau est tissé par encodage de multiples relations lexicales sans que celles-ci soient hiérarchisées.

La structure interne d'un nœud lexical est non-atomique : elle encapsule un ensemble d'informations sur la lexie vedette, allant des indications générales [l'onglet INF] (identifiant de la lexie, état d'avancement de la description, ainsi que – le cas échéant – la nature du lien de copolysémie qui la relie à un ou plusieurs copolysèmes du même vocable), par les caractéristiques grammaticales [CG], flexionnelles [MO] jusqu'aux exemples [EX]. La description sémantique dans la rubrique [DF] comprend une étiquette sémantique (cf. Polguère, 2011), une spécification d'actants sémantiques dans la forme propositionnelle et la définition structurée en composante centrale et en composantes périphériques. Enfin, les relations sémantico-syntaxiques, c'est-à-dire les fonctions lexicales, qui structurent le système lexical et qui en forme l'ossature, sont visibles dans la rubrique [FL]. Les exemples illustrent le sens et l'emploi de la lexie vedette, alors que sous [PH] sont listés – pour chaque unité lexémique – les différents types de phrasèmes dans lesquels celle-ci se trouve incluse.

Un système lexical est de nature relativiste, ce qui permet de rendre compte d'un certain flou dans l'information dont dispose le locuteur et – souvent – le lexicographe. C'est la raison pour laquelle chaque élément de la description sémantico-syntaxique est identifié avec un taux de probabilité qui va de 60 % à 100 %, lorsqu'il s'agit d'un travail effectué « manuellement »⁴⁸.

Un système lexical est une modélisation du lexique d'une langue donnée : il doit rendre compte de sa structure et de son fonctionnement, sans s'orienter d'office vers des correspondances interlinguistiques. Ainsi, le *RL-fr* et *en-LN* sont implémentés de manière autonome, en tant que modèles monolingues, leur interconnexion étant cependant envisageable.

3. STATUT LINGUISTIQUE DE LA LOCUTION ET CONSTRUCTION DU NŒUD LEXICAL

Les exemples (1) et (2) présentés dans l'Introduction suggèrent que la traduction se déroule sans encombre, lorsque le phrasème est un syntagme gouverné par le verbe *être*, alors que d'autres cotextes – les verbes attributifs (*sembler*, *paraître*, *avoir l'air*⁴⁹) et le causatif (*mettre*) – posent des difficultés⁵⁰. *Septième ciel* apparaît dans des configurations diverses : l'individu X peut, certes, *être au septième ciel*, mais il n'est pas rare de trouver également *vivre au septième ciel* ou *atteindre le septième ciel*.

Faut-il dès lors admettre plusieurs locutions verbales, plus ou moins synonymes, contenant la séquence *aux anges* et *septième ciel*, en implémentant autant de nœuds lexicaux du graphe ? Pour le vérifier, nous avons observé, dans la section 3.1, le degré d'autonomie qui caractérise les syntagmes prépositionnels *aux anges* et *au septième ciel*, en nous intéressant aux emplois sans le verbe *être* et sans la préposition *à*. Dans 3.2, nous procédons de façon analogue pour examiner deux de leurs équivalents anglais *to be over the moon* et *to be in seventh heaven*.

⁴⁸ Toute information injectée automatiquement se voit associer un taux de probabilité de 50 %.

⁴⁹ Notons que *avoir l'air* est une collocation dont *air* est la base et *avoir* le verbe support collocatif (cf. *Elle a l'air aux anges* -> *Son air aux anges*).

⁵⁰ Globalement, les résultats obtenus par Google translator sont meilleurs que ceux retournés par Reverso, dans la mesure où le premier de ces outils propose des traductions adéquates, mais qui ont perdu l'idiomaticité de l'original.

3.1 (Être) aux anges et (être) (au) septième ciel dans le RL-fr

3.1.1 Locution prépositionnelle aux anges

Le syntagme prépositionnel *aux anges* apparaît fréquemment avec des verbes attributifs (exemple 3 ci-dessous) ou bien rattaché directement à un gouverneur nominal, sans support verbal (exemple 4). Son fonctionnement syntaxique est ainsi caractéristique des adjectifs.

(3) Le couple semblait aux anges mais une consigne très précise avait été donnée à tous les journalistes présents : aucune question sur les rumeurs de grossesse. (FrWac)

(4) Assez fière du repas d'anniversaire 2005, avec des petits plats équilibrés et une famille aux anges, et surprise. (FrWac)

De même, il s'emploie régulièrement en position détachée à gauche (exemple 5), comme à droite (exemple 6).

(5) Aux anges, Géraldine berçait ce poupon mou de paroles encourageantes, expliquait comme il est plus avantageux, question vanité, d'être un peu plaqué que très cocu. (Frantext)

(6) Lambert écoutait, aux anges. (Frantext)

D'autres propriétés transformationnelles mettent en évidence la nature adjectivale de ce syntagme. Il peut en effet être modifié par des adverbes (7) et s'inscrire dans des coordinations avec d'autres adjectifs (8).

(7) « Ma petite fille Azura Sienna est née la semaine dernière ! Je suis euphorique et complètement aux anges », a-t-elle posté sur le réseau social. (Internet)

(8) Amélie N. (Hollywood Girls 4), amoureuse et aux anges, elle dévoile sa nouvelle conquête. (Internet)

Plutôt que de traiter *être aux anges* comme une locution verbale, il convient d'y voir une collocation construite sur la base *aux anges*, syntagme prépositionnel qui forme une locution à caractère adjectival. L'identification de la partie du discours permet de situer la locution dans un ensemble de relations paradigmatiques⁵¹. *Aux anges* est lié par des liens de quasi-synonymie avec, entre autres, les lexèmes *ravi*, *enchanté*, *heureux*, ainsi qu'à *sur un nuage*. *Aux anges* peut en outre commuter avec ses quasi-synonymes, en partageant leurs propriétés distributionnelles : tout comme ces adjectifs, il forme des collocations avec *être* et avec des verbes attributifs. Si cette locution s'emploie avec le verbe *se sentir*, son causatif (cf. exemple 2a dans l'Introduction) n'est pas celui que l'on utilise avec les autres adjectifs de sentiments. Les collocations formées avec *mettre* conduisent à penser que le sens de *aux anges* est mieux caractérisé par l'étiquette sémantique 'qui est dans un certain état psychique' que par 'qui éprouve un sentiment'. La figure 4 visualise l'article lexicographique de notre locution, *i.e.* le nœud lexical qu'il forme.

⁵¹ On notera à ce propos une différence notable qui oppose le traitement des locutions dans un système lexical à celui que propose WordNet (cf. Osherson et Fellbaum, 2010).

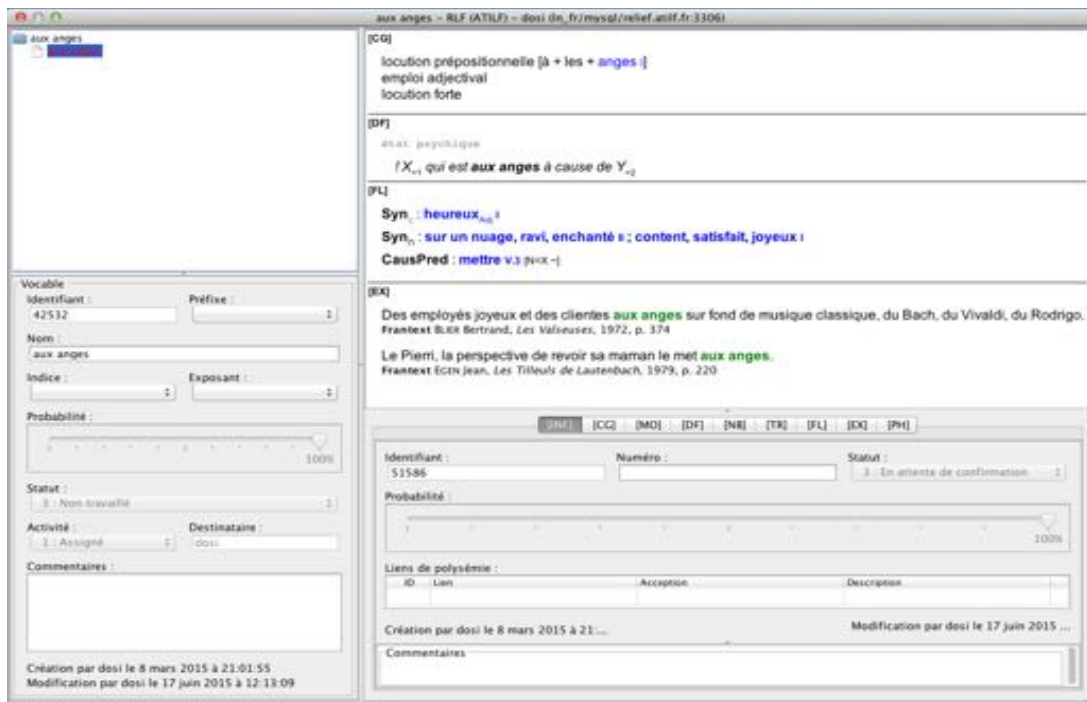


Figure 7. Locution *aux anges* dans la base *RL-fr* (vue d'article).

3.1.2 (*Être*) (*au*) *septième ciel* : locution nominale enchâssée

Le jeu de parenthèses dans l'intitulé de cette sous-section renvoie aux trois types d'emplois attestés dans les corpus. *Septième ciel* s'utilise en effet en syntagme verbal avec le verbe *être* ou bien avec des verbes attributifs tels que *paraître* dans (9).

(9) Alors que M^{lle} Lavoix souriait, paraissait au septième ciel, voulant ajouter à son ravissement, j'employai quelques épithètes flatteuses. (Sabatier, R. 1997. *Le Lit de la merveille*)

Le syntagme prépositionnel *au septième ciel* fonctionne de façon adjectivale. On le retrouve en position détachée à gauche (10) et à droite (11), ainsi qu'en coordination avec des adjectifs : dans (12), nous l'avons substitué à *aux anges* de l'exemple (8). En revanche, il se montre réfractaire à la modification adverbiale : nous n'en avons trouvé aucune attestation et l'insertion d'un modificateur dans l'exemple (9) renuméroté en (13) aboutit à une phrase difficilement acceptable.

(10) Au septième ciel, ma fille, ravie, elle veut se précipiter. (Frantext)

(11) Ils regardaient ses gestes fiévreux, ses yeux luisants, sa pose... trop penché en avant... perdant la tête, oubliant où il est, se croyant seul avec elle dans les limbes, au septième ciel... (Frantext)

(12) Amélie N., amoureuse et au septième ciel, elle dévoile sa nouvelle conquête.

(13) ? Alors que M^{lle} Lavoix souriait, paraissait complètement au septième ciel, voulant ajouter à son ravissement, j'employai quelques épithètes flatteuses.

Contrairement à *aux anges* cependant, les corpus attestent de nombreux emplois purement nominaux de *septième ciel*. Les exemples (14) et (15) en offrent une illustration :

(14) Pour rejoindre votre septième ciel, *Épices & Vous* vous propose de faire entrer l'Asie dans votre cuisine... (FrWac)

(15) Les partiels, c'est le septième ciel. (Internet)

La structure interne du nœud que forme la locution *septième ciel* sera donc celle d'une lexie nominale, insérée dans un réseau de relations de quasi-synonymie avec *bonheur*, *plaisir*, *bien-être*, *contentement*, *ravissement*, *enchantement*, *félicité*, etc. Notons cependant que, contrairement à la plupart de ses quasi-synonymes, *septième ciel* n'est pas un sentiment à proprement parler : s'il est possible de parler de *sentiment de bonheur*, de *sentiment de plaisir* ou de *sentiment de contentement*, *sentiment de septième ciel* semble difficile à accepter. Le sens de cette locution est à classer plutôt parmi les états psychiques.

Le syntagme verbal *être au septième ciel* est ainsi un exemple de double enchâssement. Pour former une collocation à valeur adjectivale, *septième ciel* sélectionne la préposition *à*. Or, à l'image des adjectifs simples, les collocations à valeur adjectivale telles que *au septième ciel*, *au désespoir de*, *en colère* se combinent avec *être* et avec les verbes attributifs. Dans la mesure où il s'agit d'une caractéristique commune aux adjectifs, aucune mention n'en est fait dans l'entrée lexicographique du nœud lexical.

La figure (6) représente le nœud *septième ciel* dans le système lexical *RL-Fr*.

The screenshot displays the 'septième ciel' entry in the RL-Fr system. On the left, a tree view shows the hierarchy: 'septième ciel' (selected), 'septième ciel 1', 'septième ciel 1.1', and 'septième ciel 1.2'. The main area is divided into several sections:

- [CG]** (Classification Grammaticale): locution nominale, masc, locution forte, déf sg.
- [DP]** (État psychologique): septième ciel de X₁, causé par Y₂.
- [PL]** (Synonymes): **Syn.**: plaisir (1), bien-être; satisfaction 2, contentement < félicité, joie, bonheur. **A.**: à (1) [e-].
- [EX]** (Exemples): Paquets, valises et nous, on se serre dans l'ascenseur, on monte dans l'immeuble. À l'extérieur, un peu H.L.M. On grimpe, on ouvre. Dedans, soudain, le luxe. Au septième ciel, ma fille, ravie, elle veut se précipiter. Frantext Dolakovsky Serge, *Le Livre d'été*, 1989, p. 305. En dehors de mes envolées vers le septième ciel, j'ai toujours eu les pieds sur la terre au sujet de Florian. Frantext Dorin Françoise, *Les Vendanges carives*, 1997, p. 71.
- Navigation**: [DNI] [CG] [MO] [DP] [NB] [TR] [PL] [EX] [PH]
- Identifiant**: 51587, **Numéro**: (H.1), **Statut**: 3 - En attente de confirmation.
- Probabilité**: 100%.
- Liens de polysémie**:

| ID | Lien | Acceptation | Description |
|------|-----------|-----------------|-------------|
| 7507 | métaphore | septième ciel 1 | comme si |
- Création par dossi** le 8 mars 2015 à 21:12:50. **Modification par dossi** le 18 juin 2015 à 13:06:35.

Figure 8. Locution *septième ciel* dans la base *RL-Fr* (vue d'article).

3.1.3 Aux anges et septième ciel dans le système lexical du français

Le sous-graphe du système lexical français organisé autour des nœuds locutionnels *aux anges* et *septième ciel* (figure 6) montre que leur interconnexion est indirecte : elle se fait par un lien de quasi-synonymie avec le nom *contentement*, lié sur le plan paradigmatique à la fois à *septième ciel* et à l'adjectif *content*, quasi-synonyme de *aux anges*.

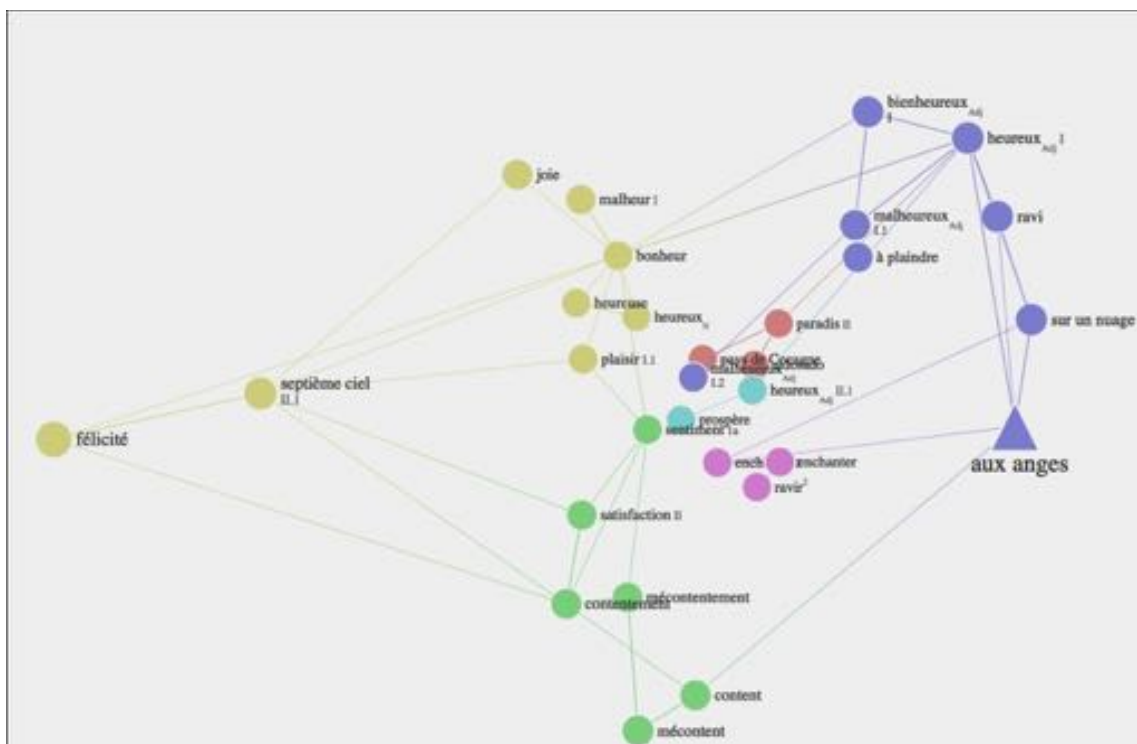


Figure 9. Interconnexion des nœuds lexicaux *aux anges* et *septième ciel* dans *RL-Fr*.

3.2 (To be) over the moon et (to be) (in) seventh heaven dans en-LN

3.2.1 (To be) over the moon dans le système lexical en-LN

Le traitement de *(to be) over the moon* dans les ressources lexicales de l'anglais n'est pas homogène. Certaines d'entre elles⁵² lui accordent un statut de locution verbale, syntagme gouverné par *to be*, alors que d'autres⁵³ décrivent le phrasème *over the moon*, c'est-à-dire un syntagme prépositionnel de forme *Prep Det N*.

Comme *(to be) over the moon* ne figure pas dans WordNet, cette locution n'a pas été injectée automatiquement dans *en-LN*. Le nœud lexical a été créé et situé dans le réseau par les lexicographes au terme d'une analyse linguistique dont l'objectif premier était de déterminer le statut lexical – locutionnel ou collocationnel – du phrasème.

En effet, si *over the moon* apparaît fréquemment en position de complément du verbe *to be*, il n'en manifeste pas moins de propriétés adjectivales. Dans la phrase (16), il se rattache directement à son support nominal, alors que (17) et (18) illustrent des emplois détachés, respectivement à gauche et à droite. Malgré une ponctuation qui en fait un énoncé à part dans (18), *over the moon* y qualifie anaphoriquement l'antécédent *mom* de la phrase précédente.

(16) Gloria Taylor's mom over the moon (Internet).

⁵² Voir, à titre d'exemple, Cambridge Dictionaries Online (dictionary.cambridge.org/dictionary/british/be-over-the-moon), Reverso, etc. Cette locution n'est pas intégrée dans WordNet, ni en tant qu'une forme verbale, ni en tant qu'une forme adjectivale).

⁵³ C'est le cas du dictionnaire Collins (www.collinsdictionary.com/dictionar/english).

(17) Although over the moon with it generally, I am disappointed with the lack of power in the naturally aspirated diesel engine. (COCA)

(18) He's fast asleep, so is mum! Over the moon. (Internet)

La possibilité de modification adverbiale dont témoigne l'exemple (19) et de coordination avec d'autres adjectifs (*humbled*, *happy* dans 20) pointe également vers un statut adjectival de la locution prépositionnelle *over the moon*.

(19) Absolutely over the moon with this job! (BNC)

(20) As the first from Great Britain to win a world title, I'm just humbled and happy, and over the moon. (Internet)

Le nœud *over the moon* du réseau *en-LN* (figure 7 ci-dessous) contient donc un ensemble d'informations propres aux adjectifs. La possibilité de sélectionner le collocatif *to be* fait partie des propriétés des lexies adjectivales, on peut donc se demander s'il est indispensable de l'indiquer dans la zone de fonctions lexicales. Une telle démarche semble en effet superflue dans le cas des nœuds monolexémiques quasi-synonymes *ecstatic*, *delighted*, *happy*, *thrilled*, *overjoyed* : en tant qu'adjectifs, ils s'emploient régulièrement avec le collocatif *to be*, tout comme les locutions *on the top of the world*, *on cloud nine*.

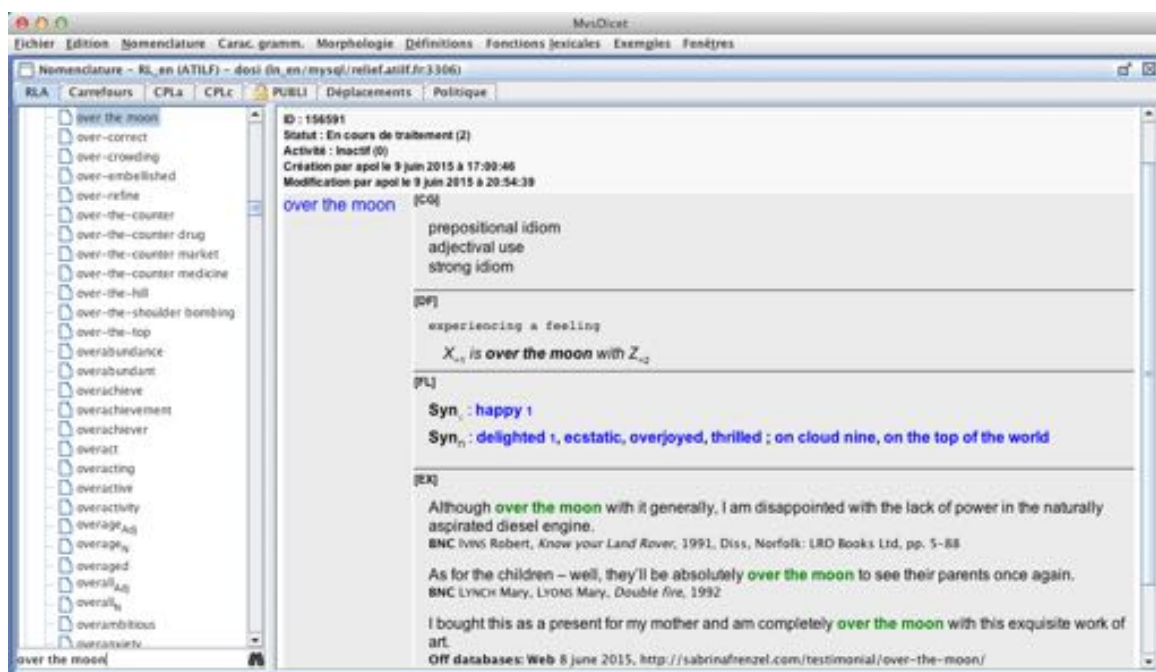


Figure 10. Locution *over the moon* dans la base *en-LN* (vue d'article).

3.2.2 (*To be*) (*in*) *seventh heaven* : une locution nominale enchâssée

Conformément à l'analyse effectuée par les lexicographes de WordNet, *seventh heaven* est une locution nominale et c'est en tant que telle qu'elle a été injectée dans le *en-LN*. À la conversion, elle s'est trouvée extraite de son synset (cf. figure 2 dans la section 2) pour former un nœud dans le système lexical de l'anglais. Le résultat de cette opération a

néanmoins nécessité une vérification manuelle permettant de s'assurer qu'il s'agit bien d'un nœud nominal.

Les corpus attestent des emplois sans le verbe *to be*, tels que l'exemple (21).

(21) It's funny to see a guy in seventh heaven because of a girl. (Internet)

Les phrases (22) et (23) montrent cependant que la préposition *in*, qui autorise un emploi adjectival, n'est nullement obligatoire. Les collocatifs *some form of ~* et *my state of ~* indiquent qu'en tant que locution nominale, *seventh heaven* dénote une entité : un état psychique qui caractérise l'individu X (instancié dans l'exemple 23 par le possessif *my*) en réaction à une situation Y.

(22) The more he wished time away, the more slowly it passed although when, at last, the ship returned to Saint-Iean-de-Luz, life assumed the aura of some form of seventh heaven. (Pike, Richard, 2006, *Seven Seas, Nine Lives. A Biography of Captain A. W. F. Sutton*, p. 144)

(23) Sarah shook her head ruefully. 'In my state of seventh heaven it's all going over the top of my head.' She sobered suddenly. 'I'm afraid I was a bit of fool.' (Armstrong, Lindsay. 2011. *Accidental Nanny*, ch. 6)

Notre analyse s'accorde donc avec celle de WordNet. La figure 8 montre le nœud lexical *seventh heaven* dans le réseau *en-LN* après un traitement lexicographique. Le lien syntagmatique vers la préposition *in* permettant de former une collocation à valeur adjectivale a été encodé dans la zone FL. Remarquons que tout comme son équivalent français, cette collocation se trouve régulièrement enchâssée dans celle qu'elle forme avec le verbe *to be*.

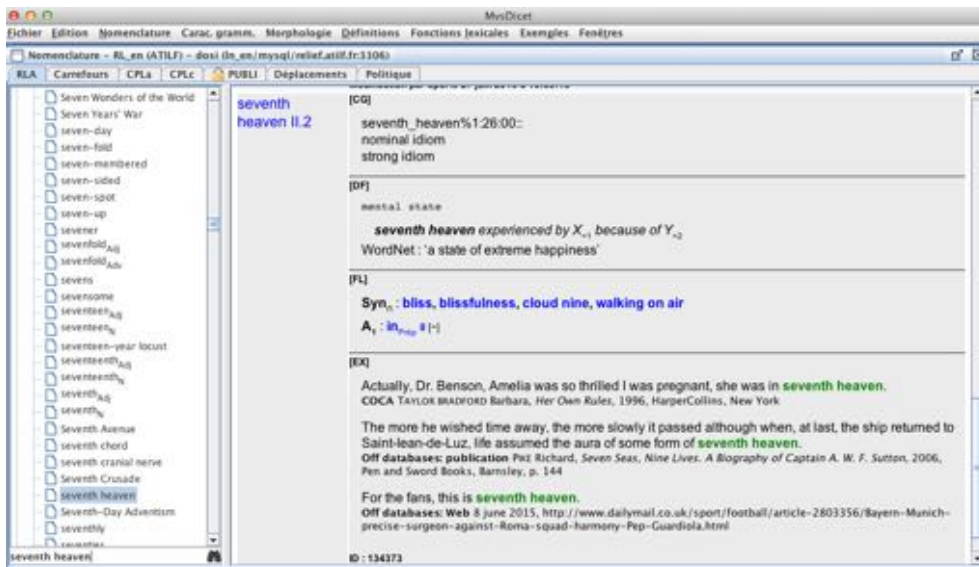


Figure 11. Locution *seventh heaven* dans la base *en-LN* (vue d'article).

4. CONCLUSIONS

Les quatre locutions dont cet article détaille l'analyse nous ont permis d'illustrer le traitement des UPs dans deux ressources lexicales de type dictionnaires virtuels. En identifiant le statut grammatical de chaque phrasème dans le cadre strictement défini d'un même modèle lexical, on vise à départager ce que celui-ci contient de locutionnel et ce qui relève de la collocation. Les liens syntagmatiques et pourront ensuite être tissés pour d'une part, situer le nœud dans le lexique et d'autre part, fournir aux utilisateurs humains et aux logiciels les informations nécessaires pour en maîtriser la combinatoire syntaxique.

Une fois qu'un nœud est inséré dans un système lexical construit de manière autonome, deux modes d'interconnexions interlinguistiques sont proposés par Polguère (2009). Premièrement, des nœuds peuvent être connectés directement entre eux : on établira ainsi un lien d'équivalence entre *pluie* I, nœud du système lexical *RL-fr* et *rain* 1 dans *en-LN*, dénotant tous les deux le même phénomène atmosphérique. Un deuxième mode d'interconnexion s'applique aux collocatifs des lexies équivalentes. La connexion s'effectue par l'intermédiaire de relations sémantico-lexicales (FL), en aboutissant à des équivalences que l'on pourrait qualifier de locales. Ainsi, l'application de la FL Magn (intensificateur) aux arguments *pluie* I en français et *rain* 1 en anglais retourne respectivement les valeurs telles que *violent(e)* II, *diluvien(ne)* en français, *heavy* 2 en anglais. Or, il est clair que des lexies telles que *diluvien(ne)* et *heavy* 2 ne peuvent être reliées qu'à travers la relation sémantico-lexicale qui se construit sur le plan collocationnel avec leurs bases respectives.

Le mode d'encodage et l'organisation des informations proposés dans le modèle de systèmes lexicaux permettra, selon notre hypothèse, de limiter, voire d'éviter certaines erreurs de traduction, notamment celles liées aux collocatifs non identifiés comme tels. Si les deux ressources sont encore trop jeunes pour des tests effectués sur l'ensemble de leurs réseaux, leurs structures autorisent des vérifications locales à partir des sous-graphes que forment des lexies bien décrites des champs lexicaux sur lesquels le travail de description a déjà été effectué, par exemple celui des sentiments.

Bibliographie :

- GADER, N., LUX-POGODALLA, V., AND POLGUERE, A., 2012. Hand-Crafting a Lexical Network With a Knowledge-Based Graph Editor. *Proceedings of the Third Workshop on Cognitive Aspects of the Lexicon (CogALex III)*, The COLING 2012 Organizing Committee, Mumbai, pp. 109–125.
- GADER, N., OLLINGER, S., AND POLGUERE, A. 2014. One Lexicon, Two Structures: So What Gives? In: H. ORAV, Ch. FELLBAUM and P. VOSSEN, eds 2014. *Proceedings of the Seventh Global Wordnet Conference (GWC2014)*. Tartu (Estonie), Global WordNet Association, pp. 163–171.
- FELLBAUM, CH. 2014. Large-Scale Lexicography in Digital Age. *International Journal of Lexicography*, 27/4, pp. 378–395.
- LUX-POGODALLA, V. AND POLGUERE, A. 2011. Construction of a French Lexical Network. Methodological Issues. *Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011. An ESSLLI 2011 Workshop*, Ljubljana, pp. 54–61.

- MCCARTHY, M. 1999. What constitutes a basic vocabulary in spoken communication? *Studies in English Language and Literature*, 1, 233-249.
- MEL'CUK, I., 2013. Tous ce que nous voulions savoir sur les phrasèmes, mais... *Cahiers de Lexicologie*, 102, 1, pp. 129-149.
- MEL'CUK, I., CLAS, A. AND POLGUERE, A. 1995. *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve: Duculot.
- OSHERSON, A. AND FELLBAUM, Ch. 2010. The Representation of Idioms in WordNet. *Proceedings of Global WordNet Conference 2002, CFILT, IIT, Bombay, Mumbai, 2010*. < [http://www.cfilt.iitb.ac.in/gwc2010/pdfs/16 Representation of Idioms in WordNet Osherson.pdf](http://www.cfilt.iitb.ac.in/gwc2010/pdfs/16%20Representation%20of%20Idioms%20in%20WordNet%20Osherson.pdf)> [Accessed 4 May 2015]
- POLGUERE, A., 2009. Lexical systems: graph models of natural language lexicons. *Language Resources and Evaluation*, 43, pp. 41-55.
- POLGUERE, A. 2011. Classification sémantique des lexies fondée sur le paraphrasage. *Cahiers de Lexicologie*, 98-1, pp. 197-211.
- POLGUERE, A. 2014. From Writing Dictionaries to Weaving Lexical Networks. *International Journal of Lexicography*, 27/4, pp. 396-418.
- POLGUERE, A. AND SIKORA, D. 2013. Modèle lexicographique de croissance du vocabulaire fondé sur un processus aléatoire, mais systématique. In: C. Masseron, C. Garcia-Debanç, Ch. Ronveaux eds, 2013. *Enseigner le lexique. Pratiques sociales, objets à enseigner et pratiques d'enseignement*. Laval (Québec): Presses universitaires de Laval, pp. 35-63.
- SIKORA, D. 2014. Brève histoire d'une erreur lexicale. Polysémie et liens lexicaux dans l'enseignement du vocabulaire. In: F. NEVEU, P. BLUMENTHAL, L. HRIBA, A. GERSTENBERG, J. MEINSCHAEFER and S. PREVOST, eds 2014. *4^e Congrès Mondial de Linguistique Française – CMLF 2014*, Paris: Institut de Linguistique Française, pp. 1157-1172.
- SVENSEN, B., 2009. *A Handbook of Lexicography: The Theory and Practice of Dictionary-making*. Cambridge, New York: Cambridge University Press.
- VAGUER, C., 2010. *Être aux anges, sortir de ses gonds...* Comment les langues traduisent-elles des états émotionnels ? *Cahiers Sens public*, 1, n° 13-14, pp. 253-269.

Websites

- BNC, BRITISH NATIONAL CORPUS. [online] Available at <http://www.natcorp.ox.ac.uk/>
- CAMBRIDGE DICTIONARIES ONLINE. [online] Available at <http://dictionary.cambridge.org>
- COCA, 1990-2012, *Corpus of Contemporary American English*. [online] Available at <http://corpus.byu.edu/coca/>
- Collins. [online] Available at < www.collinsdictionary.com/dictionar/english>
- FRANTEXT, 2014, *Base textuelle Frantext*. [online] Available at www.frantext.fr

FRWAC, 2008, *Corpus of French Web*. [online] Available at

http://nl.ijs.si/noske/wacs.cgi/first_form?corpname=frwac;lemma=;pos=

GOOGLE TRANSDUCTION, 2015. [online] Available at <https://translate.google.fr/?hl=fr>

REVERSO, 2015. [online] Available at www.reverso.net

WORDNET, 2015, *A lexical database for English*. [online] Available at

<https://wordnet.princeton.edu/>

Polysémie et homonymie des locutions dans un modèle lexical

Pausé, Marie-Sophie & Sikora, Dorota

Université de Lorraine & ATILF-CNRS, Université du Littoral Côte d'Opale & HLLI
marie-sophie.pause@univ-lorraine.fr
dorota.sikora@univ-littoral.fr

Résumé. Si les unités phraséologiques, telles que les locutions qui nous intéressent dans cet article, sont aujourd'hui de plus en plus régulièrement examinées dans les ressources lexicographiques (cf. WordNet[1], FrameNet[2] à titre d'exemple), leur traitement et la description qui en est proposée prennent rarement en compte de manière systématique leur polysémie. La base de données lexicales *Réseau Lexical du Français (RL-fr)*, développée au laboratoire ATILF – CNRS (UMR 7118)[3], compte actuellement environ 3000 entrées locutionnelles. En nous appuyant sur les recherches menées dans le cadre théorique de la *Lexicologie Explicative et Combinatoire* (Mel'čuk et al., 1995) et sur la pratique du travail de description lexicographique liée au développement de la base *RL-fr*, nous nous intéressons aux spécificités de la polysémie des vocables locutionnels et à leur traitement dans une ressource dictionnaire à large couverture. Après une présentation générale du *RL-fr*, nous discutons l'analyse et l'encodage des locutions à sens multiples. En effet, pour un vocable locutionnel tel que marcher sur la tête qui nous sert d'exemple, trois sens sont à considérer : 'dominer', 'remporter une victoire écrasante' et 'se comporter de manière déraisonnable'. L'analyse consiste alors à déterminer si nous sommes face à un seul vocable polysémique ou bien s'il s'agit d'homonymes, et par conséquent de deux, voire trois entrées de dictionnaire distinctes. Le modèle d'analyse développé et pratiqué dans le cadre du *RL-fr* prend en compte aussi bien les liens sémantiques entre les (éventuelles) acceptions copolysèmes d'une locution, que les aspects formels de son signifiant, à travers une description de sa structure lexico-syntaxique. Cette dernière s'avère plus d'une fois décisive, lorsqu'il faut juger s'il est question de polysémie ou d'homonymie. Ainsi, l'analyse de la structure lexico-syntaxique de marcher sur la tête nous amène à constater qu'il existe un vocable polysémique marcher sur la tête1 regroupant les acceptions 'dominer' et 'infliger une grosse défaite' d'une part, et de l'autre son homonyme marcher sur la tête2 avec un seul sens, celui de 'se comporter de manière déraisonnable'. En conclusion, nous considérons qu'une description de la structure lexico-syntaxique est fondamentale, si l'on tient à distinguer des cas de polysémie de ceux d'homonymie des locutions. Nous allons même jusqu'à dire que la structure lexico-syntaxique d'une locution constitue le signifiant de ces unités phraséologiques.

Abstract. The present paper focuses on a model of linguistic analysis that enables lexicographers to draw a clear distinction between polysemy and homonymy of idioms. Indeed, researchers' growing interest in phraseological units leads to give more room to idioms in lexicographic description (see for instance WordNet[1], FrameNet[2]). We therefore believe that, few dictionaries—if any—systematically account for their polysemy. The lexical database *Réseau Lexical du Français (French Lexical Network, henceforth RL-fr)* currently developed in the laboratory ATILF – CNRS (UMR 7118)[3] presently encodes about 3000 French idioms, with a systematic treatment of their polysemous structure. Building on this large-scale lexicographic

work based on theoretical et methodological principles of *Explanatory and Combinatorial Lexicography* (Mel'čuk et al. 1995), we examine the case of the verbal idiom *marcher sur la tête* (lit. 'to walk on the head'). It conveys three senses to be described in a dictionary: 'to dominate', 'to inflict a crushing defeat (in a competition)', and 'to behave irrationally'. Thus, the question that arises is whether there is one polysemous idiom *marcher sur la tête* or if we should rather consider it a case of homonymy. Semantic analysis can hardly establish any meaning relation between 'to dominate' and 'to inflict a crushing defeat' on the one side, and 'to behave irrationally' on the other. However, lexicographic treatment of idioms in *RL-fr* takes into account the signifié as well as the signifiant: the form of the later is modeled as its lexico-syntactic structure. As a matter of fact, *marcher sur la tête* reveals two different lexico-syntactic structures, corresponding respectively to the polysemous idiom *marcher sur la tête*¹ with two meanings ('to dominate', 'to inflict a crushing defeat') and its homonym *marcher sur la tête*²—'to behave irrationally'. These results lead to conclude that formal properties of idioms, modeled as their lexico-syntactic structure, should systematically be accounted for in lexicographic description. Moreover, given the available data, we suggest that lexico-syntactic structures of idioms correspond to their signifiants.

[1] <https://wordnet.princeton.edu/>

[2] <https://framenet.icsi.berkeley.edu/fndrupal/>

[3] www.atilf.fr

1 Introduction : questions de recherche et objectifs de l'article¹

La polysémie des locutions, syntagmes sémantiquement non compositionnels, est-elle différente de celle des vocables lexémiques ? Si oui, où résident les différences ? Dans les dictionnaires traditionnels, les locutions ne sont généralement pas considérées comme des unités lexicales à part entière et reçoivent une description hétérogène (Pruvost, 2011)². Seules des ressources récentes, adossées à des recherches linguistiques, telles que *WordNet* (Fellbaum, 1998, Osherson et al., 2009), *FrameNet* (Fillmore et al., 2003), et le *Réseau Lexical du Français* (voir *infra*) leur accordent ce statut. Malgré un intérêt grandissant des chercheurs pour la phraséologie, peu d'études abordent la polysémie des locutions, notamment en français. Notons qu'il existe un numéro de *Syntaxe et Sémantique* intitulé « Polysémie et Polylexicalité » (2004), mais les unités discutées sont majoritairement des mots composés, ou bien des unités telles que TELEPHONE PORTABLE, dont le statut lexical varie entre locution nominale et nom composé³.

S'appuyant sur les travaux et la pratique lexicographique liés au développement de la base de données *Réseau Lexical du Français* (*RL-fr*, cf. Gader et al., 2012, Polguère, 2014), cet article s'intéresse aux spécificités de la polysémie des vocables locutionnels et à son traitement lexicographique dans le *RL-fr*. La polysémie – on le sait – c'est le fait, pour un même signifiant, de recouvrir plusieurs signifiés sémantiquement liés. Le *RL-fr* permet de traiter et de décrire ces deux facettes – la forme et le sens – tout en prenant en compte les propriétés combinatoires des locutions.

Nous commencerons, dans la section 2, par une rapide présentation de la ressource et de la notion de *système lexical*, c'est-à-dire du modèle du lexique dont la ressource constitue l'implantation. La partie suivante sera consacrée à la discussion de l'analyse et de l'encodage de la polysémie des locutions. Nous nous arrêterons ensuite, en section 4, sur la façon dont le *RL-fr* traite leurs caractéristiques formelles, en associant à chacune d'entre elles une structure lexico-syntaxique, qui enrichit – d'une manière inédite à notre connaissance – la description lexicographique des unités polylexicales. On verra que la structure lexico-syntaxique d'une locution peut se montrer décisive pour distinguer entre la polysémie et l'homonymie. Nous étudierons le cas de MARCHER SUR LA TÊTE et des trois acceptions qu'on lui connaît : leur traitement requiert, d'une part, la prise en compte de relations sémantiques entre les signifiés, et de l'autre, l'analyse de la structure des signifiants. Nous concluons nos réflexions dans la cinquième et dernière section.

2 Présentation de la base lexicale Réseau Lexical du Français

Du point de vue de ses utilisateurs, la base de données *RL-fr* développée au sein du Laboratoire ATILF – CNRS (UMR 7118)⁴ est une ressource de type dictionnaire virtuel (Selva et al., 2003). Les informations qu'elle réunit peuvent être explorées avec des visées applicatives allant de l'enseignement de langues à la génération automatique de dictionnaires de spécialité. Dans une perspective plus large, le *RL-fr* est une implantation d'un modèle du lexique français⁵ élaboré selon les principes théoriques et méthodologiques homogènes, définis initialement dans le cadre de la Lexicologie Explicative et Combinatoire. Ce modèle a été récemment développé par Polguère (2009, 2014) sous le nom de *système lexical*.

Les propriétés formelles d'un système lexical ont été exposées et discutées entre autres dans Polguère (2014), puis dans Ollinger (2014). Faute de place, nous nous limiterons ici à relever celles d'entre elles qui sont particulièrement importantes pour l'intégration et la description de la polysémie et de l'homonymie dans le *RL-fr*.

Comme son nom l'indique, le *RL-fr* est structuré en réseau et forme un graphe au sens mathématique du terme : il est fait de sommets (nœuds) que constituent essentiellement des lexies (unités lexicales, qui sont soit des lexèmes, soit des locutions) et des arcs représentant des relations lexicales de différente nature (cf. *infra*). Précisons qu'il s'agit d'un modèle non-ontologique, puisque tous les liens ont un statut égal. C'est dire que contrairement à une ressource telle que *WordNet*, le *RL-fr* ne privilégie d'aucune façon les relations qui renvoient à une organisation conceptuelle du lexique (cf. Polguère, 2014).

La figure 1 présente le graphe *RL-fr* vu « de l'extérieur », une photo d'identité en quelque sorte, qui permet de remarquer ses traits caractéristiques, c'est-à-dire ses propriétés structurelles essentielles.

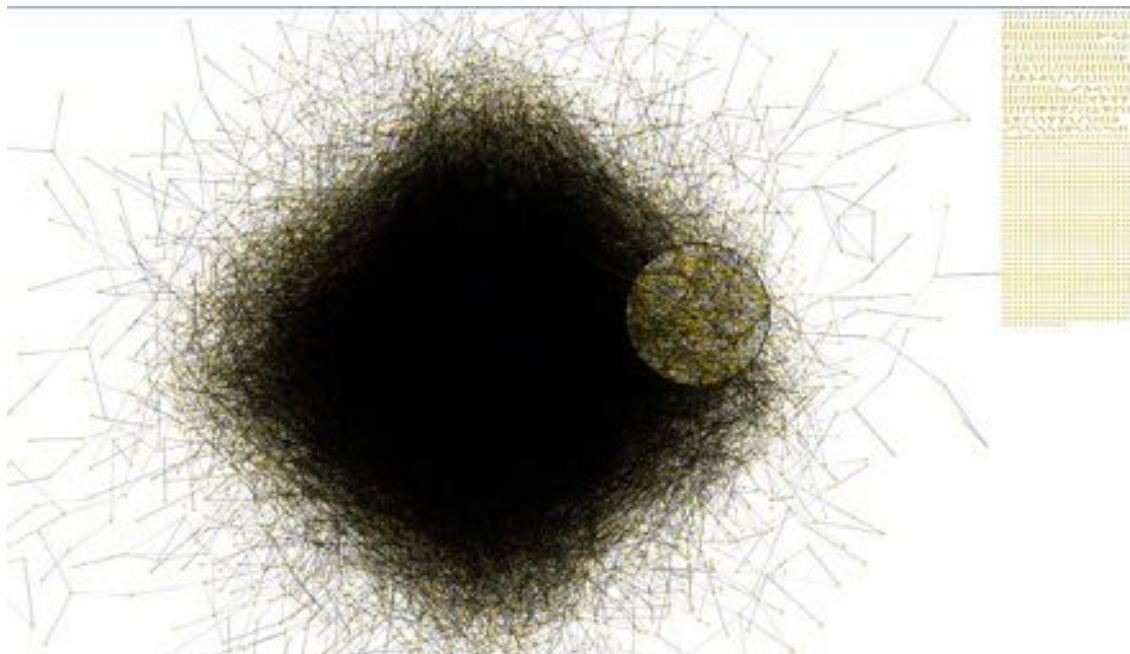


Figure 1 : Graphe *RL-fr* (état de développement du 23 novembre 2015).

Dans la mesure où il s'agit d'une ressource en cours de développement, la structure du graphe est en constante évolution. Au fur et à mesure que le travail des lexicographes progresse, de nouveaux sommets

Les nœuds sont insérés et un réseau de liens lexicaux de plus en plus dense se constitue, en rattachant les lexies faiblement, voire pas du tout connectées, à l'ensemble du graphe. À terme, chaque sommet sera rattaché à au moins un autre nœud, dans une structure de type graphe petits mondes. Les propriétés formelles des graphes petits mondes ont fait objet de nombreuses publications récentes (Gaume 2004, et surtout Ollinger 2014 pour une discussion relative au *RL-fr*). Notons simplement avec Kiss (1968) que leur organisation est à l'image de celle du lexique d'une langue. Ainsi, malgré une connectivité généralement faible de un ou deux arcs pour la plupart des sommets, il est possible de naviguer entre eux, en passant d'un nœud vers n'importe quel autre en un nombre de pas peu élevé.

Les sommets encapsulent une description lexicographique comprenant les caractéristiques grammaticales de la lexie-vedette, des informations relatives à son sens (spécification de ses actants pour les unités à caractère prédicatif, étiquette sémantique⁶ et, pour certaines, la définition⁷), les relations lexicales dont elle est source et cible, ainsi que des citations illustrant ses emplois. Au lieu de former un bloc, l'entrée lexicale est divisée en zones allouées à chaque type d'informations. Cette organisation non-atomique des données, commune à tous les sommets qu'ils soient lexémiques ou locutionnels (cf. figures 2 et 3 ci-dessous) permet une exploration « thématique » de la base.

Figure 2 : Informations stockées dans un nœud lexémique (lexème VOITURE II).



Figure 3 : Informations stockées dans un nœud locutionnel (CHEVEUX D’ANGE II).

Les sommets sont reliés par des arcs qui, grâce au système de fonctions lexicales (Mel’čuk, 1996), modélisent les relations lexicales suivantes :

1. Relations paradigmatiques : comme toute autre lexie, une locution possède des synonymes, des antonymes et des hypero- ou des hyponymes. Ainsi, la locution nominale CHEVEUX D’ANGE II est reliée à son quasi-synonyme VERMICELLE (et *vice versa*) et à son hyperonyme PATE I.1. De plus, les relations actanciennes sont systématiquement encodées de manière à connecter une lexie telle que CARTON ROUGE I.1, ‘objet servant à transmettre une information’, aux noms de ses trois actants, respectivement : ARBITRE_{N,masc} et ARBITRE_{N,fém}, JOUEUR 2 et JOUEUSE 2, et FAUTE I.1b.
2. Relations syntagmatiques, qui pointent vers les collocatifs formant la combinatoire restreinte de la lexie vedette. Dans le cas de CARTON ROUGE I.1, il s’agira d’indiquer entre autres les verbes qui dénotent l’usage qui peut en être fait : SORTIR II.1a, BRANDIR, DEGAINER.

Un système lexical tel que le *RL-fr* permet de tisser tout un ensemble d’autres liens aussi bien entre des sens qu’entre des formes. Ainsi, les lexies-acceptations d’un vocable polysémique sont connectées entre elles par des liens de copolysémie. Le vocable en question forme un graphe, représentant l’hypothèse que fait le lexicographe sur son sens de base en synchronie et sur les ponts sémantiques (Mel’čuk et al., 1995) – condition *sine qua non* de toute polysémie – entre les lexies copolysèmes. La figure 4 illustre le cas de VOITURE, avec son réseau d’acceptations copolysèmes tel qu’il s’affiche dans le *RL-fr*. Nous y reviendrons plus amplement dans la section suivante, en en proposant une version enrichie d’exemples.

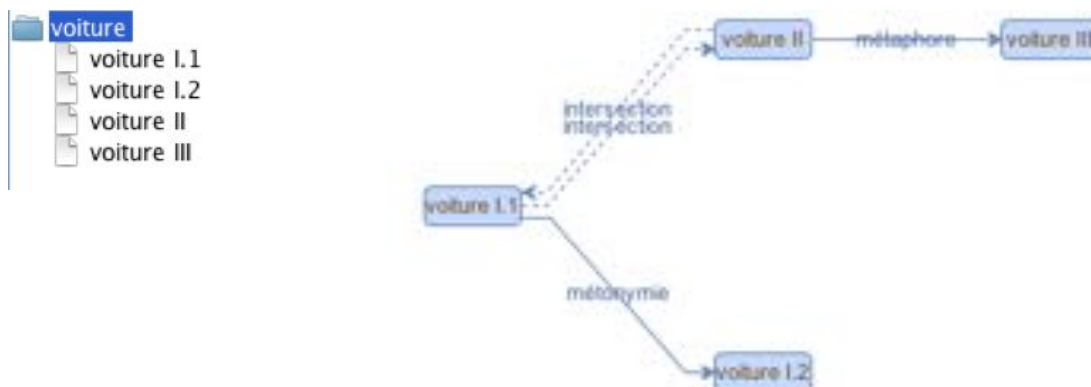


Figure 4 : Réseau des copolysèmes du vocable VOITURE.

Le travail d'élaboration des paraphrases définitoires lancé depuis peu conduit les lexicographes à tisser des liens d'inclusion définitionnelle ; à partir du *definiens* de la lexie CIGARETTE I, on atteindra les lexies essentielles pour la compréhension de ce sens : FUMER III.a, TABAC I.2, LEVRE I.1 et DOIGT I.1a (cf. figure 5).

[DF]
 chose fabriquée pour être fumée
cigarette destinée à être fumée par X_{-1}
 =
 objet I.1a destiné à être fumé III.a par l'individu X
 de forme allongée
 constitué d'un petit tube de papier très fin rempli de **tabac**¹ I.2 haché, que X fait se consumer par une de ses extrémités que X tient entre ses **lèvres** I.1 ou entre ses **doigts** I.1a quand X fume

Figure 5 : Encodage de la définition de CIGARETTE I.

De plus, quel que soit leur statut (locutions, clichés linguistiques, proverbes, etc.), les unités polylexémiques se rattachent au graphe *RL-fr* par des liens d'inclusion formelle qui leur sont spécifiques. À ce stade, nous ne faisons que les mentionner pour compléter la description générale de la ressource, en remettant leur présentation à la section 4.

3 Polysémie des locutions dans le Réseau Lexical du Français

Dans la présente section, nous exposons la façon dont le *RL-fr* décrit la pluralité des sens. En 3.1, nous présentons l'approche générale de la polysémie de la ressource. Nous observerons ensuite, en 3.2, les liens de copolysémie des locutions.

3.1 Traitement de la polysémie dans le Réseau Lexical du Français

Pour organiser la polysémie d'un vocable, le *RL-fr* adopte une approche dérivationnelle synchronique. Dans la figure 6 ci-dessous, nous reproduisons le réseau des copolysèmes du vocable VOITURE, en y intégrant des exemples illustrant chaque acception.

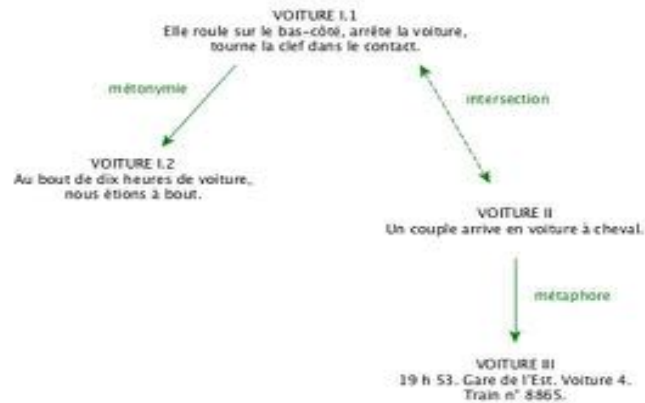


Figure 6 : Réseau des copolysèmes du vocable VOITURE (avec illustration de sens).

Si nous qualifions cette approche de dérivationnelle, c'est parce qu'elle consiste à organiser les acceptions à partir d'un sens de base (unité lexicale de base, désormais ULB). C'est l'ULB qui permet de dériver les autres copolysèmes, c'est-à-dire d'identifier et de tisser des liens sémantiques directs ou indirects entre les copolysèmes. La notion de *pont sémantique* proposée dans Mel'čuk, Clas et Polguère (1995 : 157) se trouve ainsi formellement encodée.

L'ULB d'un vocable est identifiée, nous le rappelons, dans une perspective synchronique : dans le cas d'un vocable tel que VOITURE, le graphe est organisé à partir du sens 'véhicule automobile', et non pas VOITURE II 'véhicule monté sur roues, tiré ou poussé par un animal, un homme'⁸, même si ce sens est antérieur d'un point de vue diachronique. On considère que, pour les locuteurs du français contemporain, l'ULB est à la fois le sens cognitivement le plus saillant parmi les copolysèmes (cf. Wilkins, 1996) et sémantiquement le plus simple à formuler. Cette saillance peut être facilement testée par une simple question adressée à un locuteur : *Qu'est-ce ce que veut dire VOITURE ?* Il y a fort à parier qu'en dehors d'un contexte spécifique, une réponse spontanée consistera à paraphraser le sens de VOITURE I.1, 'véhicule automobile' plutôt que celui de VOITURE II.

La métonymie qui relie l'acception VOITURE I.2, 'fait de se déplacer en voiture', à l'ULB et la relation métaphorique entre VOITURE II et VOITURE III sont des mécanismes bien connus dans le développement de la polysémie. Ils ne permettent cependant pas de modéliser la relation – clairement perceptible même pour un locuteur peu au fait des évolutions diachroniques – entre l'ULB et VOITURE II. Dans les deux cas, il s'agit de 'véhicule' ou 'moyen de transport' : cette relation sémantique particulière qui repose sur une partie commune aux deux sens est saisie en tant qu'intersection.

En terminant cette présentation très, et faute de place trop, rapide du traitement de la polysémie dans le *RL-fr*, notons que les ponts sémantiques (les arcs représentés en vert dans la figure 6) devront à terme être explicités dans les définitions des lexies copolysèmes qui en sont la source et la cible. Autrement dit, les définitions devront être formulées de manière à faire apparaître les relations de copolysémie. Il se peut donc qu'un graphe de copolysèmes dégagé à l'issue d'une première analyse soit remodelé lors de l'élaboration des paraphrases définitoires.

3.2 Liens de copolysémie des locutions

Dans l'état actuel du développement du *RL-fr*, la modélisation des relations de copolysémie autorise des comparaisons entre les structures de polysémie des vocables lexémiques et locutionnels. D'une part, il est possible d'observer certaines régularités, et de l'autre, de repérer nombre de difficultés liées

spécifiquement à la polysémie des locutions. Deux points, mis en évidence grâce à l'encodage systématique des liens de copolysémie, méritent d'être soulignés.

Premièrement, on note d'importantes différences entre le taux de polysémie des vocables lexémiques et celui des vocables locutionnels. Pour l'ensemble de la nomenclature, soit 26 540 vocables et 17 262 unités lexicales, ce taux se situe aux alentours de 1,54. Il est de 1,62 pour les vocables lexémiques et de 1,14 seulement en ce qui concerne ceux qui sont des locutions. De manière générale, même dans le cas de locutions relativement fréquentes, les réseaux de copolysèmes ont moins de sommets que ceux de vocables lexémiques fréquents. Leur taux de polysémie est généralement plus faible, dans la mesure où le nombre d'acceptions est rarement supérieur à trois. Les réseaux de copolysèmes lexémiques peuvent au contraire former des graphes atteignant, pour les plus fréquents d'entre eux, une vingtaine de sommets (cf. par exemple le verbe PASSER regroupant 19 copolysèmes⁹, et TETE avec 24 acceptions décrites à ce jour).

Deuxièmement, une comparaison des graphes que forment les acceptions copolysèmes des vocables lexémiques et locutionnels met en avant des similarités dans les structures dérivationnelles. Les arcs indiquant la nature des relations entre les copolysèmes montrent, en effet, que les mêmes mécanismes (métaphore, métonymie ou extension) sont en œuvre. Les locutions semblent donc soumises aux mêmes modes de fonctionnement cognitif que ceux qui caractérisent les lexèmes dans la plupart, sinon dans toutes les langues (cf. Sikora, 2014).

La figure 7 ci-dessous permet de comparer la polysémie de MARATHON et de COURSE D'OBSTACLES. Leurs graphes sont relativement simples, puisque ces vocables ne comptent que deux acceptions, dont la première est la source d'un développement métaphorique.



Figure 7 : Graphes de copolysèmes d'un vocable lexémique (MARATHON) et d'un vocable locutionnel (COURSE D'OBSTACLES).

Les liens de copolysémie permettent de traiter le versant sémantique de la polysémie, c'est-à-dire d'explicitier les relations de sens entre les signifiés d'un vocable lexémique ou locutionnel. Cependant, tout au long de notre travail, nous avons repéré des vocables locutionnels pour lesquels ce type de connexions est problématique à établir, tout comme il peut, par ailleurs, l'être pour certains lexèmes que l'on doit considérer dès lors comme homonymes. Il en est ainsi dans le cas de MARCHER SUR LA TETE illustré dans les exemples (1) à (3) ci-dessous. Si le lien métaphorique entre l'acception de la phrase (1), 'avoir un comportement dominant à l'égard de quelqu'un', et celle en (2), 'infliger une lourde défaite à quelqu'un', est facile à identifier, il n'en va pas de même pour MARCHER SUR LA TETE de l'exemple (3), que l'on pourrait paraphraser par 'avoir un comportement déraisonnable'.

1. La fille, surprise d'un tel débordement verbal, eut son premier sourire et murmura en s'agaçant l'oreille d'une tige de graminée : – Vous pourriez vivre avec une jeune fille qui vous **marcherait sur la tête** ? (Frantext)

2. L'équipe US version 2002 a été humiliée trois fois pour ne finir qu'à la sixième place, le plus mauvais rang des États-Unis en quatorze éditions. Dans l'ordre, l'Argentine en poule (87-80), la Yougoslavie en quart de finale (81-78) et l'Espagne en match de classement (81-75) lui **ont marché sur la tête**. (*L'Est Républicain* 10 septembre 2002)
3. Les hommes politiques **marchent sur la tête** ! Il faudrait qu'on m'explique comment on peut recapitaliser la BEI avec de l'argent qu'on n'a pas ? (Web)

Du côté du signifié, le pont sémantique entre la locution MARCHER SUR LA TÊTE dans (3) et celles des exemples (1) et (2) semble pour le moins problématique. Nous allons voir à présent que les liens d'inclusion formelle mis en évidence par la structure lexico-syntaxique des phrasèmes apportent des informations supplémentaires sur le signifiant, en permettant de trancher efficacement entre polysémie et homonymie.

4 Structures lexico-syntaxiques des locutions

Dans cette section, nous nous tournons vers le signifiant des locutions. En tant qu'unités polylexicales syntagmatiques, les locutions ont des constituants lexicaux, liés par des relations syntaxiques. Cette structure de leurs signifiants est également décrite dans le *RL-fr*. Chaque locution est alors liée à ses constituants lexicaux, et au patron syntaxique sur lequel elle est formée. Le résultat obtenu est une structure lexico-syntaxique (SLS). Dans la section 4.1, nous présentons rapidement l'intérêt qu'elle présente pour une analyse linguistique.

Précisons toutefois que l'identification des SLS n'est pas liée directement au traitement de la polysémie. Il n'en est pas moins que, chemin faisant, ce travail permet de porter un regard nouveau sur la polysémie des locutions. La section 4.2 discute des cas où le recours aux SLS conduit à une description plus précise de la pluralité des sens, notamment lorsqu'il fournit des moyens de décider s'il s'agit de polysémie (cf. MARCHER SUR LA TÊTE des exemples 1 et 2 de la section 3.2 ci-dessus) ou s'il faut opter pour l'homonymie.

4.1 Structures lexico-syntaxiques et liens d'inclusion formelle

L'objectif premier du travail de description lexico-syntaxique des locutions est de proposer modélisation, basé sur la syntaxe de dépendance, permettant d'expliquer les variations formelles des locutions (Pausé, à *paraître*). À terme, chaque locution sera, au sein de son article lexicographique, liée à un arbre de dépendance contenant des informations indiquant pourquoi certaines variations syntagmatiques se trouvent bloquées, alors que d'autres – passivation (4a), détachement (4b), extraction (4c), relativation (4d) et ajout d'un conjoint sur une composante autre que la tête du syntagme (4e) – sont autorisées (Pausé, *en préparation*).

4.
 - a. Bon, maintenant que **la glace est brisée** et que nous avons fait connaissance, [...] vous allez peut-être, monsieur Cidrolin qui semblez être du quartier, nous indiquer un autre endroit pas trop éloigné du centre de la ville, où mes chevaux [sic], ma fille, Empoigne, sa maman et moi pourrions combiner les joies du camping [sic] avec les plaisirs de la capitale. (Frantext)
 - b. Finalement, Régimbald a eu sa chance. **La glace, il l'a brisée** le 17 août dernier, et pour ajouter au stress d'un premier match dans la Ligue canadienne de football, c'est contre les Alouettes de Montréal qu'il a fait ses premiers pas. (Web)
 - c. Si on veut retrouver la confiance et le respect du peuple, **c'est devant sa porte qu'il faut** commencer à **balayer**. (Web)
 - d. Tant mieux pour Cavada, qu'il se prenne dans les gencives **la veste qu'il n'arrête pas de retourner** ! Ce type-là est à vomir ! (FrWac)
 - e. Monsieur le Président, Monsieur le Président du Conseil, vous devez **prendre par les cornes le taureau du doute sur Lisbonne** ; la France a déjà une longue

expérience dans l'art de tantôt faire avancer, tantôt faire reculer la construction européenne. (Web)

Une étude des variations paradigmatiques est également envisagée, dans le but de décrire systématiquement les commutations possibles (5) ou impossibles (6).

5.
 - a. [...] [L]e SLUC Nancy peut-il croire en ses chances de succès ce soir face à une équipe de Pau-Orthez taillée dans le roc et qui lui a fait **manger la poussière** à quatre reprises cette saison, championnat et coupe confondus ? (*L'Est Républicain* 28 mai 2002)
 - b. On veut une brute, un vrai leader qui fera **bouffer la poussière** à nos ennemis [...]. (Web)
6.
 - a. Comme beaucoup d'autres, mes parents ont adhéré très jeunes, dix-huit ans à peine, à un discours libérateur : ils en **ont payé les pots cassés**. Ils n'avaient pas fait d'études, ce n'étaient pas des intellectuels, ils avaient rompu avec leur famille... (Frantext)
 - b. *Ils en ont réglé les pots cassés.

À ce jour, plus de 2500 locutions françaises ont reçu une SLS dans le *RL-fr*. Les figures 8 et 9 montrent qu'une locution peut contenir soit uniquement des lexèmes, soit des lexèmes et des locutions.

The screenshot shows a software interface for analyzing the locution 'manger les pissenlits par la racine'. The main window displays the locution and its classification as 'locution verbale' and 'locution forte'. A detailed window titled 'locution verbale' shows the 'Structure syntaxique' as 'V Art NC Prép Art NC'. Below this, a table lists the components:

| Composant | Acception | Forme | Probabilité |
|-----------|-------------------------|------------|-------------|
| V | manger _v 113 | | 100 |
| Art | Art | les | 100 |
| NC | pissenlit | pissenlits | 100 |
| Prép | par III | par | 100 |
| Art | Art | la | 100 |
| NC | racine I | | 100 |

Figure 8 : SLS de la locution verbale MANGER LES PISSENLITS PAR LA RACINE 'être mort'.

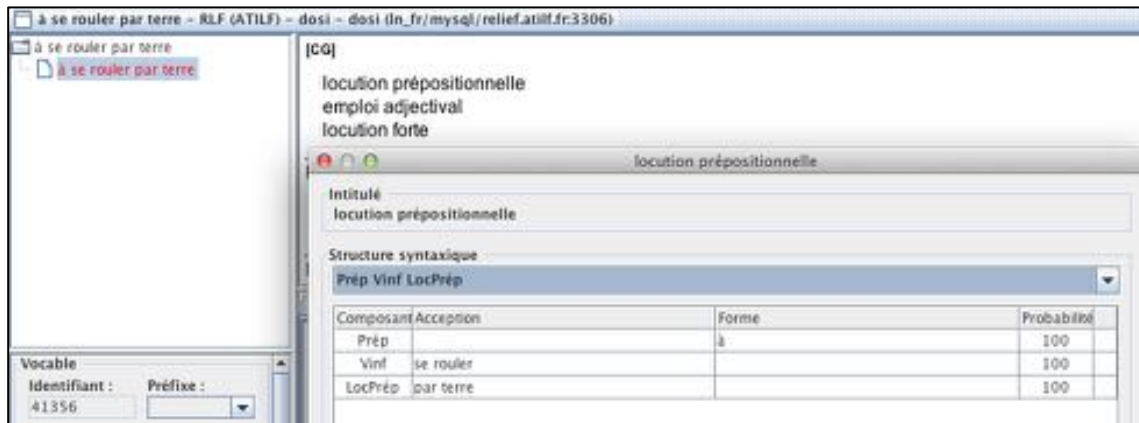


Figure 9 : SLS de la locution prépositionnelle A SE ROULER PAR TERRE¹⁰ ‘très drôle’.

La démarche d’identification des SLS permet de réunir des éléments expliquant leur comportement, tout en développant progressivement une base de patrons syntaxiques qui à terme devrait couvrir l’ensemble des locutions françaises.

L’intérêt d’identifier les constituants lexicaux des locutions est lié au fait que la combinatoire des constituants a un impact sur la combinatoire des locutions elles-mêmes. Nous savons par exemple que dans les syntagmes libres (exemples en 7), certaines lexies permettent la montée du possesseur, qui se manifeste à travers la pronominalisation du possesseur. C’est le cas des lexies désignant des parties du corps (7a) ou d’objets (7b), ou encore des vêtements (7c).

7.
 - a. Jeanne lui a cassé le bras, à Line.
 - b. Je lui ai abîmé une patte, à la chaise.
 - c. Je lui ai abîmé une manche, à ce gilet.

Cette montée du possesseur par un pronom personnel complément est également observable dans le cas de certaines locutions, par exemple pour MARCHER SUR LA TÊTE dans (1) et (2). À l’inverse, elle ne peut s’admettre pour MARCHER SUR LES TRACES ‘suivre l’exemple’ (8), puisque TRACE n’est pas liée à son actant par un lien de possession inaliénable. Elle est également impossible dans le cas de (3), étant donné que la locution a un unique actant qui est son sujet (ce cas est discuté plus précisément en section 4.2 *infra*).

- (1) La fille, surprise d’un tel débordement verbal, eut son premier sourire et murmura en s’agaçant l’oreille d’une tige de graminée : – Vous pourriez vivre avec une jeune fille qui **vous marcherait sur la tête** ?
 - (2) L’équipe US version 2002 a été humiliée trois fois pour ne finir qu’à la sixième place, le plus mauvais rang des États-Unis en quatorze éditions. Dans l’ordre, l’Argentine en poule (87-80), la Yougoslavie en quart de finale (81-78) et l’Espagne en match de classement (81-75) **lui ont marché sur la tête**.
 - (3) Les hommes politiques **marchent sur la tête** ! Il faudrait qu’on m’explique comment on peut recapitaliser la BEI avec de l’argent qu’on n’a pas ? (Web)
8.
 - a. À l’occasion des championnats académiques 2002, organisés cet après-midi à Besançon, les participants essaieront ainsi de **marcher sur les traces de leurs glorieux aînés**. (L’Est Républicain 11 décembre 2002)

- b. *Les participants essaieront de **leur marcher sur les traces**.

L'intérêt d'identifier le patron syntaxique des locutions est lié à l'hypothèse selon laquelle la structure syntaxique d'une locution a un impact sur son emploi. Ainsi, une locution construite à partir d'un verbe transitif tel que FRAPPER dans (9) ou di-transitif (PRENDRE dans l'exemple 10) sera plus encline à la passivation.

9. **Un grand coup a été frappé**, hier, à la salle polyvalente, lors de la première journée des Rencontres internationales de percussions. (Web)
10. Nous allions voir, ce que nous allions voir, **le taureau serait pris par les cornes**, demain ne serait plus comme hier. (FrWac)

Il nous est impossible de discuter, dans le cadre de cet article portant sur la polysémie des locutions, de toutes les applications possibles des SLS. Avant de quitter cette section, nous tenons néanmoins à souligner que l'identification des lexies formellement incluses dans les locutions n'est en rien incompatible avec leur non-compositionnalité sémantique. Une locution telle que MARCHER SUR LA TÊTE de l'exemple (1) est non compositionnelle dans la mesure où, en réponse à la question de savoir quel est son sens, un locuteur contemporain du français répondra plus que probablement « *Marcher sur la tête de quelqu'un*, cela signifie avoir un comportement dominant à son égard ». Une définition lexicographique du *RL-fr* reprendra cette explication dans une paraphrase construite dans un langage formel normé. L'identification des lexies incluses ne fait pas partie de l'analyse décompositionnelle du sens. Elle consiste à explorer une plus ou moins forte transparence du sens de la locution pour retrouver les éléments qui l'ont motivée. Cela revient donc à répondre à la question « Pourquoi dit-on de quelqu'un qui a un comportement dominant à l'égard d'un autre individu qu'*il lui marche sur la tête* ? ». Dans ce travail, le lexicographe étudie donc l'origine des locutions, telle qu'elle est attestée par les étymologistes, ou bien reconstituée par la communauté linguistique. Il se situe ainsi dans des perspectives que Perrin (2014) qualifie respectivement de versants historique et ascendant des sens locutionnels, par opposition au versant descendant, qui est le sens global – non compositionnel – de la locution.

4.2 Polysémie et/ou homonymie : ce que peuvent révéler les structures lexico-syntaxiques

En commentant les exemples (1) à (3) dans la section 3.2, nous avons signalé la difficulté d'identifier la relation sémantique, *ipse eo* de construire un pont sémantique, entre les sens de la locution MARCHER SUR LA TÊTE dans les phrases (1) et (2) d'une part, et (3) de l'autre. Cela revient donc à constater qu'une analyse décompositionnelle des signifiés conduit vers l'hypothèse de deux vocables locutionnels homonymes. Rappelons ces exemples avant de poursuivre notre discussion.

- (1) La fille, surprise d'un tel débordement verbal, eut son premier sourire et murmura en s'agaçant l'oreille d'une tige de graminée : – Vous pourriez vivre avec une jeune fille qui vous **marcherait sur la tête** ? (Frantext)
- (2) L'équipe US version 2002 a été humiliée trois fois pour ne finir qu'à la sixième place, le plus mauvais rang des États-Unis en quatorze éditions. Dans l'ordre, l'Argentine en poule (87-80), la Yougoslavie en quart de finale (81-78) et l'Espagne en match de classement (81-75) lui **ont marché sur la tête**. (*L'Est Républicain* 10 septembre 2002)
- (3) Les hommes politiques **marchent sur la tête** ! Il faudrait qu'on m'explique comment on peut recapitaliser la BEI avec de l'argent qu'on n'a pas ? (Web)

Observons à présent les SLS correspondant aux syntagmes que forment ces locutions. On constate en effet deux patrons syntaxiques différents que l'on associe à MARCHER SUR LA TETE des l'exemples (1) et (2) d'une part, et (3) d'autre part. Comparons les figures 10 et 11 :

| Composant | Acception | Forme |
|-----------|-------------|-------|
| V | marcher III | |
| Prép | | sur |
| Art | | la |
| NC | tête I.1a | |
| Prép | | |
| S2 | | |

Figure 10 : SLS de la locution MARCHER SUR LA TETE 'avoir un comportement dominant à l'égard de quelqu'un', et 'infliger une lourde défaite à quelqu'un' (exemples 1 et 2).

| Composant | Acception | Forme | Probabilité |
|-----------|--------------|-------|-------------|
| V | marcher I.1a | | 100 |
| Prép | | sur | 100 |
| Art | | la | 100 |
| NC | tête I.1a | | 100 |

Figure 11 : SLS de la locution MARCHER SUR LA TETE 'avoir un comportement déraisonnable à propos de quelque chose' (exemple 3).

Les SLS montrent que dans les deux cas, les locutions incluent des constituants lexicaux différents. MARCHER SUR LA TETE, 'avoir un comportement dominant à l'égard de quelqu'un' est relié à la lexie MARCHER III dont le sens peut être paraphrasé par 'poser son pied sur quelque chose' (cf. figure 10). La structure actancielle de cette lexie comprend deux actants : X marche sur Y. Il en est autrement pour la locution MARCHER SUR LA TETE qui signifie 'avoir un comportement déraisonnable' (cf. figure 11), qui renvoie à la lexie MARCHER I.1a 'se déplacer en faisant des pas', qui n'a qu'un seul actant X. Cette différence de constituants se traduit par deux systèmes de dépendances différents pour les locutions qui les contiennent. La figure 12 permet de les comparer.

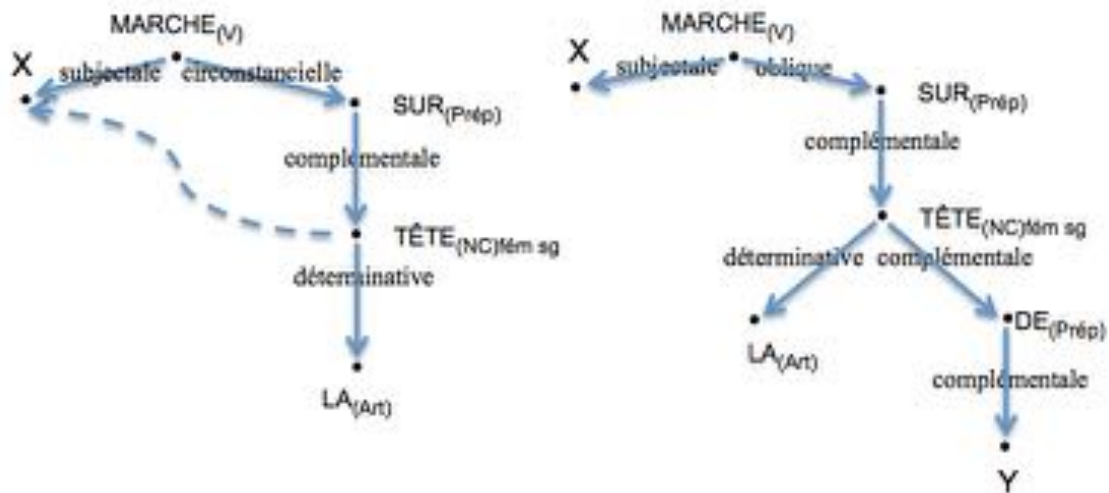


Figure 12 : Arbres de dépendance des locutions MARCHER SUR LA TÊTE ‘avoir un comportement déraisonnable’ (à gauche) et MARCHER SUR LA TÊTE ‘avoir un comportement dominant à l’égard de quelqu’un’ (à droite).

La flèche en pointillés ajoutée à l’arbre de dépendance de gauche indique que l’actant sémantique¹¹ de TÊTE correspond à l’actant de MARCHER. À l’inverse, dans l’arbre de droite, il est bien exprimé comme dépendant syntaxique de TÊTE et non pas du verbe. Les deux locutions ont alors une structure actancielle différente ; ce qui explique notamment la différence de comportement face à la montée du possesseur observée à la fin de la section 3.2.

Revenons à l’objet qui nous intéresse, à savoir la question de la polysémie et/ou de l’homonymie de la locution MARCHER SUR LA TÊTE. Du côté du signifié, il est facile – nous l’avons dit dans la section 3.2 – d’identifier la relation sémantique entre la locution de l’exemple (1) et celle de (2). Il est en revanche problématique de construire un pont sémantique reliant les deux acceptions illustrées dans les exemples (1) et (2) et MARCHER SUR LA TÊTE en (3). Ajoutons à cela des SLS différentes : ces syntagmes comprennent des constituants lexicaux et des réseaux de dépendances propres (figure 12). C’est dire que, malgré l’homophonie et l’homographie, le signifiant n’a pas la même forme. L’absence de lien sémantique de copolysémie s’accompagne de différences structurelles dans la forme du signifiant. Force est dès lors de reconnaître l’existence de deux vocables construits sur des lexies différentes du verbe MARCHER : un MARCHER SUR LA TÊTE¹, qui possède deux acceptions copolysèmes (MARCHER SUR LA TÊTE¹ I et MARCHER SUR LA TÊTE¹ II dans les exemples 1 et 2) et son homonyme MARCHER SUR LA TÊTE² (exemple 3) avec un seul sens. MARCHER SUR LA TÊTE¹ et MARCHER SUR LA TÊTE² sont morpho-phonologiquement identiques, mais lexico-syntaxiquement différentes puisque construites sur des lexies distinctes, avec des combinatoires propres¹².

Lorsque l’on conclut, comme pour MARCHER SUR LA TÊTE¹, à la polysémie, se pose naturellement la question de savoir laquelle des acceptions copolysèmes, en l’occurrence ‘infliger une lourde défaite à quelqu’un’ ou bien ‘avoir un comportement dominant à l’égard de quelqu’un’ doit être considérée comme ULB. Dans la mesure où le sens lexical décrit dans le *RL-fr* est abordé dans une perspective synchronique (cf. section 3.1 ci-dessus), il s’agit d’identifier l’acception la plus saillante pour un locuteur contemporain du français. C’est elle qui permettra ensuite de construire des ponts sémantiques entre les copolysèmes d’un vocable, et par conséquent, structurer le graphe qu’ils forment.

Nous reconnaissons volontiers qu'il n'est pas toujours facile de décider laquelle des deux acceptions de MARCHER SUR LA TÊTE¹ constitue l'ULB de ce vocable. On pourrait ainsi se tourner vers les corpus disponibles, en admettant que le sens le plus fréquent est le plus saillant pour l'ensemble des locuteurs. Le travail lexicographique développé dans le cadre du *RL-fr* exploite régulièrement trois corpus : Frantext (restreint aux ouvrages écrits à partir de 1950), le Corpus Français du Web (FrWac) et celui de l'Est Républicain (ER), sans renoncer aux exemples venant d'autres sources telles que les médias, le cinéma, le théâtre, le Web, la presse et diverses lectures personnelles. L'exploration de ces ressources a pour objectif de ramener des illustrations d'emplois.

Face à une pluralité de sens qu'il s'agit d'organiser en un réseau de copolysèmes, on pourrait utiliser des données fréquentielles, en admettant que l'ULB correspondra à l'emploi le plus fréquent de la locution. Le recours à la fréquence d'attestation en corpus présente cependant de nombreuses difficultés. Notons qu'aucun des corpus utilisés n'est désambiguïté, ce qui implique un dépouillement manuel des résultats. Quoiqu'abordable pour une séquence telle que MARCHER SUR LA TÊTE, appliquée à l'échelle d'une ressource à large couverture, cette démarche ralentirait sérieusement l'encodage des données lexicales, sans garantir une réponse certaine à notre question. Aussi, la séquence *marcher sur la tête* est-elle attestée 44 fois dans nos trois corpus de référence, respectivement 13 fois dans Frantext, 16 fois dans le corpus de l'ER, 15 fois dans le FrWac. Il nous a fallu cependant écarter onze occurrences, soit parce qu'elles se sont avérées de nature collocationnelle (référant par exemple à des comportements peu louables de footballeurs à l'égard des membres de l'équipe adverse), soit parce qu'elles correspondaient à des doublons fréquents dans le FrWac. Nous avons également éliminé une métaphore non lexicalisée, créée *ad hoc*, et deux emplois entre guillemets. Parmi les 33 occurrences retenues, une majorité (26) instancie le sens de 'avoir un comportement déraisonnable' du vocable monosémique MARCHER SUR LA TÊTE². Les données relatives aux deux sens copolysèmes de MARCHER SUR LA TÊTE¹ sont ainsi trop peu nombreuses pour servir de base à un raisonnement.

Plutôt que de recourir à la fréquence d'emploi d'un sens, le lexicographe dépouille donc l'organisation de la polysémie sur le plan sémantique. En effet, les deux sens de MARCHER SUR LA TÊTE¹ sont reliés par un lien métaphorique dont l'ULB constitue la source. Cette métaphore s'explique, si l'on considère comme source l'acception 'infliger une lourde défaite à quelqu'un'. Son copolysème MARCHER SUR LA TÊTE¹ II se laisse ainsi paraphraser 'avoir un comportement dominant à l'égard de quelqu'un, comme si on lui infligeait une lourde défaite'. Il serait, au contraire impossible d'explicitier le fonctionnement de l'extension polysémique à partir de 'avoir à l'égard de quelqu'un un comportement'.

Notons au passage qu'un vocable lexémique, tel que DOMINER, contient des acceptions quasi-synonymes de ces locutions, employées dans les exemples (11) et (12). L'encodage systématique de liens de copolysémie dans le *RL-fr* permet de comparer leur organisation sémantique, même si celle-ci devra faire objet d'une autre étude.

11. Les leçons des révolutions du XVIII^e siècle, les déclarations des droits de l'homme, les États européens les ont enseignées dans les écoles qu'ils ont ouvertes au profit des enfants des pays qu'ils **dominaient**. (Frantext)
12. Énorme coup de bambou pour l'équipe de France qui **a été** complètement **dominée** individuellement et collectivement au Stade de France pour le match d'ouverture de « Sa » Coupe du Monde, par une équipe d'Argentine impressionnante de maîtrise (17-12). (FrWac)

5 Conclusions

Nous avons présenté ici la modélisation particulière de la polysémie dans le *RL-fr*, plus précisément celle des locutions. Nous avons montré à travers le cas de MARCHER SUR LA TÊTE comment la structure lexico-syntaxique des locutions peut aider à trancher entre polysémie et homonymie. Nous sommes par conséquent amenés à envisager une conception élargie du signifiant d'une locution. En tant que syntagme, la locution est constituée de lexies. Son signifiant ne se limite alors pas au versant morpho-phonologique (face graphique et sonore du signe saussurien), mais s'analyse à un niveau lexico-syntaxique : deux locutions peuvent se prononcer et s'écrire de la même façon sans être complètement identiques d'un point de vue formel. Elles peuvent en effet être construites sur des lexies différentes appartenant au même vocable, comme c'est le cas de MARCHER SUR LA TÊTE. Elles sont alors homographes et homophones, sans pour autant être lexicalement identiques (cf. la notion de *quasi-signe* chez Kahane, 2015 : 45, pour désigner notamment les constituants lexicaux des locutions). Ceci nous permet d'avoir un critère additionnel pour trancher de façon sûre entre polysémie et homonymie ; critère qui n'a pas cours pour les lexèmes.

Il existe toutefois des cas, rares à notre connaissance, où des locutions sont entièrement identiques — graphiquement, phonologiquement et lexicalement — sans pour autant avoir un lien sémantique identifiable. C'est le cas de LOUP DE MER, qui a au moins trois acceptions¹³ : 1) 'marin endurci et expérimenté', 2) 'espèce de poisson' et, par métonymie, 3) 'chaire de ce poisson, que l'on consomme'. S'il est aisé de lier sémantiquement les sens 2 et 3, il est moins facile de les rattacher au sens 1, qui doit alors être considéré comme appartenant à un autre vocable.

Il est fort à parier que l'exploitation des structures lexico-syntaxiques réserve aux lexicographes plus d'une surprise !

Références bibliographiques

- Bosredon, B. (2004). Le paramètre catégoriel dans les unités polylexicales : de la polysémie à la néologie, *Syntaxe et sémantique*, 1 (5), 47–58
- Collectif (2004). *Syntaxe et sémantique*, 1(5). Caen : Presses Universitaires de Caen.
- Dubois, J., Dubois-Charlier, F. (1997). *Les Verbes français*. Paris : Larousse-Bordas.
- Fellbaum, C. (éd.) (1998). *WordNet : An Electronic Lexical Database*. Cambridge (MA) : MIT Press.
- Fillmore, C., Johnson, C., Petrucci, M. (2003). Background to Framenet, *International Journal of Lexicography*, 16 (3), 235–250.
- Gader, N., Lux-Pogodalla, V., Polguère, A. (2012). Hand-Crafting a Lexical Network With a Knowledge-Based Graph Editor. *Proceedings of the Third Workshop on Cognitive Aspects of the Lexicon (CogALex III), The COLING 2012 Organizing Committee, Mumbai*, 109–125.
- Gaume, B. (2004). Balades aléatoires dans les petits mondes lexicaux. *Information interaction intelligence*, 4(2), 39–96.
- Kahane S. (2015). Les trois dimensions d'une modélisation formelle de la langue : syntagmatique, paradigmatique et sémiotique, *TAL*, 56 (1), 39–63.
- Kiss, G. R. (1968). Words, associations, and networks. *Journal of Verbal Learning and Verbal Behavior*, 7, 707-713.
- Lux-Pogodalla, V. Polguère, A. (2011). Construction of a French Lexical Network. Methodological Issues. *Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011. An ESSLLI 2011 Workshop, Ljubljana*, 54–61.
- Mel'čuk, I. (2004a). Actants in Semantics and Syntax I : Actants in Semantics, *Linguistics*, 42 (1), 1–166.
- Mel'čuk, I. (2004b). Actants in Semantics and Syntax II : Actants in Syntax, *Linguistics*, 42 (2), 247–291.
- Mel'čuk, I. (1996). Lexical Functions : A Tool for the Description of Lexical Relations in the Lexicon. In : Wanner, L. (éd.) *Lexical Functions in Lexicography and Natural Language Processing, Language Companion Series*, 31, Amsterdam/Philadelphia : John Benjamins, 37–102.
- Mel'čuk, I., Clas, A., Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve : Duculot.
- Ollinger, S. (2014). Regroupement de structures de dérivations lexicales par raisonnement analogique. In : Bigi, B. (éd.) *Actes de Récital 2014, Faculté Saint Charles, Aix Marseille Université (Marseille), du 1er au 4 juillet 2014*, 92–103.
- Osherson, A., Fellbaum, C. (2009). The representation of idioms in WordNet. *Actes de ACL-IJCNLP 09*.
- Pausé, M.-S. (à paraître). Modélisation de la structure lexico-syntaxique des locutions au sein d'un réseau lexical. In : Kauffer, M., Keromnes, Y. (éd.) *Approches théoriques et empiriques en phraséologie. Actes du colloque international (Nancy, 11-12 décembre 2014)*. Tübingen : Stauffenburg.
- Pausé, M.-S. (en préparation). *Modélisation de la structure lexico-syntaxique des locutions françaises au sein d'un réseau lexical*. Thèse, Université de Lorraine, ATILF-CNRS (UMR 7118), Nancy.
- Perrin, L. (2014). Citation, lexicalisation et interprétation des expressions idiomatiques. In : Authier-Revuz, J., Doury, M. (éd.) *Parler des mots : le fait autonymique en discours*. Paris : Presse Sorbonne Nouvelle, 281–290.
- Polguère, A. (1997). Meaning-Text Semantic Networks as a Formal Language. In : Wanner, L. (éd.) *Recent Trends in Meaning-Text Theory*. Amsterdam, Philadelphie : John Benjamins, 1–24.
- Polguère, A. (2009). Lexical systems: graph models of natural language lexicons. *Language Ressources and Evaluation*, 43, 41–55.
- Polguère, A. (2011). Classification sémantique des lexies fondée sur le paraphrasage. *Cahiers de Lexicologie*, 98 (1), 197–211.

- Polguère, A. (2014). From Writing Dictionaries to Weaving Lexical Networks. *International Journal of Lexicography*, 27 (4), 396–418.
- Pruvost, J. (2011).) À travers siècles, un article de dictionnaire et ses séquences figées. In : Anscombe, J.-C., Mejri, S. (éd.) *Le figement linguistique : la parole entravée*. Paris : Honoré Champion, 341-362.
- Selva, T., Verlinde, S., Binon, J. (2003). Vers une deuxième génération de dictionnaires électroniques. *Traitement Automatique des Langues*, 44 (2), 177–197.
- Sikora, D. (2014). Brève histoire d'une erreur lexicale. Polysémie et liens lexicaux dans l'enseignement du vocabulaire. In : Neveu, F., Blumenthal, P., Hriba, L., Gerstenberg, A., Meinschafer, J., Prévost, S. (éd.) *4e Congrès Mondial de Linguistique Française – CMLF 2014*. Paris : Institut de Linguistique Française, 1157–1172.
- Wilkins, D. (1996). Natural Tendencies of Semantic Change and the Search for Cognates. In : Durie, M., Ross, M. (éd.) *The Comparative Methods Reviewed. Regularity and Irregularity of Semantic Change*. New York / Oxford : Oxford University Press, 264–304.

Ressources en ligne :

Frantext : <http://www.frantext.fr/>

FrWac : <http://nl.ijs.si/noske/index.html>

Est Républicain : <http://www.cnrtl.fr/corpus/estrepublikain/>

¹ Nous remercions Sandrine Ollinger pour les précieuses statistiques de la base lexicale qu'elle nous a fournies, sur lesquelles nous nous sommes appuyées pour ce travail. Nous remercions également Alain Polguère pour les discussions fructueuses que nous avons eues ensemble, ainsi que Véronika Lux-Pogodalla pour ses retours sur une première version de cet article. Nos remerciements vont aussi aux relecteurs anonymes : leurs remarques nous ont permis d'améliorer le texte. Nous sommes entièrement responsables de toute erreur pouvant subsister.

² Le Wiktionnaire, ressource collective enrichie et modifiée par la communauté internet, se distingue des grands dictionnaires tels que le Robert et le Larousse, en accordant aux locutions leurs propres articles. (<http://fr.wiktionary.org>)

³ TELEPHONE PORTABLE est identifiée comme locution dans Antidote et Wiktionnaire. Aucune catégorie grammaticale n'est spécifiée dans le Robert. Néanmoins, ce dernier considère POMME DE TERRE, unité polylexicale également analogue à un syntagme, comme un nom commun. Quant au Larousse, dans sa version en ligne il n'offre aucune correspondance pour cette unité. Pour une description morphologique – et donc une conception comme nom composé – voir Bosredon (2004).

⁴ Pour un descriptif succinct, voir : <http://www.atilf.fr/spip.php?article908>

⁵ Un système lexical de l'anglais est également en cours d'élaboration. Bien que dans un état de développement moins avancé, des bases du coréen et de l'espagnol ont été développées.

⁶ Pour une présentation de la notion d'*étiquette sémantique*, voir Polguère (2011). Rappelons simplement qu'il s'agit d'une caractérisation minimale classifiante du sens décrit : les lexies telles que AUTO, AUTOCAR, BUS, CAMION, TRAMWAY, VOITURE, etc. portent l'étiquette 'véhicule'.

⁷ Le travail d'élaboration des définitions est actuellement dans une phase initiale. Dans la mesure où il s'agit d'une démarche strictement normalisée et contrôlée, il est en effet indispensable de mettre en place une procédure de rédaction explicite et une feuille de route précise.

⁸ Nous empruntons cette paraphrase définitoire au *Petit Robert 2011* dans sa version électronique.

⁹ Notons que ce score est encore bien inférieur à celui de la base *Les verbes français* de Dubois et Dubois-Charlier (1997), qui décrit 60 acceptions du verbe PASSER.

¹⁰ Notons que la polysémie de certaines lexies dites grammaticales n'est à ce jour pas suffisamment travaillée. C'est la raison pour laquelle, dans certaines SLS, nous ne pointons pas encore vers la lexie, mais affichons uniquement la forme.

¹¹ Sur l'opposition entre actants sémantiques et syntaxiques, voir notamment Mel'čuk (2004a et 2004b)

¹² MARCHER SUR LA TÊTE n'est, bien entendu, pas un cas isolé. Nous pouvons citer, entre autres, l'exemple de COUP DE BARRE. La locution a au moins deux sens : 'fatigue subite' (le fameux coup de barre provoqué par la digestion post-méridionale) et 'changement' (le coup de barre que l'on donne à sa vie en changeant de profession). Si le second sens est très nettement construit par analogie avec le changement de cap d'un bateau à l'aide de la barre actionnant le gouvernail, le premier est vraisemblablement construit par analogie avec un coup de barre (de fer, par exemple) assommant, reçu sur la tête. Les SLS seront alors différentes et nous aurons deux vocables homonymes.

¹³ Au moins, car les encyclopédies utilisent ce syntagme comme nom vernaculaire de plusieurs espèces, qu'il s'agisse de phoques ou de poissons.

DOROTA SIKORA
ATILF – CNRS (UMR 7118)
dorota.sikora@atilf.fr

DÉFINIR LE SENS DANS UN RÉSEAU LEXICAL*

Abstract. Dorota Sikora, *Définir le sens dans un réseau lexical* [How to define a word in a lexical network?], *Studia Romanica Posnaniensia*, Adam Mickiewicz University Press, Poznań, vol. XXXIX/3: 2012, pp. 63-79. ISBN 978-83-232-2469-3. ISSN 0137-2475. eISSN 2084-4158.

As underlined in Wierzbicka (1996: 239), a « linguistic definition is a scientific hypothesis about the concept encoded in a given word ». The present paper focuses on the problem of what such hypothesis is about. Lexicon of any language is structured with a set of connections relating lexical items. That is to say that speakers of any language need to master not only separate words, but rather a cluster of connected senses. If so, then a lexical definition should lay out the concept the word lexicalizes in the way that gives access to the whole grasp of its conceptual, paradigmatic and syntagmatic relations. We discuss the topic of definitional structures developed in the context of the French Lexical Network (Réseau Lexical du Français).

Key words : lexical definition, Explanatory and Combinatorial Lexicology, lexical resources

INTRODUCTION

Dans cet article, nous présentons les définitions lexicographiques du Réseau Lexical du Français (désormais RLF). Cette ressource lexicale se pose deux objectifs indissociablement liés. D'une part, elle fournit à l'utilisateur toutes les informations sémantiques, syntaxiques, morphologiques nécessaires pour maîtriser l'usage d'une unité lexicale. De l'autre, elle se construit en tant que modèle représentant la structuration du lexique du français sous forme d'un graphe.

Dans la section qui suit, nous présenterons brièvement le projet RLF et son cadre théorique. La section (2) nous permettra de décrire le patron général des définitions, en exposant les principes qui régissent l'élaboration des paraphrases définitoires. Dans la troisième partie, nous nous servirons de quelques exemples de définitions du RLF pour montrer les relations conceptuelles qu'elles permettent d'établir. Nous discuterons ensuite les liens paradigmatiques et syntagmatiques qu'elles décrivent (section 4). Toutes ces réflexions nous amèneront vers des conclusions exposées dans la dernière section.

* Nous tenons à remercier chaleureusement Alain Polguère pour sa relecture critique d'une première version de cet article.

Précisons dès à présent deux notions essentielles pour notre exposé. Dans les pages qui suivent, nous employons systématiquement les termes de *vocable* et d'*unité lexicale*. Le vocable est un ensemble comprenant minimalement une, et potentiellement plusieurs unités lexicales. Nous parlerons indifféremment d'*unité lexicale*, *lexie*, *acception* ou *sens*.

1. CADRE THÉORIQUE DU RLF ET PRÉSENTATION GÉNÉRALE DE LA RESSOURCE

1.1. LEXICOLOGIE EXPLICATIVE ET COMBINATOIRE

L'impulsion pour la conception du RLF vient de la réflexion menée par les linguistes de la Lexicologie Explicative et Combinatoire (désormais LEC) dans le cadre plus général de la Théorie Sens-Texte développée par Mel'čuk (pour une introduction, voir, entre autres, Mel'čuk, 1997, Milićević, 2006, Polguère 2011b). Cette théorie se caractérise par une orientation résolument lexicaliste, en plaçant l'ensemble des unités lexicales au cœur du système linguistique : « En exagérant quelque peu, on pourrait même dire que l'ensemble de lexies est la langue. En effet, une langue est constituée de lexies et de règles servant à la manipulation de celles-ci » (Mel'čuk, I., Clas, A. et A. Polguère, 1995 : 17).

Il est clair qu'une telle vision du lexique n'est pas indifférente pour le travail lexicographique. Depuis les années quatre-vingt, elle a conduit à l'élaboration de tout un ensemble de principes déterminant¹ la démarche descriptive, que l'on désigne avec Polguère (*à paraître*), sous le nom de Lexicographie Explicative et Combinatoire (LgEC).

La LEC envisage l'unité lexicale en tant que signe linguistique sous ses deux facettes saussuriennes – signifiant et signifié – en les complétant par une troisième : sa combinatoire syntaxique. La pratique de la LgEC consiste donc à analyser et à décrire ces trois aspects constitutifs, et indissociables dans le discours, du signe linguistique. Voyons à présent comment une ressource conçue dans cet esprit répond aux postulats de la LEC.

1.2. RÉSEAU LEXICAL DU FRANÇAIS

Le RLF n'est pas la première entreprise réalisée selon les préceptes de la LEC². Il suffit de rappeler les quatre volumes du Dictionnaire Explicatif et Combinatoire (Mel'čuk et *alii* 1984, 1988, 1992, 1999) qui présentent des discussions focalisées

¹ Il nous est impossible, dans le cadre du présent article, ne serait-ce que de les énumérer. Le lecteur pourra se reporter à Mel'čuk, Clas et Polguère (1995).

² Pour une discussion approfondie de ces ouvrages et ressources, voir Polguère (*à paraître*).

sur des problèmes lexico-sémantiques précis illustrées par des descriptions lexicographiques de plus de 600 vocables. Il s'agit néanmoins d'ouvrages à destination de linguistes plutôt que d'un utilisateur non spécialiste. *Le lexique actif du français* (Mel'čuk et Polguère 2007) s'adresse, au contraire, à un public plus large non initié. Il se conjugue également en version de base lexicale consultable en ligne³.

Le RLF⁴, développé actuellement au Laboratoire ATILF CNRS (UMR 7118) sous la direction scientifique d'Alain Polguère, est une base de données lexicales destinée à assurer, au terme du projet, une large couverture du lexique français. Sa première étape de trois ans consiste à proposer une description du vocabulaire de base (McCarthy, 1999), c'est-à-dire d'environ 3700 vocables dont les lexies se caractérisent par un potentiel communicatif le plus élevé dans le lexique français. Néanmoins, cette nomenclature doit rapidement s'accroître selon des règles bien déterminées (Polguère et Sikora, à paraître).

Les données lexicales du RLF permettent de traiter toutes les facettes du signe linguistique. En cas de polysémie (c'est-à-dire dans la plupart des cas), toutes les acceptions du vocable sont clairement distinguées et décrites séparément. Chaque unité lexicale (lexie-vedette) est identifiée en termes de ses caractéristiques grammaticales. Une définition analytique (voir la section suivante), les liens lexicaux (*fonctions lexicales*) qu'elle contrôle constituent la partie sémantiquement essentielle de l'article. Viennent ensuite des exemples tirés de Frantext, d'un corpus journalistique de presse régionale et d'un autre comprenant des tranches du Web. Enfin, la rubrique intitulée « Nota bene » permet de fournir d'autres informations, par exemple les connotations véhiculées par l'unité lexicale.

On imagine facilement la taille que représenterait une telle entrée dans un dictionnaire traditionnel. Par ailleurs, l'exploitation de toutes ces données saisies sous forme d'un texte serait extrêmement lourde pour les utilisateurs de la ressource. Il est donc primordial de préciser que le RLF est un dictionnaire virtuel (Selva, Verlinde et Binon, 2003). Les données sont atomisées et stockées dans des tables d'une base de données de type SQL via l'éditeur MVS_Dicet⁵. Cette structure de données permet d'envisager la génération automatique des dictionnaires de différents types. La figure 1 offre une vue d'article pour une entrée dans le RLF.

L'objectif que se pose le RLF va pourtant au-delà de création de ressources dictionnairiques linéaires de par leur nature. Il s'agit de créer un modèle du lexique du français aussi fidèle que possible à l'« original ». Conformément à son nom, le RLF représente le réseau que constitue le lexique, en l'occurrence celui du français. Cela signifie que les données saisies dans la base de données doivent permettre de créer un graphe connexe d'unités lexicales.

³ <http://olst.ling.umontreal.ca/dicopop/>

⁴ Pour une présentation du projet, voir : <http://www.atilf.fr/spip.php?article908>

⁵ L'éditeur MVS_Dicet est développé par Nabil Gader de la société Micro Video Soft, partenaire du Laboratoire ATILF dans le projet RELIEF/RLF.

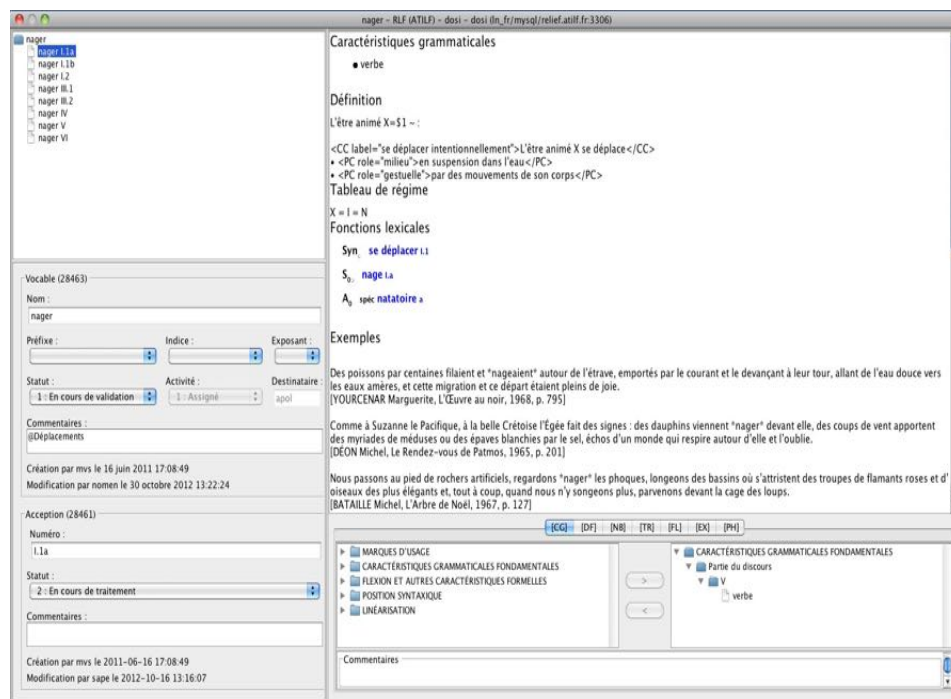


Fig. 1. Vue d'articles des unités lexicales du vocable NAGER dans le RLF

Dans les paragraphes qui suivent, nous nous concentrerons sur le volet définitionnel du RLF. Nous expliquerons comment se construisent les paraphrases définitoires dans un réseau lexico-sémantique.

2. STRUCTURE DES DÉFINITIONS DANS LE RLF

2.1. STRUCTURE DU *DEFINIENDUM*

À l'exception des ressources lexicographiques réalisées dans le cadre de la LEC que nous venons de mentionner, les dictionnaires traditionnels, indépendamment de leur mode d'existence électronique ou papier, ne diffèrent guère quant à l'objet des définitions qu'ils proposent. Le *definiendum* équivaut habituellement à la forme du vocable accompagnée d'indications grammaticales.

Conformément aux principes de la LEC, pour un lexicographe du RLF, le premier pas vers une définition consiste à dégager la forme propositionnelle propre à chaque sens à paraphraser. La forme propositionnelle est une formule explicitant sous forme de variables X, Y, Z, etc., tous les arguments sémantiques d'un prédicat. Ainsi, le verbe NAGER des exemples (1) et (2), tirés de Frantext, se voit associer deux formes propositionnelles distinctes.

- (1) Des poissons par centaines filaient et *nageaient* autour de l'étrave, emportés par le courant et le devançant à leur tour, allant de l'eau douce vers les eaux amères, et cette migration et ce départ étaient pleins de joie.
- (2) Puis, avec vigueur, il *nageait* la brasse vers le large... pas très loin.

Dans (1), il s'agit d'un prédicat monoargumental $X \sim$, alors que pour le sens (2), un second argument Y peut être sélectionné, ce qui nous amène à lui attribuer la forme propositionnelle $X \sim Y$.

La forme propositionnelle est ainsi un premier indice conduisant à distinguer deux acceptions du verbe *NAGER*, respectivement **Ia** et **Ib**, là où les dictionnaires ne traitent que d'un seul sens qu'il s'agisse d'un être humain ou d'un animal aquatique.

Nous verrons, dans la sous-section suivante, que l'explicitation de la forme propositionnelle entraîne de nombreuses conséquences pour le *definiens*.

2.2. STRUCTURE DU *DEFINIENS*

Les acceptions **Ia** et **Ib** ont reçu les paraphrases définitoires suivantes dans le RLF :

NAGERIa

L'animal $X \sim$:

- <CC label="se déplacer intentionnellement">L'animal X se déplace</CC>
- <PC role="milieu">en suspension dans l'eau</PC>
 - <PC role="gestuelle">par des mouvements de son corps</PC>

NAGERIb

L'individu $X \sim$ (en utilisant la technique Y) :

- <CC label="se déplacer intentionnellement">L'individu X se déplace</CC>
- <PC role="milieu">en suspension dans l'eau</PC>
 - <PC role="gestuelle">par des mouvements de son corps (conformes à la technique Y)</PC>
 - <PC role="background">ce mode de déplacement ayant été appris par X</PC>

Au prime abord, ces définitions peuvent paraître quelque peu déroutantes, notamment en raison des balises segmentant leurs contenus. Précisons que l'interface d'utilisateur prévue dans le RLF débarrassera le lecteur de ces marques. Il lira ainsi le texte de la définition de **NAGERIa** sous forme d'une phrase : « L'animal X se déplace en suspension dans l'eau par des mouvements de son corps ». Si nous avons choisi de garder les balises dans notre présentation, c'est parce que nous espérons d'une part, mettre en avant la structuration rigoureuse des paraphrases définitoires et de l'autre faciliter, le cas échéant, le renvoi à une composante définitionnelle précise. Autrement dit, elles nous permettront de naviguer plus aisément à travers les sous-réseaux

du lexique français, en suivant les liens qu'elles permettent d'établir entre les unités lexicales.

De fait, comme toutes les définitions du RLF, celles qui paraphrasent le sens de **NAGERIa** et de **NAGERIb** ont une structuration analytique : elles articulent une composante centrale (CC), qui exprime le genre prochain, et des composantes périphériques (PC) qui expriment la différence spécifique. La CC de chaque définition est porteuse d'une étiquette sémantique (*label*) qui correspond à « la forme normalisée du genre prochain de la définition qui établit l'appartenance de L à une classe de lexies d'au moins deux éléments » (Polguère 2011 : 202). Quant aux PC, le type d'information apportée par la composante est caractérisé par un attribut (*role*).

Notons que, dans le *definiens*, figurent obligatoirement toutes les variables actancielles présentes dans la forme propositionnelle avec un typage sémantique, si nécessaire. Ainsi, les actants X de **NAGERIa** et de **NAGERIb** relèvent de deux types sémantiques différents, respectivement *animal* et *individu*. Il s'agit certes, d'une contrainte pour le lexicographe, mais également d'un outil précieux, puisque le rédacteur introduit ainsi dans la définition tous les participants de la situation, en spécifiant les rapports entre eux.

Toutes les PC, qui expriment la différence spécifique associée au genre prochain de la définition, doivent impérativement se rattacher directement à la CC. Il est interdit de hiérarchiser ces composantes entre elles, en leur adjoignant d'autres PC.

En plus de la CC et de PC, il est nécessaire, pour certaines lexies, d'introduire une composante présuppositionnelle. C'est ainsi que nous avons procédé en construisant la paraphrase de l'acception **NAGERVI** illustrée par l'exemple (3) :

(3) Débutante en latin, je *nage* complètement !!

L'unité lexicale **NAGERVI** a été définie de la façon suivante :

L'individu X ~

<PP role="contexte">Dans une situation complexe α que X doit maîtriser</PP>||
 <CC label="être dans un état psychologique">L'individu X ne comprend pas α </CC>
 <PC role="effet">ce qui fait que X éprouve une gêne</PC>

Selon notre analyse, l'information contextuelle concernant la complexité de la situation α que X est censé maîtriser, ne fait pas partie du sens verbal décrit. Il ne s'agit pas non plus d'un participant du fait dénoté par **NAGERVI**. C'est un contenu présupposé indispensable pour comprendre et utiliser cette lexie.

Le schéma ci-dessous (figure 2) réunit l'ensemble de composantes susceptibles d'apparaître dans la partie définitoire, en visualisant sa structuration. Il nous est impossible, dans le cadre de cet article, de citer l'inventaire complet des CP employées dans les définitions selon les attributs qu'elles portent. Celles dans la figure 2 ne sont que quelques exemples parmi d'autres, choisis dans le but d'illustrer notre propos sur la structure du *definiens*.

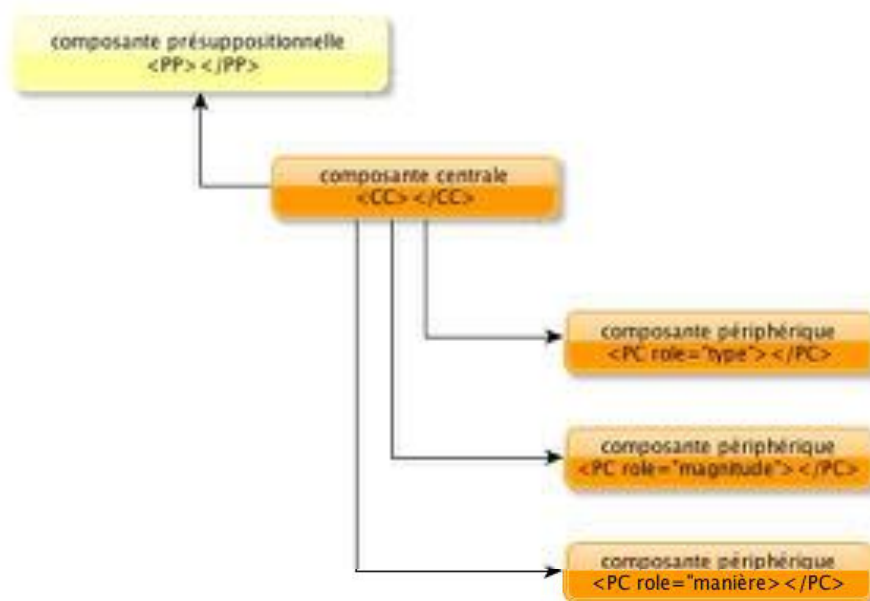


Fig. 2. Structure générale des définitions du RLF

3. DÉFINITIONS ET LIENS CONCEPTUELS

Dans cette section, nous expliquerons comment la structuration des définitions du RLF permet de reconstituer et de suivre les liens lexicaux de nature conceptuelle. Nous nous intéresserons tout particulièrement à ceux qui sont à la base des classes de sens (section 3.1) et à ceux qui relient les acceptions d'un vocable polysémique (section 3.2).

3.1. COMPOSANTES CENTRALES, COMPOSANTES PÉRIPHÉRIQUES ET CLASSES DE SENS

Commençons par le rôle que joue l'étiquette sémantique dans l'architecture d'une définition et bien au-delà de celle-ci. La CC du verbe **NAGER****la** porte la même étiquette « se déplacer intentionnellement », la même que celle que l'on retrouve dans la définition de **VOLER****la** présentée ci-dessous. L'exemple (4) illustre ce sens.

(4) Il regarda le ciel encore clair où *volaient* des martinets.

L'animal X ~ :

- <CC label="se déplacer intentionnellement">L'animal X se déplace</CC>
- <PC rôle="milieu">dans l'air</PC>
- <PC rôle="moyen">supporté par des parties spécifiques de son corps avec lesquelles il fait des mouvements qui lui permettent de se propulser</PC>

La classe de sens que constituent les verbes de déplacement français compte par ailleurs nettement plus de membres que le minimum requis par Polguère (*cf. supra*), puisqu'elle comprend les lexies telles que *courir*, *marcher*, *ramper* et bien d'autres.

Les principes définitionnels retenus dans le RLF permettent ainsi de situer la lexie décrite dans le sous-système conceptuel dont elle relève. Le paraphrasage rigoureux qu'impose une étiquette sémantique sert à connecter la lexie avec les autres membres de sa classe de sens. Nous avons pu constater, en effet, que les définitions de **NAGER_{Ia}** et de **VOLER_{Ia}** reposent sur une même paraphrase minimale. Dans une optique plus large, l'étiquette sémantique permet d'indiquer clairement la place qu'occupe une unité lexicale selon l'axe vertical de la hiérarchie des sens codés dans la langue. En remontant, on trouvera successivement des étiquettes de plus en plus vagues : « se déplacer », « être en mouvement », « exercer une activité », et ainsi de suite jusqu'à ce que le sommet de la hiérarchie soit atteint avec « fait » et « entité » qui se rejoignent sous l'étiquette commune « quelque chose ».

Parallèlement, les PC exprimant la différence spécifique d'une définition sont toujours caractérisées par un attribut. Une comparaison ne serait-ce que très rapide de **NAGER_{Ia}** et de **VOLER_{Ia}** permet de constater la présence des PC du même type dont le contenu véhicule les éléments de sens propre à chacune des deux lexies. Les attributs (milieu, gestuelle dans le cas de **NAGER_{Ia}**) qui caractérisent les PC apparaissent, bien que dans des configurations différentes, dans les définitions des verbes de la même classe de sens. Ainsi, pour définir la lexie **VOLER_{Ia}**, on recourt à la PC milieu, alors que la "gestuelle" est remplacée par "moyen" pour inclure le vol plané des oiseaux.

3.2. STRUCTURE DES DÉFINITIONS ET TRAITEMENT DE LA POLYSÉMIE DANS LE RLF

Il n'est pas rare que les lexies d'un même vocable polysémique appartiennent à des classes de sens bien différentes. Il suffit de penser au verbe **ROULER** dont la première acception **ROULER_{I1}** dénote un déplacement non intentionnel (exemple 5), **ROULER_{I2}** de (6) est un causatif de déplacement, alors qu'avec **ROULER_{II}** dans (7), il s'agit plutôt d'un mouvement corporel.

(5) Les pierres *roulent* sous les pieds ; on se cramponne ; on se rattrape aux branches ; mon sabre devient un alpenstock.

(6) Un paysan *roule* un tonneau.

(7) Il grince des dents et *roule* des yeux féroces d'assassin de film muet.

Avant de présenter le traitement lexicographique de la polysémie dans le réseau lexical décrit par le RLF, quelques précisions terminologiques s'imposent.

Nous désignerons désormais les unités lexicales (lexies, acceptions, sens) d'un vocable polysémique sous le nom de **co-polysèmes**.

Les liens reposant sur des composantes sémantiques communes aux co-polysèmes d'un vocable polysémique seront appelés **ponts sémantiques**⁶.

L'un des co-polysèmes constitue l'unité lexicale de base (ULB), c'est-à-dire le sens auquel tous les autres sont connectés par des ponts sémantiques directement ou indirectement, *via* une autre acception.

La position résolument synchronique du RLF est décisive pour le traitement de la polysémie. Il faut en effet que les ponts sémantiques entre les co-polysèmes soient accessibles à l'utilisateur sans détour étymologique. Prenons le cas du verbe ROULER dans la phrase (8) ci-dessous :

(8) Nous étions en plein procès avec l'ancien propriétaire qui nous *avait roulés* et nous allions perdre.

Dans les dictionnaires⁷ qui décrivent la polysémie en s'appuyant sur le « parcours » étymologique du vocable, la lexie ROULER de l'exemple (8) est un co-polysème de ROULER employé dans (5), (6) et (7). Dans le RFL cependant, on sera amené à les traiter en termes d'homonymie, puisqu'à moins d'une construction conceptuelle très complexe et artificielle, il est impossible de construire un pont sémantique en synchronie susceptible de relier le sens de l'exemple (8) aux précédents. Il convient par conséquent d'en faire deux articles séparés⁸, respectivement ROULER1 et ROULER2.

Les composantes définitionnelles sont l'outil essentiel pour relier les co-polysèmes. La première étape consiste à identifier l'ULB, car c'est elle qui constitue le point de départ du raisonnement structurant toutes les autres acceptions. Autrement dit, c'est en partant de l'ULB et de sa paraphrase définitoire que le lexicographe – et à sa suite l'utilisateur du RLF – pourra construire un premier pont sémantique, puis un deuxième et tous ceux qui le mèneront d'une acception à l'autre pour comprendre la rationalité qui sous-tend la structuration d'un vocable polysémique. Si la définition de l'ULB est le point de départ, il faut que le pont sémantique soit solidement ancré à l'autre bout, c'est-à-dire du côté du co-polysème cible. La seconde tâche dans le traitement de la polysémie est donc d'élaborer des paraphrases définitoires qui explicitent les liens entre eux.

Dans la section 3.2, nous avons présenté les lexies NAGER1a et NAGER1b reliées par la même CC et par la PC 'milieu'. Cependant, les connexions sont parfois plus complexes. Observons les définitions de deux acceptions qui appartiennent à la très

⁶ Nous adoptons ainsi les notions proposées et définies par Mel'čuk, Clas et Polguère (1995 : 157).

⁷ Cf. à titre d'exemple *Le Petit Robert* (2011) ou bien le site du *Centre National des Ressources Textuelles et Lexicales* (<http://www.cnrtl.fr/definition/roule>).

⁸ L'enquête que nous avons menée auprès de francophones natifs indique clairement qu'il n'est pas possible de rattacher ROULER2 à ROULER1 *via* la locution *rouler quelqu'un dans la farine*. Il semblerait en effet qu'un locuteur contemporain puisse utiliser le verbe *rouler* avec le sens de *tromper* tout en ignorant la locution en question.

riche polysémie du vocable **ROULER**. En effet, **ROULER^{I1}** et **ROULER^{I2}**, ont été illustrées, au début de la présente section, respectivement par les exemples (5) et (6).

ROULER^{I1}

X ~ sur Y

- <CC label="être en déplacement">X est en déplacement</CC>
- <PC role="causateur">par des mouvements autour de son axe</PC>
- <PC role="milieu">sur le support Y</PC>

ROULER^{I2}

W ~ X sur Y

- <CC label="déplacer">W déplace X</CC>
- <PC role="manière">en faisant rouler^{I1} X</PC>

La lexie **ROULER^{I1}**, ULB du vocable **ROULER**(1), est sémantiquement la plus simple parmi les co-polysèmes et elle sert de base pour la construction de ponts sémantiques. Le premier d'entre eux, celui qui la relie avec le sens **I2**, est une inclusion. **ROULER^{I2}** est plus riche, puisqu'il lexicalise une relation causale, en sélectionnant un actant agentif. Autrement dit, il inclut le sens de base, en y ajoutant du contenu informationnel.

Il n'est pas rare qu'au fur et à mesure que l'on s'éloigne de l'ULB, les ponts deviennent plus difficiles à identifier. Il est parfois impossible de détecter des composantes communes avec la lexie de base. **ROULER^{I1}** de l'exemple (7) ci-dessous semble effectivement bien différent de celui que nous venons de définir pour la phrase (5).

Le lien qui relie **ROULER^{I1}** à l'ULB est de toute autre nature que dans le cas du causatif **I2**. L'hypothèse que nous avons adoptée consiste à admettre que le pont entre **ROULER^{I1}** et **I1** employée dans l'exemple (7) se construit par métaphore. Contrairement à l'ensemble de dictionnaires que nous avons pu consulter, le RLF ne se contente pas de laisser à l'utilisateur le soin d'identifier les liens métaphoriques. Une métaphore lexicalisée étant partie intégrante du sens, elle est systématiquement explicitée dans la définition-cible à travers une PC spécifique porteuse de l'attribut 'métaphore' :

ROULER^{I1}

L'individu X ~ la partie de son corps X1

- <CC label="remuer">L'individu X remue la partie X1 de son corps</CC>
- <PC role="manière">par des mouvements circulaires</PC>
- <PC role="métaphore">comme si X1 *roulait*.1</PC>

Construire des ponts sémantiques en définissant est ainsi une tâche qui génère une double contrainte. D'une part, en paraphrasant le sens de départ sans perdre de vue les co-polysèmes auxquels il doit nous mener par la suite. De l'autre, il faut éviter la redondance dans les définitions des sens ciblés. Cette discipline vaut pourtant l'effort qu'elle requiert dans la mesure où les mécanismes des évolutions sémantiques à l'ori-

gine de la polysémie produisent, certes, des effets spécifiques à chaque langue, mais sont identiques quant à leur nature.

4. DÉFINITIONS ET LIENS PARADIGMATIQUES ET SYNTAGMATIQUES

Jusqu'à présent, nous avons employé le terme de paraphrase, en la qualifiant de définitoire, comme synonyme de définition. Nous adopterons à présent une perspective plus large pour nous occuper des liens lexicaux qui autorisent le paraphrasage tel qu'il est placé au cœur du modèle Théorie Sens-Texte. Par paraphrasage, on entend la synonymie entre les phrases (*textes*) qui expriment le même sens, par exemple entre (9) et (10) :

(9) Jean nage très bien.

(10) Jean est un très bon nageur.

Nous présentons le traitement des relations paradigmatiques dans le RLF dans la sous-section (4.1), pour passer ensuite, en (4.2), aux liens syntagmatiques.

4.1. DÉFINITIONS ET LIENS PARADIGMATIQUES

Dans la section 3, nous avons présenté les définitions des lexies **NAGERIa** et **NAGERIb** qui dénotent respectivement le déplacement d'un animal aquatique, typiquement un poisson, et celui qui peut être maîtrisé par un être humain. Ce choix peut paraître étonnant dans la mesure où les dictionnaires aussi bien traditionnels qu'informatisés n'y voient habituellement qu'un seul sens.

Certes, les deux verbes ont des structures actancielles différentes, traduites dans deux formes propositionnelles distinctes dans le *definiendum*. Cela n'est pourtant pas une raison suffisante d'en faire deux lexies-vedettes. On pourrait en effet considérer que le second actant Y de **NAGERIb** (style de natation) correspond à une composante sémantique faible qui n'est pas systématiquement actualisée dans le discours, notamment lorsque X est un animal⁹. Dans cette optique, on aurait affaire à une seule acception **NAGERI** dont la forme propositionnelle serait la suivante :

L'être animé X ~ (en utilisant la technique Y).

En adoptant cette hypothèse, il faudrait proposer une paraphrase définitoire suffisamment large pour couvrir les deux cas. Cette solution a été envisagée à une étape antérieure de notre travail avec la définition suivante :

<CC label="se déplacer intentionnellement">L'être animé X se déplace</CC>

<PC role="milieu">en suspension dans l'eau</PC>

<PC role="gestuelle">par des mouvements de son corps (conformes à la technique Y)</PC>

⁹ Remarquons que le second actant Y correspond à une composante faible même lorsque l'on sépare les acceptions **NAGERIa** et **NAGERIb**, dans la mesure où un être humain peut très bien se déplacer en suspension dans l'eau sans que ses gestes forment un style quelconque.

Rappelons cependant que la démarche définitoire adoptée dans le cadre du RLF vise un triple objectif. Il s'agit de :

- i. décrire le sens,
- ii. identifier le sous-réseau des unités lexicales auxquelles la lexie-vedette est connectée,
- iii. situer la lexie-vedette dans le graphe que forme le lexique du français.

Si une acception commune **NAGERI** est satisfaisante en ce qui concerne le postulat i., elle est loin de répondre aux exigences ii. et iii. En effet, elle ne prend pas en compte les liens paradigmatiques et syntagmatiques (sous-section 4.2) qui structurent le lexique de la langue.

Or, en ce qui concerne les liens lexicaux, le verbe **NAGERIa** de l'exemple (1) et **NAGERIb** de (2) forment deux réseaux lexicaux sensiblement différents. Il suffit de penser aux exemples (9) et (10) dans la partie introductive de cette section : **NAGEUR** et **NAGEUSE** sont bien des noms de l'actant X mentionné dans la définition de **NAGERIb**, mais pas de celui de **NAGERIa**. Ainsi, dire d'un poisson rouge qu'il est un excellent nageur est toujours possible, mais relève plutôt d'une plaisanterie.

Le tableau 1 réunit les lexies connectées avec les acceptions **NAGERIa** et **NAGERIb** par des relations paradigmatiques. On désignera ces lexies sous le nom de dérivés sémantiques.

Tableau 1. Dérivés sémantiques des lexies **NAGERIa** et **NAGERIb**

| Type de lien paradigmatique | NAGERIa | NAGERIb |
|-----------------------------|-------------------------------------|--|
| quasi-synonymes | — | CRAWLER FLOTTER, BARBOTER, PATAUGER |
| nominalisation | NAGEIa | NAGEIb |
| adjectivation | NATATOIREa | NATATOIREb |
| noms pour l'actant X | <i>typiquement</i> : POISSON | NAGEURa, NAGEUSEa |
| nom pour l'actant Y | — | NAGEII ; BRASSE, BRASSE PAPILLON, CRAWL, DOS BRASSÉ, DOS CRAWLÉ, NAGE INDIENNE, PAPILLON |

Il est frappant de constater la disparité quantitative des dérivés sémantiques de ces deux acceptions. Le sens « humain » développe dans la langue un réseau de connexions nettement plus vaste que le sens « animal ». Ce dernier donne lieu à une conversion nominale, adjectivale et à un nom d'actant X.

Les lexies nominales **NAGEIa** et **NATATOIREa** sémantiquement construites sur **NAGERIa** dénotent le même fait que leur source de dérivation. Les manipulations que nous effectuons sur l'exemple (11) conduisent aux phrases (12) et (13) qui expriment un sens similaire.

- (11) Les poissons-pilotes ne guident pas les requins, comme on le pensait autrefois, ils profitent du flux hydrique créé par la *nage* du requin. (Internet)
 (12) Les poissons-pilotes ne guident pas les requins, comme on le pensait autrefois, ils profitent du flux hydrique créé par le requin qui *nage*.
 (13) Les poissons-pilotes ne guident pas les requins, comme on le pensait autrefois, ils profitent du flux hydrique créé par l'activité *natatoire* du requin.

Pour éviter la redondance, tout en indiquant clairement le lien paradigmatique, nous proposerons dès lors les définitions suivantes de ces dérivés sémantiques :

NAGEIa

~ de l'animal X

<CC label="action de se déplacer intentionnellement">Fait que l'animal X **NAGEIa**</CC>

NATATOIREa

[X] ~

<CC label="relatif au fait de se déplacer intentionnellement">Relatif au fait que l'animal X **NAGEIa**</CC>

De la même manière, la lexie **NAGERIb** servira de base pour définir ses dérivés sémantiques réunis dans la colonne de droite de notre tableau. Faute de place, nous laisserons au lecteur le plaisir de cet exercice.

Cette démarche rigoureuse conduit par ailleurs à éviter des rapprochements trop intuitifs. Ainsi, le nom **NATATION** ne figure pas parmi les dérivés sémantiques de **NAGERIb**, car on ne saurait définir ces lexies en partant de ce verbe. Pour s'en convaincre, il suffit de rappeler l'exemple (2) : **NAGERIb** la brasse vers le large, même vigoureusement, ne revient point à faire de la natation. La lexie **NATATION** semble connectée par un lien paradigmatique à **NAGERII** qui est un verbe de réalisation que nous qualifierions de sportive, elle-même sémantiquement construite sur **NAGERIb**.

Le travail de construction de définitions poursuivi dans le RLF est donc en fait un incessant va-et-vient entre la lexie-vedette et ses dérivés sémantiques. D'une part, la définition ouvre l'accès aux lexies connexes dans la mesure où elle permet à l'utilisateur de prévoir les dérivés sémantiques. On comprend dès lors à quel point il est important de spécifier, dans la paraphrase définitoire, tous les actants représentés dans la forme propositionnelle (voir section 3) et les rapports entre eux. C'est ainsi qu'il est possible d'identifier le sous-réseau d'unités lexicales dont le locuteur a besoin pour désigner un fait donné par la phrase qu'il pourra choisir parmi celles qui conviennent.

D'autre part, le lexicographe parcourt régulièrement le chemin en sens inverse, en allant du dérivé vers la lexie-vedette. En effet, une définition des lexies **NAGEa** et **NAGEb** ne peut se faire qu'en fonction de celles qui constituent leurs sources de dérivation sémantique. Cela signifie *a contrario* qu'en construisant la définition de la lexie source (en l'occurrence celles de **NAGERIa** et de **NAGERIb**), le lexicographe ne perd pas de vue le fait qu'elle sera au cœur des paraphrases définitives de ses dérivés sémantiques.

4.2. DÉFINITIONS ET LIENS SYNTAGMATIQUES

Une définition du RLF doit également permettre à l'utilisateur d'explorer les liens syntagmatiques propres à la lexie décrite. Or, l'analyse de la combinatoire syntaxique de **NAGERIa** et de **NAGERIb** fournit des arguments solides en faveur d'un traitement séparé des deux sens, *ipse eo* de deux paraphrases définitives distinctes.

On constate tout d'abord que le réseau syntagmatique de **NAGERIb** est nettement plus riche que celui de l'ULB **NAGERIa**. En ce qui concerne Jean, excellent nageur des exemples (9) et (10), il est possible d'apprécier ses talents en disant qu'il nage comme un poisson. Certes, on pourra toujours dire d'un serpent de mer qu'il **NAGEa** comme un poisson. Il s'agira néanmoins non pas d'une évaluation positive mais d'une comparaison de modes de déplacement.

Cela signifie que l'unité lexicale **NAGERIb** constitue la base des collocations formées avec des collocatifs *comme un poisson, comme un boulet, comme un caillou, comme un fer à repasser*, etc., incompatibles avec le sens **NAGERIa**. Or, une définition du RLF doit compter avec ce type de liens. En ce qui concerne **NAGERIb**, les CP "gestuelle" et "background" indiquent que le degré de maîtrise des gestes lié tant à l'apprentissage qu'à la pratique trouve une expression dans la langue. De même, le caractère inné de ce mode de déplacement chez un poisson bloque la possibilité de former la collocation *apprendre à nagerIa*, alors qu'elle n'est pas rare dans le cas de **NAGERIb**.

4.3. DÉFINITIONS, LIENS LEXICAUX ET STRUCTURE DU SOUS-RÉSEAU LEXICAL

La troisième exigence à laquelle répondent les définitions du RLF est celle de identifier le sous-réseau de la lexie-vedette. Avant de conclure nos propos, il est temps de présenter les graphes des deux lexies qui nous ont particulièrement intéressées dans ces pages. Les arcs continus marquent les liens paradigmatiques, ceux en pointillés indiquent les relations syntagmatiques. La disparité entre la densité des liens paradigmatiques et syntagmatiques dans les graphes (3) et (4) est désormais frappante.

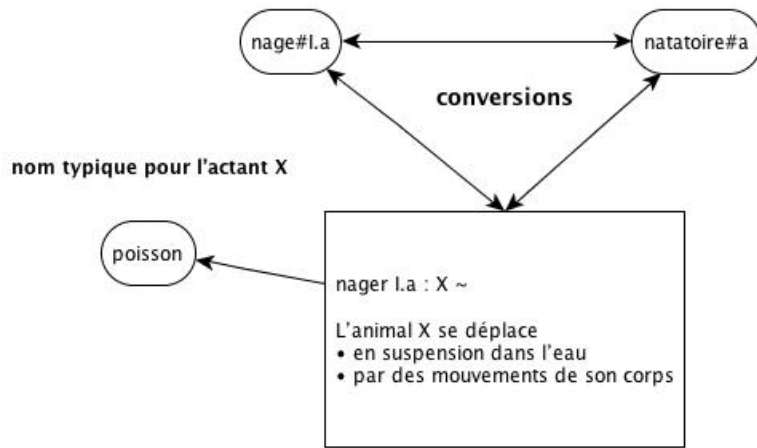


Fig. 3. Réseau de relations lexicales de NAGERIa

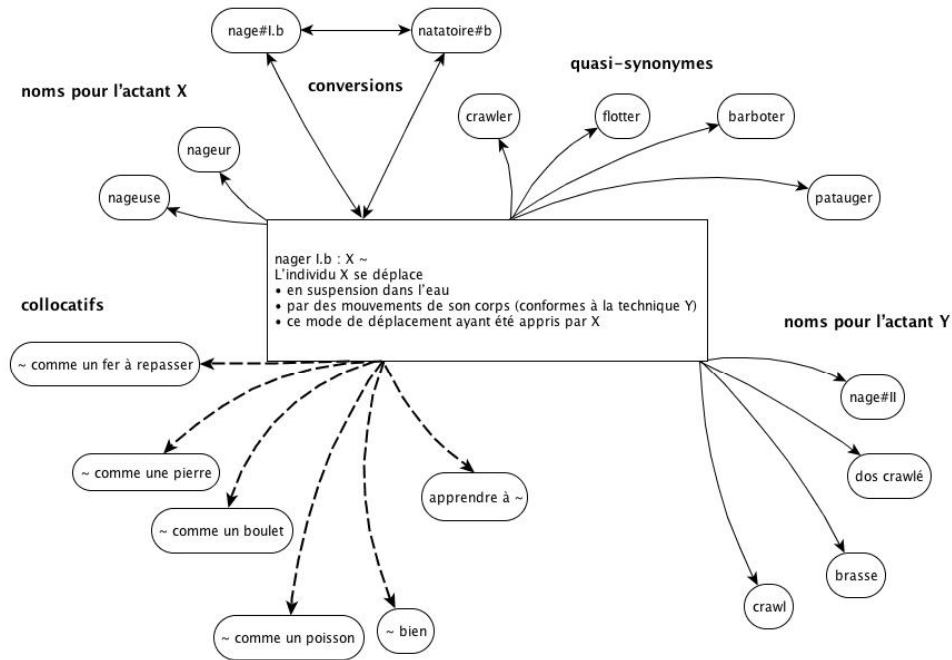


Fig. 4. Réseau de relations lexicales de NAGERIb

5. POUR CONCLURE

Définir en suivant fidèlement les liens lexicaux quelle que soit leur nature (cf. sections 3 et 4) est un défi qui – à notre connaissance – n’a pas encore été relevé dans une ressource offrant une large couverture du français. Au fil de ces pages, nous avons vu le nombre de contraintes à respecter pour construire les définitions qui tissent un réseau de liens reproduisant la structure du lexique, en l’occurrence celui du français. Elles sont, certes, très fortes mais débouchent en réalité sur une richesse d’informations inestimables pour tout utilisateur. Le linguiste y trouvera, entre autres, de quoi explorer les phénomènes de polysémie. Un élève de langue maternelle pourra explorer le réseau à la recherche du « mot juste » qui lui permettra d’éviter le verbe *être* omniprésent, mais invariablement mal vu par son professeur de français. Les apprenants étrangers de différents niveaux le parcourront à la recherche du collocatif manquant. À chacun son réseau ... lexical.

BIBLIOGRAPHIE

- LUX-POGODALLA, Veronika ; POLGUÈRE, Alain (2011): *Construction of a French Lexical Network : Methodological Issues. Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011. An ESSLLI 2011 Workshop*. Ljubljana, 54-61.
- MCCARTHY, Michael (1999): « What constitutes a basic vocabulary for spoken communication ? ». *English and Literature* 1: 233-249.
- MEL’ČUK, Igor Aleksandrovič et *alii* (1984, 1988, 1992, 1999): *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques*. Volumes I–IV. Montréal : Les Presses de l’Université de Montréal.
- MEL’ČUK, Igor Aleksandrovič (1996): « Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon ». In: WANNER, Leo [ed.]. *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam/Philadelphie : John Benjamins, 37-102.
- MEL’ČUK, Igor Aleksandrovič (1997): *Vers une linguistique Sens-Texte*. Leçon inaugurale. Paris : Collège de France.
- MEL’ČUK, Igor Aleksandrovič ; CLAS, André ; POLGUÈRE, Alain (1995): *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve : Duculot.
- MEL’ČUK, Igor Aleksandrovič ; CLAS, André ; POLGUÈRE, Alain (2007): *Lexique actif du français. L’apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. Bruxelles : De Boeck.
- MILIČEVIĆ, Jasmina (2006): « A Short Guide to the Meaning-Text Linguistic Theory ». *Journal of Koralex* 8: 187-233.
- POLGUÈRE, Alain (2011a): « Classification sémantique des lexies fondée sur le paraphrasage ». *Cahiers de Lexicologie* 98: 197-211.
- POLGUÈRE, Alain (2011b) : « Perspective épistémologique sur l’approche linguistique Sens-Texte ». In: *L’architecture des théories linguistiques, les modèles et leurs interfaces*. Mémoires de la Société de Linguistique de Paris, t. XX: 79-114.
- POLGUÈRE, Alain (*à paraître*): « Lexicographie des dictionnaires virtuels ». In: *Meaning, Text and Other Exiting Things. Hommages à Igor Mel’čuk*.

- POLGUÈRE, Alain ; SIKORA, Dorota (*à paraître*): « Modèle lexicographique de croissance de vocabulaire fondé sur un processus aléatoire, mais systématique ». In : *Enseigner le lexique. Pratiques sociales objets à enseigner et pratiques d'enseignement*. Metz : AiRDF.
- SELVA, Thierry ; VERLINDE, Serge ; BINON, Jean (2003): « Vers une deuxième génération de dictionnaires électroniques ». *Traitement Automatique des Langues (T.A.L.)* 44(2): 177-197.
- WIERZBICKA, Anna (1996): *Semantics. Primes and Universals*. Oxford – New York : Oxford University Press.

Ressources électroniques

URL: <<http://www.cnrtl.fr>>. (Last accessed at: 8 juin 2012)

URL: <<http://www.frantext.fr>>. (Last accessed at: 8 juin 2012)

URL: <<http://olst.ling.umontreal.ca/dicopop/>>. (Last accessed at: 8 juin 2012)

URL: <<http://olst.ling.umontreal.ca/dicouebe/>>. (Last accessed at: 8 juin 2012)

Modélisation de connaissances lexicales dans une ressource lexicale

Dorota Sikora, Université du Littoral-Côte d'Opale

Résumé : Découvrir le lexique d'une langue, c'est à la fois accéder au sens et maîtriser le fonctionnement d'unités lexicales (lexies) dans les relations (formelles, paradigmatiques, syntagmatiques, associatives, etc.) qu'elles tissent entre elles. Comment décrire, dans une ressource dictionnaire, cet énorme réseau, tout en apportant des informations relatives à des lexies particulières ? L'article présente la structure informationnelle de la base de données Réseau Lexical du Français (RL-fr), ressource qui constitue l'implémentation du modèle de type système lexical (Polguère 2014, 2016). Dans une deuxième partie, nous discutons des exploitations possibles du RL-fr.

Abstract : Exploring lexicon implies access to the meaning of lexical units, as well as mastery of a whole set of formal, paradigmatic, syntagmatic or simply associative relations between them. Thus, how can a lexical resource manage to account for the global structure of the lexicon on the one hand, providing detailed lexicographic description of single lexical units on the other? With such a double perspective, the paper discusses the data structure encoded in the French Lexical Network (RL-fr), resource that implements lexical system model (Polguère 2014, 2016). Further, a handful of its possible applications are considered.

Mots-clés : connaissances lexicales, ressources, modèles lexicaux

Keywords : lexical knowledge, resources, lexical system

1. Introduction

Cet article présente le *Réseau Lexical du Français* (désormais RL-fr), base de données lexicales à large couverture pour le français¹. Sans être un dictionnaire électronique, le RL-fr constitue néanmoins une source d'informations riche et structurée, qui offre une multitude d'exploitations pour les utilisateurs humains, qu'ils soient linguistes, enseignants, étudiants ou amoureux de la langue, comme pour le traitement automatique.

Nous en proposons un tour aussi complet que possible dans cet article. La partie 2 présentera rapidement la structure globale de la ressource et l'état de son développement. Le modèle lexical dont le RL-fr constitue l'implémentation y sera également exposé. La section 3 permettra de décrire l'organisation des données encodées dans la base, en y accédant par l'entrée des artistes² : nous monterons qu'un travail lexicographique consistant non pas à rédiger des articles de dictionnaire, mais à tisser un réseau, implique une démarche spécifique qui conduit parfois à des choix inédits. Nous adopterons ensuite le point de vue de l'utilisateur pour discuter, en section 4, quelques applications et exploitations du RL-fr, avant de conclure dans la cinquième et dernière section.

¹ Le projet RELIEF (2011-2014) a été mené sous la direction scientifique d'Alain Polguère (ATILF-CNRS et Université de Lorraine) et administrative de Jean-Marie Pierrel (ATILF-CNRS et Université de Lorraine).

2. Présentation du Réseau Lexical du Français

La construction du RL-fr s'inscrit dans une double perspective. D'une part, comme toute ressource dictionnaire, celle-ci offre une description des éléments du lexique : actuellement, environ 28 000 unités lexicales (lexies) y ont été intégrées². D'autre part, la structure des données encodées doit rendre compte de l'organisation du lexique en tant qu'entité constitutive de la langue. Le RL-fr s'appuie ainsi sur une certaine conceptualisation de ce module et sur la modélisation qui en est issue : celle d'un système lexical que nous présentons en 2.1. La sous-section 2.2. décrit la structure globale de la ressource en tant qu'implémentation de ce modèle.

2.1. Le RL-fr en tant que système lexical

Un système lexical (Polguère 2014, 2016) est un modèle, qui par l'organisation des informations encodées rend compte à la fois de la nature relationnelle et du caractère multidimensionnel du réseau que forme le lexique d'une langue. Les caractéristiques suivantes du module lexical sont postulées :

1. Dans un système lexical, modèle élaboré dans le prolongement des travaux réalisés dans le cadre de la *Lexicologie Explicative et Combinatoire-LEC* (Mel'čuk & al. 1984, 1988, 1992, 1999 ; Mel'čuk & Polguère 2007)³, le réseau se construit entre des lexies (unités lexicales) considérées comme éléments de base du lexique et, par conséquent, l'objet de l'analyse et de l'activité lexicographique. Un système lexical est ainsi un réseau construit essentiellement⁴ à partir de liens entre des lexies, et non pas entre des ensembles de lexies tels que des synsets dans *WordNet* ou des vocables, c'est-à-dire des regroupements d'acceptions reliées par des relations de copolysémie.
2. L'organisation interne du lexique, autrement dit sa structure logique, repose sur des relations de nature diverse (paradigmatiques, syntagmatiques, sans oublier les liens d'inclusion formelle entre des unités polylexicales et les lexèmes qui les composent, ou parfois de simples associations, cf. la section 2.2. ci-dessous), dont aucune n'a un statut privilégié. Ce traitement démocratique de tous les liens repérés dans le lexique distingue un système lexical d'un réseau de type *WordNet*, structuré par des relations d'inclusion catégorielle entre des synsets (ensemble de quasi-synonymes).
3. Les propriétés structurelles d'un système lexical sont celles d'un graphe petits-mondes. Le nombre de sommets y est élevé, mais ces nœuds sont dans l'immense majorité des cas faiblement connectés : deux liens en moyenne les relient à l'ensemble du graphe. Cependant, un nombre relativement restreint d'entre eux est caractérisé par un taux de connectivité très fort : ce sont des sommets-carrefours, essentiellement des collocatifs tels que FAIRE II.1, qui sert de verbe support à de très nombreuses lexies aussi éloignées sémantiquement que FOOT I, VAISSELLE II ou ACTE DE PRESENCE. La moyenne du plus court chemin qui mène de n'importe quel sommet à un autre est faible. L'analyse topologique d'un tel graphe révèle enfin que des sommets qui correspondent aux lexies sémantiquement proches se regroupent dans des *clusters*.

² Le 12 mars 2017 le nombre de lexies décrites était très exactement de 27 725, mais il augmente quotidiennement.

³ Les conventions d'écriture de la LEC seront adoptées tout au long de ce texte : les lexies seront notées par de petites majuscules et accompagnées, en cas de polysémie, de numéros indiquant leur place dans la structure du vocable.

⁴ D'autres types d'entités y sont néanmoins intégrées (cf. section 2.2.).

2.2. Structure globale du RL-fr

La structure du graphe RL-fr peut être visualisée. Il est possible dès lors d'en extraire une sorte de portrait-robot du lexique français⁵. La figure 1 ci-dessous présente le RL-fr tel qu'il était en juillet 2015⁶ (à noter que le nombre de composants isolés visibles à droite a été sensiblement réduit par l'implémentation systématique de liens (cf. section 3.4.)).

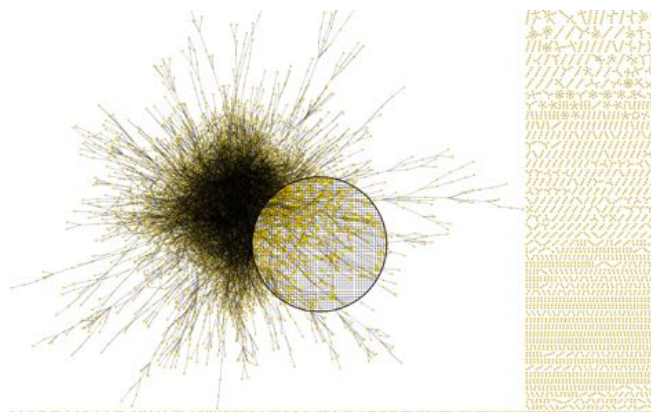


Figure 1. Graphe lexical RL-fr (juin 2015).

Si, pour une large majorité, les nœuds du RL-fr sont des lexies, d'autres entités aux statuts lexicaux divers y sont également intégrées : des constructions lexicalisées telles que *à [Y] près* ('*ajuster au centimètre près*') où une position syntaxique doit être saturée par une lexie, des constructions syntaxiques, c'est-à-dire des configurations peu ou pas encore lexicalisées, telles que *Et [X] de [Y]* ('*Et l'auteur de conclure*'), ou des proverbes⁷, etc. (cf. section 3.5.). Les sommets sont connectés par environ 73 000 liens de nature diverse. L'immense majorité d'entre eux – plus de 57 000 – sont des liens de fonctions lexicales (FL, cf. Mel'čuk & al. (1995), Polguère (2007)) : nous en reparlerons dans la section suivante, mais précisons d'ores et déjà que les FL modélisent les relations paradigmatiques et syntagmatiques récurrentes dans le lexique.

La figure 2 représente un sous-graphe du RL-fr visualisé *via* l'interface d'utilisateur actuellement en cours de développement⁸. Il est organisé autour de la lexie AUX ANGES, qui était objet de notre requête. Cette locution, dont le signifiant prend la forme d'un syntagme prépositionnel, connaît essentiellement des emplois adjectivaux : parmi ses quasi-

⁵ Bien qu'à une moindre échelle, des travaux d'implémentation d'autres réseaux lexicaux sont actuellement menés, notamment pour l'anglais, l'arabe, le coréen, l'espagnol et le russe.

⁶ Compte tenu de sa croissance, aussi bien pour ce qui est du nombre de nœuds lexicaux que de la densité des liens, une visualisation récente devient difficilement lisible dans un format réduit. C'est la raison pour laquelle nous avons choisi de reproduire ici le graphe tel qu'il était en 2015, plutôt que sa version actuelle.

⁷ Dans la mesure où il ne s'agit pas de lexies (cf. Mel'čuk 2013), nous n'employons pas de petites majuscules pour les transcrire.

⁸ Développée par *KodexLab* (<<http://kodexlab.com>>), cette interface est destinée à donner accès au RL-fr sous licence libre.

synonymes, on trouve donc les adjectifs HEUREUX I et RAVI, ainsi que la locution SUR UN NUAGE⁹.

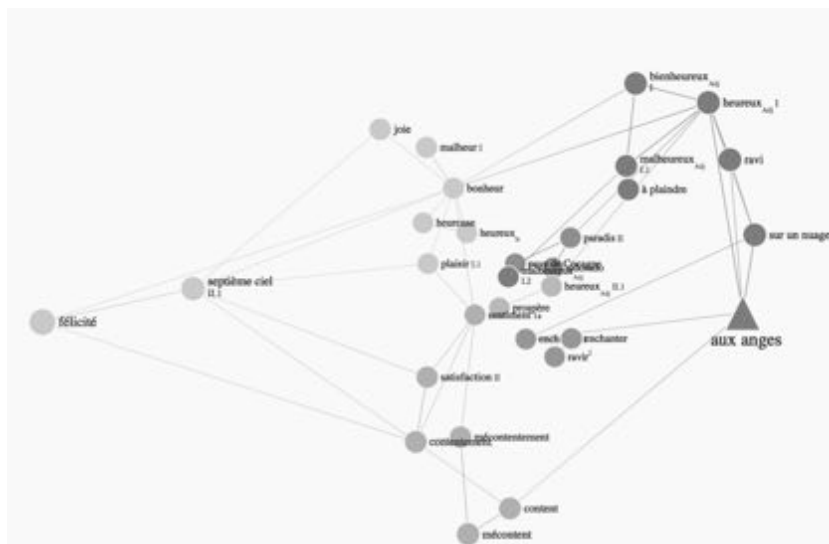


Figure 2. Sous-graphe autour de la lexie AUX ANGES.

3. Structure de l'information lexicale

Lorsque l'on construit un graphe de lexies, la première étape consiste à les identifier en tant que nœuds du réseau. Autrement dit, en créant l'entrée AMOUREUX, MOUSTIQUE ou PEUR, il est essentiel de repérer toutes les lexies que recouvrent ces signifiants. Il faut en effet dégager la polysémie d'un vocable pour procéder ensuite à l'analyse et à la description de chacune des acceptions. Cette section suivra l'ordre des opérations : nous présenterons le traitement de la polysémie dans le RL-fr, ainsi que l'analyse des formes qui, comme AMOUREUX, correspondent aux lexies adjectivales et nominales. Nous observerons ensuite le contenu informationnel d'un nœud lexical.

3.1. Analyse et modélisation de la polysémie

En tant que graphe de lexies, le RL-fr est, à notre connaissance, la seule ressource qui modélise les relations de copolysémie, c'est-à-dire celles qui relient les acceptions d'un vocable. Un vocable polysémique forme ainsi un réseau dont les arcs représentent les mécanismes d'extension de sens. La figure 3 ci-dessous montre la polysémie de MOUSTIQUE, qui comprend quatre acceptions : la lexie de base MOUSTIQUE I 'insecte', est la source de deux extensions construites par métaphore MOUSTIQUE II ('enfant') et MOUSTIQUE III.1 ('catégorie sportive'). Cette dernière se voit associer une métonymie lexicalisée MOUSTIQUE III.2, dénotant un individu qui appartient à cette catégorie sportive.

⁹ Nous considérons en effet que les locutions AUX ANGES et SUR UN NUAGE, analysées parfois comme des locutions verbales ETRE AUX ANGES/SUR UN NUAGE (Vaguer 2010), se limitent aux seuls syntagmes prépositionnels ; les verbes ETRE, SEMBLER, METTRE sont des collocatifs sélectionnés par la base. Pour une discussion de problèmes liés à l'identification des locutions et leur nature, voir Sikora (2016).

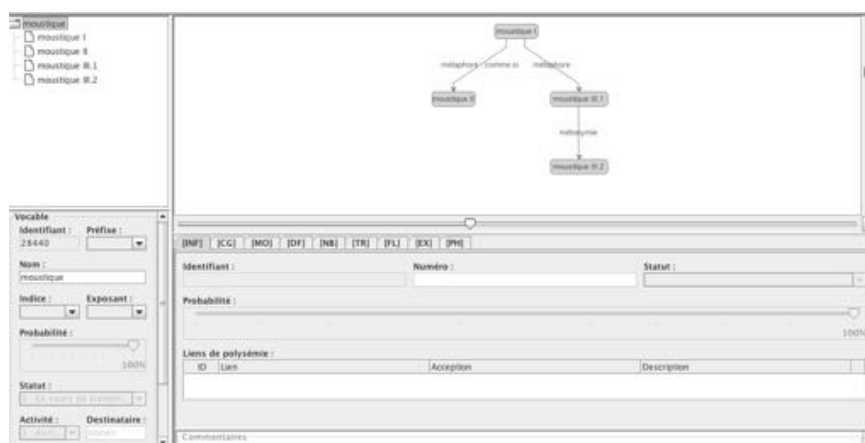


Figure 3. Modélisation de la polysémie dans le RL-fr : vocable MOUSTIQUE.

L'encodage explicite de la polysémie et des relations qui la structurent différencie le RL-fr des autres réseaux lexicaux tels que *WordNet* (<<http://wordnet.princeton.edu/>>) ou le *Dictionnaire des Synonymes* (<<http://www.crisco.unicaen.fr/des/>>), où celles-ci doivent être restituées par une réanalyse de données¹⁰.

3.2. Identification des lexies : analyse de caractéristiques grammaticales

Intégrer une lexie en tant que nœud du réseau suppose de la connecter d'une manière appropriée, conforme à sa nature, à d'autres sommets. C'est ainsi que la description lexicographique qui en sera proposée pourra rendre compte non seulement de son sémantisme, mais également de son fonctionnement dans le lexique. Le cadre théorique et méthodologique adopté conduit dès lors à des choix susceptibles de bousculer les habitudes des usagers de dictionnaires traditionnels, que ce soit en version électronique ou papier. La figure 4 ci-dessous reproduit l'entrée AMOUREUX dans le *Petit Robert* électronique.

L'information grammaticale sur la (les ?) partie(s) du discours est associée au signifiant présenté dans sa graphie et dans sa prononciation. Elle indique que ce signifiant correspond à la fois à des signifiés adjectivaux et nominaux, qui par ailleurs ne sont pas clairement séparés dans les paraphrases définitionnelles. Une telle organisation de l'information conduit à penser – notamment dans un contexte scolaire – que les lexies correspondant à ce signifiant sont de nature instable, tantôt adjectivale, tantôt nominale, susceptibles de fléchir en genre et en nombre.

Une requête portant sur AMOUREUX conduira l'utilisateur du RL-fr à découvrir un mini-réseau de sept lexies. Trois d'entre elles sont des adjectifs. Les exemples (1) à (4)¹¹ illustrent respectivement les sens de AMOUREUX I.1 ('[X] qui éprouve de l'amour pour Y'), AMOUREUX I.2 ('[X] relatif à l'amour') et AMOUREUX II ('[X] qui a un goût marqué pour Y').

¹⁰ Pour une discussion, on pourra se reporter à Polguère (2016).

¹¹ Tous les exemples, parfois raccourcis, sont extraits du RL-fr (cf. la section 4.3. ci-dessous).

Chacun d'entre eux s'est vu attribuer une table morphologique propre aux adjectifs, c'est-à-dire comprenant quatre formes flexionnelles selon le genre et le nombre.



Figure 4. Entrée AMOUREUX dans le Petit Robert électronique 2011.

Dans le cas des adjectifs, le signifiant AMOUREUSE dans l'exemple (2) recouvre simplement une forme du féminin.

- (1) Il était amoureux de cette fille.
- (2) Lynda est amoureuse d'un autre.
- (3) C'est par dépit amoureux qu'il a fait cela.
- (4) Amoureux des beaux produits de qualité, nous sommes restés fidèles aux techniques traditionnelles du cousu main.

Dans les exemples (5) à (8), toutes les lexies sont au contraire de nature nominale : AMOUREUX I ('individu X, qui éprouve de l'amour pour Y') et AMOUREUX II ('individu X, qui a un goût pour Y'). Elles sont reliées par une relation de quasi-synonymie respectivement à AMOUREUSE I ('individu X, de sexe féminin, qui éprouve de l'amour pour Y') et AMOUREUSE II ('femme qui a un goût pour Y'). En effet, les noms masculins, sous-spécifiés selon le genre, subissent un sens générique : on les considère comme sémantiquement moins riches. Comparativement, le sens des noms féminins comprend une information supplémentaire sur le sexe de l'individu dénoté¹². Le signifiant AMOUREUSE correspond alors à deux signifiés distincts, bien que reliés par quasi-synonymie aux noms masculins correspondants. Compte tenu de leur nature nominale, toutes ces lexies sont porteuses d'un genre grammatical, alors que leur flexion repose seulement sur la catégorie du nombre.

- (5) L'amoureux éconduit ayant annoncé son intention de se venger, la gendarmerie a préféré l'interpeller.
- (6) Lonyon était un amoureux de la musique.
- (7) Quentin veut se marier avec Hélène, son amoureuse.
- (8) Travailler et vivre en région parisienne, pour une amoureuse de la nature comme moi, est une véritable souffrance.

Ce raisonnement conduit à intégrer au RL-fr trois mini-réseaux d'acceptions copolysèmes correspondant aux vocables AMOUREUX_{Adj}, AMOUREUX_N, AMOUREUSE.

¹² Pour une discussion détaillée de cette relation sémantique récurrente dans le lexique, on pourra se reporter à Delaite & Polguère (2013).

3.3. Informations définitionnelles dans un nœud du réseau

Un nœud lexical renferme tout un ensemble d'informations décrivant le sens de la lexie vedette : une étiquette sémantique (Polguère 2011), la forme propositionnelle (FP) et – à terme – une définition. A ce stade du travail, seules les deux premières sont encodées pour toutes les lexies. Une trentaine de définitions seulement ont été élaborées. Elles apparaissent sous une forme classique, organisée autour d'une composante centrale (CC) exprimant le genre prochain (cf. figure 5 'sentiment 1a désagréable de X face à Y') et de composantes périphériques (CP) précisant les différences spécifiques¹³.

En l'absence de définition (cf. plus bas les figures 7 et 8, présentant respectivement la locution NOYER LE POISSON et le lexème MOBYLETTE), chaque lexie est caractérisée à l'aide d'une étiquette sémantique (Polguère 2011) : paraphrase minimale normalisée, qui renvoie à une classe sémantique inscrite dans une hiérarchie. Ainsi, 'sentiment' que l'on repère dans la figure 5 pour PEUR 1a, est une étiquette-fille de 'état psychique', incluse dans une classe plus générale des 'états', elle-même comprise dans celle de 'faits'. Les sens de NOYER LE POISSON (figure 7) est classé parmi les comportements, et celui de MOBYLETTE (figure 8) porte l'étiquette 'véhicule'.

Malgré sa forme sobre, qui ne remplace pas une définition à proprement parler telle qu'elle est illustrée pour PEUR 1a, l'étiquette sémantique constitue un outil descriptif puissant. Son choix repose tout d'abord sur une analyse du sens lexical, conduisant notamment à différencier les 'sentiments' (PEUR 1a) des 'états d'esprits' (tels que PRESSEMENT). Deuxièmement, lorsqu'elle est attribuée à un quasi-prédictat¹⁴, l'étiquette véhicule une information sur le nombre potentiel et la nature des actants, ainsi que sur les relations entre eux : une lexie classée parmi les 'sentiments' est le plus souvent bi- ou triactancielle, ce qui est explicité dans la forme propositionnelle (FP) : PEUR 1a de X causée par Y (cf. figure 4), GRATITUDE de X envers Y pour Z.

On remarque dans la figure 5 que malgré un apport informationnel non négligeable, l'étiquette sémantique et la forme propositionnelle se situent du côté du *definiendum* qui doit être décrit dans la paraphrase définitoire par la suite. Le *definiens* précise – si besoin est – le type sémantique des actants, ce typage (surligné en gris) faisant partie des présuppositions. La relation prédicative entre les actants est explicitée dans la CC et précisée dans les CP. Si dans l'affichage actuellement disponible, elle prend une forme textuelle, la définition lexicographique du RL-fr est en réalité non-linéaire. Elle constitue un réseau sémantique (Mel'čuk & Polguère 2016), tissé autour de nœuds sémantiquement dominants : 'sentiment désagréable', 'dangereux', 'se protéger', 'réaction physique'.

Les informations sémantiques forment ainsi un ensemble structuré et cohérent, dans lequel la paraphrase définitoire n'a plus une position centrale. Elle rend compte du sens, tout en s'inscrivant dans une structure complexe de données, avec lesquelles elle se doit de rester cohérente. L'organisation interne de la paraphrase définitoire dépend de l'étiquette sémantique et de la FP : un sentiment se caractérise en effet par une intensité, se laisse évaluer selon toute une échelle de ressentis positifs ou négatifs, possède des manifestations physiques, etc.

¹³ La structure de définitions est celle préconisée par la LEC. Elle est soumise aux mêmes principes d'élaboration. Néanmoins, compte tenu de la nature du modèle de système lexical dans lequel elle doit s'insérer, la définition a récemment fait l'objet d'une réflexion et de mises à jour importantes (cf. Mel'čuk & Polguère 2016).

¹⁴ Un quasi-prédictat (cf. Mel'čuk & Polguère 2008) est un nom qui dénote une entité (par exemple un artefact tel que MOBYLETTE) et qui sélectionne des actants. Le premier actant d'un quasi-prédictat de type 'véhicule' (qui est bien un artefact) est son utilisateur.



Figure 5. Informations définitionnelles du nœud PEUR 1a.

De plus, une définition, tout comme la FP qui spécifie les actants, servent de référence pour l’encodage des fonctions lexicales (cf. § 3.4.), c’est-à-dire à l’implémentation des arcs du réseau.

3.4. Fonctions lexicales : relations structurantes d’un système lexical

Les fonctions lexicales jouent un rôle particulièrement important dans le graphe RL-fr. Par leur nombre, qui a dépassé 57 000 sur la quantité globale des 73 000 liens (cf. § 2), elles constituent la charpente du système lexical RL-fr, tout en assurant sa croissance. En effet, la création de chaque nouveau nœud doit conduire à l’implémentation d’au moins une FL pour éviter l’apparition de composant isolé dans le graphe. Si la lexie ciblée n’est pas présente dans le réseau, il convient de la créer, en la connectant à d’autres nœuds lexicaux. Ainsi, PEUR 1a a conduit à créer son dérivé adverbial relatif au premier actant (FL Adv₁) PEUREUSEMENT, lié à son tour à CRAINTIVEMENT par une relation de quasi-synonymie (FL QSyn). Des FL syntagmatiques dont la lexie vedette est la source pointent vers ses collocatifs. PEUR 1a est ainsi relié entre autres à ses verbes supports AVOIR II, EPROUVER, RESENTIR¹⁵, etc.

Dans la figure 6, les cibles de FL dont PEUR 1a est la source s’affichent dans la partie centrale. La fenêtre ‘Fonctions lexicales’ montre l’outil d’encodage de ces liens. Les collocations les plus fréquentes, comme celle que forme PEUR 1a avec son verbe support AVOIR II, sont illustrées dans la zone d’exemples extraits de trois corpus (*Frantext*, un corpus français du Web et celui de presse régionale) ou repérés par veille lexicale¹⁶.

¹⁵ Notons que le lien avec ces verbes peut être inféré à partir de l’étiquette sémantique ‘sentiment’ attribuée à PEUR 1a.

¹⁶ Nous reviendrons à la question de corpus dans la section 4 ci-dessous, consacrée entre autres aux ressources qu’il est possible de générer à partir du RL-fr.

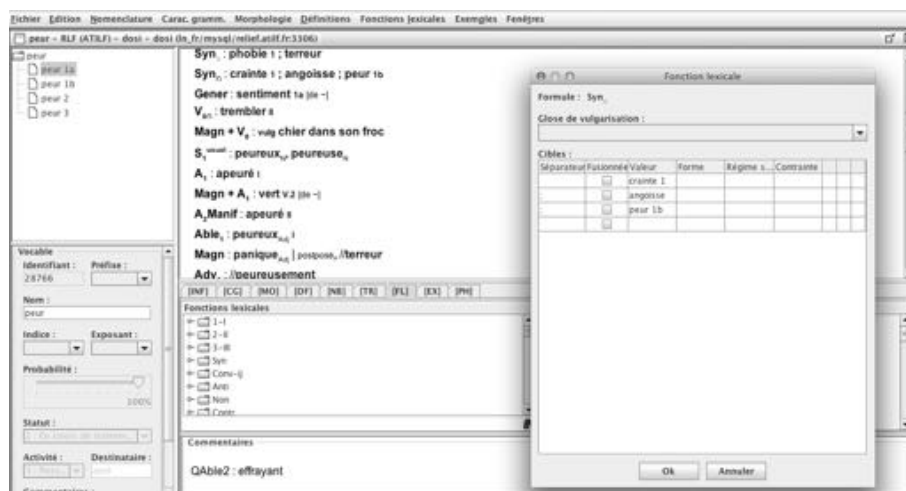


Figure 6. Encodage de liens de fonctions lexicales dans le RL-fr.

3.5. Structures lexico-syntaxiques des unités polylexicales

Tout comme les lexèmes, les unités polylexicales aux statuts divers (lexies composées, constructions, locutions, clichés, pragmatèmes¹⁷, proverbes, etc.) forment des nœuds du système. Sur le plan sémantique, leur description comprend les mêmes éléments : une locution verbale telle que NOYER LE POISSON est porteuse d'une étiquette sémantique qui la situe parmi les comportements, ses actants sont explicités dans la FP. Elle est liée à toute une série de quasi-synonymes locutionnels et lexémiques (cf. la figure 7).

Formellement, les unités polylexicales sont néanmoins des syntagmes et cette spécificité de leurs signifiants est également prise en charge dans le RL-fr. Au moment de l'attribution des caractéristiques grammaticales – locution verbale pour NOYER LE POISSON –, elles se voient associer une structure lexico-syntaxique (SLS) construite à partir de lexèmes que contient le syntagme¹⁸. Dans la mesure où les locutions n'autorisent pas un calcul compositionnel, ces liens sont considérés comme asémantiques. Leur rôle est de mettre en évidence l'organisation syntaxique, qui pourra dès lors être analysée en tant que réseau de dépendances entre lexèmes (Pausé, à paraître). La figure 7 présente l'encodage de SLS dans le RL-fr¹⁹.

¹⁷ Les pragmatèmes sont des phrasèmes (syntagmes) compositionnels qui – *Chien méchant, A consommer de préférence avant...* – s'emploient dans des situations spécifiques et sont contraints par elles. Pour une discussion approfondie, voir Mel'čuk (2013).

¹⁸ Plus de 3 300 phrasèmes ont déjà reçu une SLS.

¹⁹ Ce travail se poursuit et fait objet d'une recherche doctorale 'Structure lexico-syntaxique des locutions françaises et incidence sur leur combinatoire' menée par M.S. Pausé.



Figure 7. Encodage des structures lexico-syntaxiques des unités polylexicales.

3.6. Structure de l'information dans le Réseau Lexical du Français : bilan

Le travail lexicographique est focalisé sur les lexies ; un nœud du RL-fr possède une structure non-atomique d'informations structurées et interdépendantes, qui le positionnent dans le graphe. Il est cependant essentiel de préserver une vision globale du réseau, qui s'étoffe à chaque nouvelle implémentation, que ce soit d'un sommet lexical ou d'un arc. Tout comme les connaissances humaines, le RL-fr est de nature dynamique : des liens dont la validité est mise en question peuvent être réanalysés, puis défaits et reportés vers d'autres cibles. Cette double exigence imposée par la conceptualisation du lexique et une analyse linguistique multidimensionnelle exclut une démarche collaborative de type *crowdsourcing* sur laquelle s'appuie entre autres le réseau *Jeux de Mots* (<<http://www.jeuxdemots.org/jdm-accueil.php>>).

4. Quelques domaines d'application du RL-fr

La richesse d'informations lexicales qu'offre le RL-fr, ainsi que l'architecture des données (stockage dans des tables séparées d'une base SQL) permettent d'envisager une multitude d'applications et d'exploitations. Dans cette section, nous présentons le mode de consultation du RL-fr pour un usager non spécialiste (en 4.1.), l'intérêt de la ressource pour l'enseignement, tout particulièrement pour la didactique du lexique (4.2.), avant de décrire un corpus sémantiquement annoté, qui se développe parallèlement au graphe lexical (4.3.).

4.1. Navigation dans un réseau lexical

L'interface d'utilisateur permet de visualiser la structure du réseau lexical, en extrayant des sous-graphes locaux, comme celui dans la figure 2 (§ 2.1.) : une requête portant sur la locution AUX ANGES ramène tout une grappe de lexies auxquelles celle-ci est connectée par diverses relations. Ce qui compte pour l'utilisateur cependant, c'est également – et sans doute prioritairement – les informations auxquelles il accèdera, ainsi que le mode de consultation proposée. En sélectionnant MOBYLETTE dans la fenêtre d'interrogation de l'interface, on déclenche l'extraction d'un sous-graphe construit automatiquement par calcul de proxémie (Gaume & Duvignau 2004). Le nœud lexical qui fait objet de la requête a une forme triangulaire. C'est un objet tridimensionnel que l'utilisateur peut manipuler : il lui suffit de positionner le curseur sur un

nœud pour retourner le sous-réseau, en ramenant au premier plan les éléments du second. La figure 8 montre le sous-graphe affiché dans l'outil de visualisation.

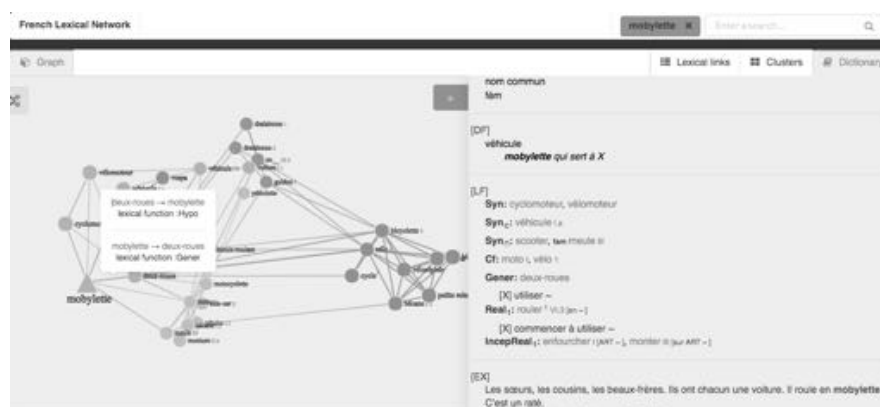


Figure 8. Visualisation des informations lexicales dans l'interface d'utilisateur.

La lexie MOBYLETTE constitue une porte d'accès au réseau. L'utilisateur peut dès lors en explorer le contenu informationnel affiché dans le volet de droite, mais il peut également entamer tout un parcours selon ses besoins communicatifs ou selon sa curiosité. Il est possible de naviguer d'un sommet à l'autre, par exemple de MOBYLETTE à DEUX-ROUES, en suivant les arcs. Ces derniers sont étiquetés de manière à expliciter la nature et l'orientation du lien qui connecte les deux nœuds. Ainsi, comme on le lit dans la figure 8, MOBYLETTE est relié à DEUX-ROUES, son hyperonyme, par un lien de fonction lexicale *Gener* (celui qui relie un hyponyme à son terme générique). Dans la mesure où cette relation n'est pas symétrique, il est indiqué que les deux nœuds lexicaux sont également connectés par un lien inverse partant de l'hyperonyme DEUX-ROUES dont l'hyponyme MOBYLETTE est la cible. La navigation dans le réseau peut être planifiée ou entièrement aléatoire. A chaque étape, il est possible de faire une halte, en cliquant sur un nœud lexical pour accéder aux informations qu'il contient.

Cependant, l'aspect quelque peu formel des informations disponibles peut présenter un obstacle pour un non-spécialiste. La caractérisation des connexions à l'aide de fonctions lexicales (aussi bien dans un sous-réseau de la partie gauche de la figure 8 qu'à l'intérieur du nœud à droite) est peu lisible. Un important travail reste à effectuer pour faciliter l'exploitation des données réunies par un public plus large. Notons qu'il a déjà été entamé en ce qui concerne les formules de vulgarisation des FL. En effet, dans la partie dictionnaire de la figure 8, les FL syntagmatiques *Real₁* et *IncepReal₁*, qui relient MOBYLETTE respectivement à *ROULER V1.3* et *ENFOURCHER I*, *MONTER III*, sont paraphrasées en des termes plus courants. *Rouler en mobylette*, c'est l'utiliser, *enfourcher une mobylette* ou *monter sur une mobylette*, c'est commencer à l'utiliser. Nous verrons dans la section suivante qu'une exploration systématique des liens de FL encodées dans une ressource à large couverture peut apporter des bénéfices pédagogiques non négligeables.

4.2. Exploitations pédagogiques d'un système lexical

Les informations lexicales disponibles dans le graphe RL-fr offrent des exploitations pédagogiques multiples, notamment pour l'enseignement du vocabulaire en langue maternelle (L1) et en langue étrangère (L2). Certaines démarches et activités ont déjà été testées et

présentées (Sikora 2014, 2016), d'autres pourront être conçues en fonction de besoins des élèves et de contraintes imposées par les programmes. Plutôt que de citer des exemples d'exercices, il nous semble important de proposer une réflexion sur une démarche globale conduisant à une pédagogie organisée en fonction de l'objet sur lequel elle porte : le lexique d'une langue.

La consultation d'un dictionnaire a pour objectif de découvrir un sens jusque là inconnu. L'intérêt porte alors sur une lexie particulière et le recours à une ressource permet de trouver l'information nécessaire. Si en plus celle-ci est mémorisée et réutilisée par la suite, on pourra considérer que le bénéfice est quasi immédiat. Le RL-fr apporte des réponses à des questionnements ponctuels, mais la structure de ses données et le mode de consultation qu'il offre sont surtout à considérer à moyen et long terme.

La navigation dans le graphe permet tout d'abord de prendre progressivement conscience de la structure du module lexical de la langue et de sa nature relationnelle, tout en favorisant chez l'élève une attitude d'explorateur. L'utilisateur est conduit à identifier plus ou moins consciemment ce qui en est l'unité de base : la lexie. Cette prise de conscience est en effet indispensable – tant pour les enseignants que pour les élèves – pour repérer les relations récurrentes qui structurent le lexique : celles de copolysémie d'abord, puis les FL paradigmatiques et syntagmatiques qui se tissent entre les acceptions. Compte tenu de la taille du graphe lexical RL-fr, il est certes difficile de le visualiser dans son ensemble (cf. § 2.), mais une observation systématique des sous-graphes qui se forment autour de telle ou telle lexie permet de repérer des liens et des structures récurrents.

Deuxièmement, exploiter en classe un outil de type système lexical revient à adopter une position résolument lexicaliste, en plaçant le lexique au cœur des démarches didactiques. Il ne s'agit aucunement d'abandonner les autres dimensions et faits de langue, mais d'organiser leur traitement à partir du module lexical. Notre discussion des informations disponibles dans un nœud du RL-fr montre qu'il est possible, et même souhaitable, d'articuler l'enseignement de la grammaire sur celui du lexique, puisque les caractéristiques grammaticales telles que la partie du discours et le paradigme flexionnel sont celles d'une lexie, et non pas d'un vocable.

Enfin, apprendre à traiter le lexique de sa langue maternelle est crucial pour le développement du vocabulaire en L2, notamment pour les stratégies de sa découverte. Les didacticiens définissent ces stratégies comme un ensemble d'opérations (Cyr 1998) ou procédures (Oxford 1990) mises en place par l'élève pour améliorer l'apprentissage. Elles supposent une attitude active et une implication dans ses démarches. Notre présentation des stratégies directes et indirectes, empruntée à Oxford (*ibid.*), n'est pas complète : elle évoque uniquement celles qui peuvent s'appuyer sur l'exploration d'une ressource de type système lexical. Ainsi, les stratégies directes, qui portent sur la langue elle-même, reposent sur trois types de procédures. Elles consistent notamment à établir des liens conceptuels entre ses divers éléments, en les regroupant, en les classant, en les associant (stratégies mnémoniques). La pratique de la langue est liée en grande partie à l'identification des patrons récurrents, elle suppose leur analyse et leur comparaison avec d'autres langues connues (stratégies cognitives). Enfin, la communication en langue étrangère doit beaucoup à des stratégies compensatoires, c'est-à-dire le devinement à partir d'éléments déjà connus. Par ailleurs, les stratégies indirectes, celles qui visent l'activité d'apprentissage elle-même, comprennent un volet métacognitif : en examinant la langue à maîtriser, l'étudiant cherche à découvrir son fonctionnement.

En amont, la mise en place de ces stratégies est favorisée et guidée par le traitement que l'enseignant propose de la matière même de l'apprentissage, en l'occurrence du lexique et de ses unités constitutives. Un modèle du lexique d'une langue et une ressource qui en est l'implémentation constituent un outil intéressant pour concevoir tout un programme

d'enseignement de lexique. Ses objectifs sont à définir à court comme à long terme : enrichir le stock lexical de l'apprenant, tout en stimulant des stratégies d'apprentissage efficaces.

4.3. Génération de ressources dérivées

En présentant le contenu informationnel d'un nœud lexical dans la section 3, nous avons évoqué les citations illustrant chaque sens. Elles sont sélectionnées majoritairement, mais non exclusivement, dans trois corpus (cf. *supra*). 31 000 exemples en ont été extraits – ainsi que de diverses publications, des écrits du Web, des films, des chansons –, ou tout simplement repérés dans des conversations quotidiennes. En poursuivant sa croissance, la ressource développe donc un corpus qui compte à ce jour environ 1 300 000 mots. Conformément au protocole de travail, lorsqu'une nouvelle lexie est intégrée au réseau, les exemples illustrant son sens sont recherchés d'abord dans le corpus de citations déjà disponibles, avant de se tourner vers des bases textuelles externes.

Si le corpus de la base RL-fr est encore relativement modeste du point de vue quantitatif, il présente un intérêt incontestable en ce qui concerne la désambiguïsation de ses occurrences. Progressivement, elles se trouvent associées aux lexies correspondantes, et de ce fait, sémantiquement annotées. L'exemple qui s'affiche à droite dans la figure 9 est l'un de ceux qui illustrent MOBYLETTE (cf. figure 8), mais il est également associé à neuf autres nœuds lexicaux : AVOIR_v, BEAU-FRERE, CHACUN, COUSIN¹, RATE, ROULER¹VI.3, SŒUR I, VOITURE I.I. A terme, toutes ces occurrences, y compris celles à valeur grammaticale, pourront renvoyer à l'acception dont ils relèvent.

The screenshot shows a software interface with a list of examples on the left and a detailed view of a citation on the right. The citation is: "Les sœurs, les cousins, les beaux-frères. Ils ont chacun une voiture. Il roule en mobylette. C'est un raté." The interface also shows a list of related lexical nodes: avoir, beau-frère, chacun, cousin, mobylette, raté, rouler, sœur, and voiture.

Figure 9. Exemple de citation désambiguïsée de la base RL-fr.

5. Conclusion

En présentant la base RL-fr, nous avons tenu à décrire le modèle conceptuel du lexique dont elle est la réalisation matérielle. A travers les données déjà disponibles, et celles qui continuent à l'alimenter quotidiennement, il a été possible de saisir la spécificité du travail lexicographique en réseau : il consiste à tisser des liens pour développer une structure complexe d'un graphe lexical. Le RL-fr offre ainsi de multiples exploitations à ses utilisateurs

humains, linguistes ou non-spécialistes, tout en fournissant des données nécessaires pour une génération automatique de ressources dérivées.

Dorota Sikora
 Université du Littoral-Côte d'Opale / Département Lettres et Arts / Centre universitaire de la Citadelle
 220, av. de l'Université, BP 5526, 59379 Dunkerque Cedex 1
 <Dorota.Sikora@univ-littoral.fr>

Références

- Cyr, P. (1998). *Les stratégies d'apprentissage*. Paris, Clé International.
- Delaite C. & Polguère A. (2013). Sex-Based Nominal Pairs in the French Lexical Network: It's Not What You Think. In Apresjan, V., Iomdin, B. & Ageeva, E. (eds), *Proceedings of the 6th International Conference on Meaning-Text Theory (MTT'13)*, Prague, 29-40.
- Gaume, B. & Duvignau, K. (2004). Pour une ergonomie cognitive des dictionnaires électroniques. *Document numérique*, 8 (3), 157-181.
- Mel'čuk, I. (2013). Tout ce que nous voulions savoir sur les phrasèmes, mais... *Cahiers de Lexicologie*, 102, 1, 129-149.
- Mel'čuk, I., Clas, A. & Polguère, A. (1995). *Introduction à la Lexicologie Explicative et Combinatoire*. Louvain-la-Neuve, Duculot.
- Mel'čuk, I. & al. (1984, 1988, 1992, 1999). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques* (vol. I-IV). Montréal, Presses de l'Univ. de Montréal.
- Mel'čuk, I. & Polguère, A. (2007). *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. Bruxelles, De Boeck & Larcier.
- Mel'čuk, I. & Polguère, A. (2008). Prédicats et quasi-prédicats dans une perspective lexicographique. *Revue de linguistique et de didactique des langues (Lidil)*, 37, 99-114.
- Mel'čuk, I. & Polguère, A. (2016). La définition lexicographique. *Cahiers de Lexicologie*, 109, 61-92.
- Oxford, R. (1990). *Language Learning Strategies: What Every Teacher Should Know*. New York, Newbury House.
- Pausé, M.S. (*à paraître*). Modélisation de la structure lexico-syntaxique des locutions au sein d'un réseau lexical. In *Actes du colloque 'Approches théoriques et empiriques de la phraséologie'*, Université de Lorraine (Nancy), 11-12 décembre 2014.
- Polguère, A. (2007). Lexical function standardness. In L. Wanner (ed.), *Selected Lexical and Grammatical Issues in the Meaning-Text Theory. In Honour of Igor Mel'čuk*, Amsterdam, Benjamins, 43-95.
- Polguère, A. (2011). Classification sémantique des lexies fondée sur le paraphrasage. *Cahiers de Lexicologie*, 98-1, 197-211.
- Polguère, A. (2014). From Writing Dictionaries to Weaving Lexical Networks. *International Journal of Lexicography*, 27-4, 396-418.
- Polguère A. (2016). La question de la géométrie du lexique. SHS Web of Conferences, 27 (2016) 01002, DOI:<<https://doi.org/10.1051/shsconf/20162701002>>, <<http://www.shs-conferences.org/articles/shsconf/abs/2016/05/contents/contents.html>> (21 mars 2017).
- Sikora, D. (2014). Brève histoire d'une erreur lexicale. Polysémie et liens lexicaux dans l'enseignement du vocabulaire. SHS Web of Conferences, 8, 1157-1172. DOI: <<https://doi.org/10.1051/shsconf/20140801229>>, <<http://www.shs-conferences.org/articles/shsconf/abs/2014/05/contents/contents.html>>, (21 mars 2017).
- Sikora, D. (2016). Identification d'unités phraséologiques et équivalence sémantique dans la traduction. In Corpas Pastor, G. (ed.) *Computerised and Corpus-based Approaches to Phraseology: Monolingual and Multilingual Perspectives*. Geneva, Tradulex, 218-231. <<http://www.tradulex.com/varia/Europhras2015.pdf>> (20 mars 2017).
- Vaguer, C. (2010). Être aux anges, sortir de ses gonds... Comment les langues traduisent-elles des états émotionnels ? *Cahiers Sens public*, 13-14, 253-269.



CLASSIQUES
GARNIER

POLGUÈRE (Alain) et SIKORA (Dorota), « Introduction », *Cahiers de lexicologie*, n° 109, 2016 - 2, *La définition*, p. 9-11

DOI : [10.15122/isbn.978-2-406-06861-7.p.0009](https://doi.org/10.15122/isbn.978-2-406-06861-7.p.0009)

La diffusion ou la divulgation de ce document et de son contenu via Internet ou tout autre moyen de communication ne sont pas autorisées hormis dans un cadre privé.

© 2017. Classiques Garnier, Paris.
Reproduction et traduction, même partielles, interdites.
Tous droits réservés pour tous les pays.

INTRODUCTION

Definition has been practised and discussed for nearly two and half millenniums. It has been more widely adopted, and less often reviled, than any other part of the original theory of logic drawn up by Aristotle. [...] The word 'definition' is more often used by the general public than any other peculiarly logical term except the word 'logic' itself. [...] If a man's car is insured against 'collisions', and is then damaged by running into a body of floodwater pouring across the road, must the insurer pay? Can one 'collide' with water? What is the definition of 'collision'?

Cette ouverture de l'étude de Robinson (1954 : 1)¹, référence incontournable pour qui s'intéresse à la définition, pourrait d'office faire apparaître ce numéro des *Cahiers de lexicologie* sur la définition comme présomptueux. D'une part, est-il possible aujourd'hui d'approcher la définition autrement qu'à travers son histoire de plus de deux millénaires? Certes, les études réunies dans Chaurand et Mazière (1990)² montrent que depuis *La Métaphysique* aristotélicienne, la définition a bien pris la couleur des siècles et que des analyses linguistiques permettent de saisir les évolutions de sa forme. Mais peut-on considérer que le quart de siècle qui s'est écoulé depuis la parution du volume des Actes du colloque « La définition » a apporté des changements notables?

D'autre part, la définition fait effectivement partie de notre quotidien : le numéro 11 de la revue électronique *Publif@rum* intitulé *Autour de la définition* (2010)³ présente la diversité des approches (diachroniques, discursives, textuelles) possibles, tout en explorant leurs applications aux études littéraires, à la pédagogie, voire ses détournements.

1 Richard Robinson (1954) : *Definition*, Oxford, Clarendon Press.

2 Jacques Chaurand et Francine Mazière (dir.) (1990) : *La définition. Actes du colloque La définition*, Université Paris 13, 18-19 novembre 1988, Paris, Larousse.

3 Hélène Giaufret, Michele Prandi et Micaela Rossi (dir.) (2010) : *Autour de la définition*, *Publif@rum*, 11, [En ligne], http://www.publiforum.farum.it/show_issue.php?iss_id=6, (consulté le 6 mai 2016).

Dans ce volume, nous proposons un regard à la fois théorique et appliqué sur la définition lexicographique et sur les évolutions qu'elle a connues dans les dictionnaires et autres ressources lexicales développés pour plusieurs langues, selon des approches théoriques variées. Du côté des utilisateurs, les changements sont également notables, surtout pour ce qui est du mode de consultation et de l'accès à l'information. Plutôt que de feuilleter des volumes, l'utilisateur n'a besoin que d'un clic pour naviguer entre les entrées lexicales, notamment à travers les termes de la définition, devenus des liens actifs. Cet accès rapide à un ensemble d'unités lexicales interconnectées rend toute inconséquence définitionnelle immédiatement détectable, même pour un non-spécialiste.

Par ailleurs, l'implémentation récente de ressources lexicales multidimensionnelles et modulaires a nettement modifié le statut de la définition au sein de l'article lexicographique. Les ressources et bases de données électroniques n'accordent pas nécessairement à la définition le statut privilégié dont elle bénéficiait jadis dans les dictionnaires papier. Elle forme souvent un module descriptif parmi d'autres : il suffit de penser à une ressource telle que WordNet et à l'importance de sa structuration en *synsets*, face à laquelle la paraphrase définitoire fait presque figure de parente pauvre. La question qui se pose dès lors est celle de savoir comment inscrire la définition lexicographique dans une structure de données linguistiques complexe. Souvent, une définition doit se positionner non plus seulement sur le plan purement lexical, mais rester cohérente à l'intérieur d'un système taxinomique tel qu'une ontologie informatique. Quelles sont dès lors les nouvelles contraintes qui pèsent sur elle ?

Le présent numéro des *Cahiers de lexicologie* réunit des études qui, nous semble-t-il, permettent d'envisager les évolutions possibles des méthodes et formats de définition lexicale, aussi bien dans le champ originel de la lexicologie et lexicographie, que dans ses applications, notamment en enseignement/apprentissage du vocabulaire, terminographie et traitement automatique de la langue.

L'article de C. Goddard, qui ouvre ce volume, présente et discute l'état actuel du développement du Métalangage Sémantique Naturel (*Natural Semantic Metalanguage*).

Les outils du Métalangage Sémantique Naturel sont mis à contribution par C. Levisen dans son étude consacrée à la lexicographie postcoloniale et à l'expression des émotions en bichelamar, créole parlé en milieu urbain de Port-Vila, capitale de Vanuatu (île d'Éfaté dans le Sud du Pacifique).

I. Mel'čuk et A. Polguère présentent l'approche de la définition lexicale développée dans le cadre de la Lexicologie Explicative et Combinatoire au sein de la théorie Sens-Texte, en exposant les propositions les plus récentes quant à la formalisation des définitions.

C'est selon les principes théoriques de la Lexicologie Explicative et Combinatoire que l'article de J. Milićević présente des réflexions issues du projet « Learner-Friendly Lexicographic Definitions in a French Learner's Dictionary » réalisé à l'Université de Dalhousie (Canada). La double approche adoptée par J. Milićević – lexicologique et pédagogique – vise un format de définition lexicographique à la fois rigoureux et accessible.

L'article de D. Sikora situe la définition dans ses deux dimensions d'activité (action de définir) et de type d'énoncé qu'un étudiant de langue étrangère doit construire. Une séquence de classe comprenant une série d'activités définitionnelles de niveau A2 du *Cadre européen commun de référence pour les langues* vient compléter la partie théorique de l'article.

En s'appuyant sur l'expérience lexicographique liée au développement des ressources *Le dictionnaire fondamental de l'environnement* (DiCoEnviro)⁴ et *Le dictionnaire fondamental de l'informatique et de l'Internet* (DiCoInfo)⁵, l'article de M.-C. L'Homme et A. San Marín s'intéresse à la microstructure des définitions terminologiques.

S. Seppälä, A. Ruttenberg, Y. Schreiber et B. Smith discutent, dans leur article, les deux types de définitions – textuelles et logiques – construites dans les ontologies informatiques. Les auteurs montrent que les définitions textuelles et logiques, destinées respectivement aux utilisateurs humains et aux systèmes informatiques, doivent être complémentaires pour améliorer l'ontologie dont elles font partie.

Ce numéro des *Cahiers de lexicologie* sur la définition ne saurait se passer d'une réflexion métalexicographique sur les propositions récentes des dictionnaires traditionnels. P. Frassi se penche sur un ensemble de dictionnaires de langue française et anglaise, en comparant les paraphrases définitionnelles qui y sont proposées.

Il apparaît ainsi, au fil de ces textes, que malgré son âge respectable, la définition lexicale continue à évoluer pour répondre aux exigences qu'imposent les nouveaux défis lexicographiques.

Alain POLGUÈRE
Université de Lorraine
et ATILF CNRS (UMR 7118)

Dorota SIKORA
Université du Littoral – Côte d'Opale
et HLLI (EA 4030)

4 <http://olst.ling.umontreal.ca/cgi-bin/dicoenviro/search.cgi>.

5 <http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi>.

III. Sémantique lexicale et didactique de vocabulaire

Ce troisième axe qui organise mes recherches se caractérise par sa dimension applicative dans l'enseignement. Il s'agit avant tout d'applications et d'implications d'une certaine modélisation du lexique à la didactique du vocabulaire dans l'enseignement de langues étrangères. Les études [20] à [25] rendent compte des recherches que j'ai réalisées et qui se poursuivent dans ce domaine.

- [20] Modèle lexicographique de croissance du vocabulaire fondé sur un processus aléatoire, mais systématique. Dans : Masseron, C., Garcia-Debanc, C. et Ch. Ronveaux (éds) *Enseigner le lexique. Pratiques sociales, objets à enseigner et pratiques d'enseignement*. Namur : Presses universitaires de Namur, 2013, 35-61 (avec A. Polguère).
- [21] Brève histoire d'une erreur lexicale. Polysémie et liens lexicaux dans l'enseignement du vocabulaire. Dans Neveu F., Blumenthal, P. Hriba, L., Gerstenberg, A., Meinschaefer, J. et S. Prévost (éds.), *4^e Congrès Mondial de Linguistique Française – CMLF 2014*, Paris : Institut de Linguistique Française, 2014, 1157-1172.
- [22] Activité(s) de définition dans l'apprentissage d'une langue seconde. Dans : Polguère, A. et D. Sikora, (dir) «La définition», *Cahiers de Lexicologie*, 109, 2, 2016, 117-143.
- [23] Locutions en apprentissage de langue seconde. *6^e Congrès Mondial de Linguistique Française, SHS Web of Conferences*, 46, 2018, 05002.
- [24] Entre sens et compétences métalinguistiques : locutions en phraséodidactique du FLE. *Lettre de l'AIRDF*, 64, 2018, 60-66.
- [25] De quelques aspects de compétence phraséologique en L2 : locutions en test. *Linguistica Atlantica*, 37 (2), 2019
- [26] (avec C. Haman) L'espace dans le roman : bilan et perspectives. Dans : C. Haman, D. Sikora et G. Kleiber (éds), *L'espace dans le roman français*

contemporain, Presses universitaires de Rennes, collection *La Licorne*, à paraître.

- [27] Partition de l'espace dans «Les Âmes grises» de Philippe Claudel. Dans : C. Haman, D. Sikora et G. Kleiber (éds), *L'espace dans le roman français contemporain*, Presses universitaires de Rennes, collection *La Licorne*, à paraître.

L'article [20], co-signé avec A. Polguère, rapporte et discute l'élaboration de la nomenclature, soit la phase initiale de l'implémentation d'un réseau lexical. S'il figure dans cette partie didactique du volume plutôt que dans la précédente, c'est parce que, chargée de sélectionner les items qui forment le noyau central du lexique d'une langue, je me suis appuyée sur des travaux et sur des listes de vocabulaire à destination pédagogique. En effet, dès le départ, cette modélisation du lexique devait avoir des applications dans l'enseignement.

L'étude [21] propose une analyse de certaines erreurs de combinatoire attestées dans des productions d'apprenants de français L2. Il est aujourd'hui largement admis en didactique de langues que les erreurs sont une importante source de renseignements pour l'enseignant, à condition néanmoins qu'elles soient diagnostiquées de manière pertinente. Elles peuvent alors orienter la démarche à suivre. Cet article est destiné à montrer que la sémantique lexicale fournit des outils pour l'évaluation de nature des erreurs et pour leur exploitation en classe de langue. Mon travail sur les activités de définition [22] développe plus encore cette dimension applicative, en montrant le rôle du paraphrasage du sens dans la structuration des connaissances lexicales.

Les articles [23] à [25] présentent des recherches et des réflexions menées dans le cadre du programme ICaL. Les enjeux sont, certes, didactiques, mais ces études constituent le prolongement des analyses proposées dans les articles [15] et [16] de la deuxième partie du présent volume. La thèse sous-jacente est celle que développent les chapitres 6 et 7 du volume de synthèse de mes travaux : les objets des apprentissages lexicaux, tout particulièrement lorsqu'il s'agit des unités poly-lexicales, doivent être abordés, décrits et enseignés en tenant compte de leurs caractéristiques linguistiques. Ces études montrent également l'importance des compétences méta-lexicales et phraséologiques pour le développement du vocabulaire d'apprenants.

On pourrait dès lors se demander la raison d'intégrer les deux dernières études ([26] et [27]) dans cette partie didactique, alors qu'elles ne traitent pas d'enseignement de vocabulaire. Il s'agit de l'Introduction, co-écrite avec C. Haman, du volume qui fait suite au colloque international «L'espace dans le roman contemporain français : approches linguistiques et littéraires» (Dunkerque, 11-12 octobre 2018). Issues d'un projet commun avec des collègues littéraires, elles montrent une autre facette applicative de la sémantique lexicale telle que je la pratique. En effet, je me suis tournée vers *Les Âmes grises* de Ph. Claudel avec l'objectif de montrer que l'espace dans un roman n'est pas simplement une représentation, mais une organisation linguistique à analyser sur le plan tant lexical que syntaxique et discursif, susceptible d'accorder une place précise au lecteur.

Modèle lexicographique de croissance du vocabulaire fondé sur un processus aléatoire, mais systématique

Alain Polguère* et Dorota Sikora**

* Université de Lorraine, ATILF, UMR 7118, Nancy, F-54000, FRANCE

** Université du Littoral-Côte d'Opale, HLLI, EA 4030, Dunkerque, F-59140, FRANCE

1. Introduction

1.1 Perspective lexicographique sur l'acquisition du vocabulaire

La réflexion sur le vocabulaire que nous proposons est non conventionnelle, dans le contexte de la didactique des langues, car elle s'insère dans une perspective lexicographique de construction d'une ressource lexicale – le *Réseau Lexical du Français* (Lux-Pogodalla & Polguère 2011 ; Gader et coll. 2012). Nous nous focalisons ici sur un élément du travail lexicographique qui, nous le pensons, est particulièrement pertinent pour une réflexion sur le vocabulaire, son apprentissage et son enseignement : il s'agit de la spécification et du processus de croissance de la nomenclature du Réseau Lexical du Français.

Comme l'indique son nom, le Réseau Lexical du Français, désormais *RLF*, modélise le lexique non pas par un texte dictionnaire, mais sous forme d'un réseau extrêmement riche et complexe d'unités lexicales. Le travail lexicographique de construction du RLF vise à créer un réseau lexical répondant aux trois objectifs suivants :

1. il permet de contrôler de façon rationnelle la croissance du RLF dans le cours de sa construction ;
2. son processus de croissance reflète une logique possible de la croissance du vocabulaire d'un individu ;
3. il peut servir de support à une navigation véritablement linguistique dans le système lexical de la langue.

Avant d'entrer dans le vif du sujet, il est nécessaire d'introduire quelques précisions terminologiques portant sur les notions fondamentales dont nous ferons usage dans le reste de l'article (1.2). Nous résumerons ensuite les principales caractéristiques du RLF, en tant que ressource lexicale (1.3).

1.2 Précisions terminologiques

Le cadre linguistique théorique dans lequel nous nous situons est celui de la *Lexicologie Explicative et Combinatoire* (Mel'čuk et coll. 1995), qui est le module lexical de la *théorie Sens-Texte* (Mel'čuk 1997 ; Milićević 2006). Dans ce contexte, nous faisons usage d'une terminologie bien spécifiée, dont un petit sous-ensemble doit être brièvement introduit. Nous nous concentrons bien évidemment sur les termes dénotant les différents types d'entités lexicales manipulées dans le cadre d'une activité lexicographique¹.

1 On pourra se reporter à Polguère (2008) pour une présentation des notions fondamentales (concepts et termes

L'unité d'étude en lexicologie et de description en lexicographie est l'*unité lexicale*, désignée dorénavant de façon plus compacte par le terme de *lexie*. Il existe deux types formels majeurs de lexies :

1. les *lexèmes*, qui se manifestent dans les énoncés en tant que signes linguistiques du type mots-formes – JUTEUX I [*couper une pêche juteuse dans son verre de vin.*], JUTEUX II [*tremper dans une affaire juteuse*], etc. ;
2. les *locutions*, qui se manifestent dans les énoncés en tant que syntagmes non compositionnels – HUILE DE COUDE [*Ce travail va demander un peu plus d'huile de coude.*], VIDER SON SAC [*Elle a vidé son sac devant les journalistes.*], etc.

Comme on le voit ci-dessus, nous écrivons les noms de lexies – lexèmes ou locutions – en petites majuscules, accompagnés si nécessaire de numéros lexicographiques distinctifs. Les lexies qui, comme JUTEUX I et JUTEUX II ci-dessus, se manifestent par les mêmes signifiants linguistiques et possèdent une intersection de sens significative sont regroupées au sein de *vocables polysémiques* – ici, le vocable JUTEUX, qui est égal à l'ensemble de lexies { JUTEUX I, JUTEUX II }. On identifie dans chaque vocable polysémique une *lexie de base*, qui est le « noyau sémantique » autour duquel se développe la structure polysémique du vocable. Dans le cas du vocable JUTEUX, il s'agit de JUTEUX I, dont JUTEUX II est dérivé par métaphore. Bien entendu, la numérotation des acceptions doit refléter le caractère primordial de l'acception de base ainsi que les écarts sémantiques entre toutes les lexies du vocable².

Les dictionnaires et autres ressources lexicales ont normalement comme *entrées* dans leur *nomenclature* les vocables de la langue. Il est cependant essentiel de bien garder à l'esprit que ce sont les lexies elles-mêmes – entités tout à fait circonscrites sur le plan sémantique – qui sont les unités de connaissance lexicale et, donc, les unités d'étude en lexicologie et lexicographie.

La lexie (lexème ou locution) peut-être vue comme une généralisation du signe linguistique et elle se modélise selon le même patron de description que ce dernier ; il s'agit, selon la conception Sens-Texte du signe linguistique, de décrire les trois composants fondamentaux suivants :

- *signifié* ;
- *signifiant* ;
- *combinatoire restreinte*.

Le dernier composant, la combinatoire restreinte, correspond à toutes les propriétés de combinatoire du signe linguistique – donc de la lexie – qui ne peuvent être entièrement déduites ni de son signifié ni de son signifiant (partie du discours, genre pour les noms en français, régime syntaxique, collocations, etc.).

Nous nous arrêtons là dans notre bref survol, ayant couvert la terminologie dont nous aurons à faire usage dans la discussion qui suit. Une fois cette terminologie élémentaire introduite, il convient notamment de ne plus utiliser le terme *mot* qu'avec une extrême prudence, car il ne s'agit justement pas d'un élément de la terminologie linguistique. Il pourra éventuellement s'employer pour désigner des entités dont la nature linguistique exacte n'est pas clairement identifiée. Par exemple, quand on parle des « listes de mots » construites pour l'enseignement du vocabulaire par des chercheurs ou pédagogues, qui n'indiquent pas toujours avec précision de quel type d'entités lexicales il s'agit : vocables, lexies, mots-formes, etc.

associés) de lexicologie.

2 Dans l'approche de la Lexicologie Explicative et Combinatoire, la numérotation en chiffres romains indique des écarts de type métonymique ou métaphorique.

1.3 Caractérisation générale du RLF

Le projet lexicographique sur lequel nous appuyons notre réflexion se situe dans la lignée des travaux menés sur les dictionnaires dits explicatifs *et combinatoires*, notamment le *Dictionnaire explicatif et combinatoire du français contemporain* (Mel'čuk et coll. 1984, 1988, 1992, 1999). Le RLF se distingue cependant radicalement des travaux antérieurs de ce type par sa structure et la méthodologie lexicographique mise en œuvre pour le construire.

Sur le plan de la structure, nous avons mentionné en 1.1 que le RLF n'est pas un dictionnaire, mais un *réseau lexical* dont les nœuds sont les lexies de langue et dont les arcs correspondent aux liens paradigmatiques (c'est-à-dire, sémantiques) et syntagmatiques (c'est-à-dire, de combinatoire restreinte) connectant ces lexies. Une telle structuration présente de nombreux avantages par rapport aux modèles traditionnels textuels que sont les dictionnaires, notamment :

- Un graphe lexical est une structure qui, selon toute vraisemblance, se rapproche beaucoup plus de la structure du *lexique mental* (Aitchison 2003) que la « plate » page de texte. On doit donc pouvoir effectuer sur une telle structure des opérations de navigation, d'inférence analogique, etc., qui s'approchent de celles effectuées de façon quasi instantanée par le locuteur lorsqu'il accède à ses connaissances lexicales.
- Un tel modèle n'est pas lié à un format de présentation spécifique, forçant un choix et un agencement fixes de l'information lexicale. Par contre, s'il est bien conçu, il peut être exploité pour la génération automatique de pages de dictionnaire aux contenus et formats variés. L'approche lexicographique du RLF relève ainsi de ce que l'on appelle la *lexicographie des dictionnaires virtuels* (Atkins 1996 ; Polguère 2012a).

Sur le plan de la méthodologie lexicographique, le projet de construction du RLF se distingue des travaux antérieurs en Lexicographie Explicative et Combinatoire par le recours à un éditeur lexicographique dédié – appelé *Dicet* – qui permet de construire la ressource lexicale non en écrivant des articles lexicographiques, mais en tissant des liens lexicaux : il s'agit donc d'un éditeur de graphes. La présentation de l'éditeur Dicet et des implications qu'a son utilisation sur le travail lexicographique sort du cadre de la présente discussion. On pourra se reporter à Gader et coll. (2012) pour la description de Dicet dans son état actuel et à Polguère (2012b) pour des réflexions sur la nature particulière du nouveau « geste lexicographique » que Dicet permet de mettre en œuvre.

Un des premiers problèmes auxquels il a fallu faire face dans le cadre de la construction du RLF a été celui de l'identification de sa nomenclature cible, sachant qu'il était impératif de fonctionner, sur le long terme, par étapes successives : des lexies les plus centrales du français, en s'étendant graduellement vers la périphérie du lexique de cette langue. La réflexion menée sur la nomenclature projetée pour le RLF et les différentes stratégies qui ont été élaborées dans ce contexte alimentent la discussion qui va suivre. Nous entamons cette dernière par un examen de la notion de vocabulaire de base, considérée dans une perspective didactique.

2. Vocabulaires de base en didactique de langues

2.1 Des listes de mots aux vocabulaires de base

La construction de listes de vocabulaire à usage pédagogique est une pratique qui a pris beaucoup d'importance au siècle dernier, pour ce qui est de l'enseignement du français et de l'anglais³. Ainsi,

3 Il ne nous appartient pas de faire ici un historique de la question. Pour le français, Galisson (1971 : 8–9) remarque que l'idée d'établir des listes de vocabulaires pédagogiques est ancienne puisque au XVIII^e, notamment, l'Abbé de l'Épée a dressé une liste de 5 400 mots nécessaires selon lui pour l'éducation des enfants sourds-muets. On trouvera d'autres informations sur l'historique des listes de vocabulaires pédagogiques pour le français ou l'anglais dans

Odgen (1930) propose un anglais de base, c'est-à-dire une sélection de 850 vocables d'anglais simplifié à usage international. Dans les années cinquante, les auteurs du *Français fondamental* ont entrepris de sélectionner un vocabulaire qui, à la fois, « fournit des moyens d'expression suffisants » et constitue « une première étape en vue de l'acquisition du “ français complet ” » (Gougenheim et coll. 1959 : 11).

Si les méthodes mises en œuvre par aussi bien par Odgen (1930) que par Gougenheim et coll. (1959) ont suscité de nombreuses critiques⁴, l'utilité de listes de vocabulaire dans l'enseignement d'une langue première (L1) et dans celui d'une langue seconde (L2) semble difficilement contestable. L'exemple de la *Oxford Wordlist* (Lo Bianco et coll. 2008)⁵, par exemple, illustre le chemin parcouru. L'outil, mis en place au terme des recherches menées entre 2007 et 2009 auprès des élèves et des étudiants australiens, permet de générer des listes des vocables les plus fréquents selon des critères géographiques, sociologiques, ethniques ou selon le parcours scolaire.

Le développement des études sur corpus et des méthodes statistiques d'analyse de fréquences ont permis d'affiner la notion de vocabulaire de base. Les vocables dont les occurrences ont les taux de fréquence les plus élevés peuvent à juste titre être considérés comme les plus utiles aussi bien pour un locuteur natif que pour un apprenant de L2.

Une démarche purement statistique aurait cependant des effets pervers. En s'appuyant sur des recherches menées sur l'anglais, McCarthy (1999) observe par exemple que *Saturday* a une fréquence d'emploi nettement supérieure à celle de *Wednesday*. Or, il serait étrange de dire que *Saturday* est plus important pour la maîtrise de l'anglais que *Wednesday*, alors que les noms des jours de la semaine forment un paradigme qui n'est pertinent que si on le prend comme un tout. McCarthy (1999) parle ainsi de **vocabulaire de base** en tant que partie du lexique d'une langue qui se caractérise par le potentiel communicatif le plus élevé : il est déterminé sur des bases fréquentielles, certes, mais revu et corrigé sous l'angle d'une pertinence pédagogique.

2.2 Évaluation quantitative d'un vocabulaire de base

La stratégie d'élaboration d'une nomenclature minimale pour une ressource lexicale telle que le RLF repose sur la thèse selon laquelle un vocabulaire de base recouvre le noyau dur du lexique d'une langue. Les études apportant une évaluation quantitative d'un vocabulaire de base dans le contexte de l'apprentissage d'une L2, notamment du français, et de l'acquisition d'une L1 ont en commun l'emploi du terme *mot* : la taille des vocabulaires est mesurée en nombres de « mots ». Il convient de réinterpréter *mot* dans la terminologie de la Lexicologie Explicative et Combinatoire, qui est celle mise en œuvre dans le projet lexicographique du RLF.

Dans le contexte de l'acquisition du français langue maternelle, Picoche (1995 : 369) indique que « l'essentiel des échanges verbaux s'effectue à partir de 3 000 mots du vocabulaire fondamental ». Elle précise par ailleurs, dans Picoche (2007 : 95⁶), que « le vocabulaire dont un individu a besoin pour s'exprimer sans difficulté dans la vie courante n'est que de quelques milliers de mots et que 8 ou 10 000 mots sont déjà une belle richesse »⁷. En discutant les différents aspects de la compétence lexicale, Picoche (1995, 2007) insiste sur le caractère polysémique des « mots ». On ne peut connaître un « mot » qu'une fois qu'on maîtrise ses différents sens, la combinatoire de chacun

Déro (1998 : Chapitre 3) et Chetouani (1997).

4 Pour une discussion critique, notamment du *Français fondamental*, voir Galisson (1971) et, plus récemment, Cortier & Bouchard (2008).

5 <http://www.oxfordwordlist.com/pages/>

6 Il s'agit du numéro de page du fichier électronique de Picoche (2007) ; cette édition électronique d'un ouvrage paru originellement en 1993 n'est pas paginée.

7 Le premier chiffre avancé par Picoche est très proche de celui que proposaient Gougenheim et coll. (1964) : leur français fondamental de premier et de second degré comprend en effet environ 3 500 entrées.

d'entre eux, ses dérivés morphologiques, ses connotations, etc.. Cela indique que le « mot » chez Picoche correspond approximativement au vocable tel qu'il est défini par la Lexicologie Explicative et Combinatoire. Il en ressort donc qu'un vocabulaire fondamental du français, celui qui permet d'assurer les échanges quotidiens, compte environ 3 000 vocables.

Dans le domaine de l'apprentissage d'une L2, le « mot » recouvre généralement une *famille de mots*, notion introduite par Nation (1982) pour servir d'unité de mesure dans l'évaluation des connaissances lexicales de l'apprenant, la croissance de son vocabulaire et la progression de la maîtrise d'une L2 qui en résulte. Une famille de mots comprend toutes les formes fléchies, ainsi que les dérivés forgés à l'aide des affixes les plus fréquents. Soit le cas du verbe DIRIGER ; il conditionne la famille de mots suivante : le verbe REDIRIGER, le verbe pronominal SE DIRIGER, les adjectifs DIRIGEANT_{Adj}, DIRECTEUR_{Adj}, DIRIGÉ, les noms DIRECTION, DIRIGEANT_N, DIRECTEUR_N, DIRIGEANTE, DIRECTRICE. Par ailleurs, on s'accorde à considérer, dans le domaine de l'apprentissage d'une L2, que la connaissance d'un vocable comme DIRIGER englobe la maîtrise des propriétés de chacune de ses acceptions – sens et combinatoire restreinte –, même si ces différents composants ne sont pas nécessairement désignés en ces termes.

La famille de mots est une famille de vocables, pour utiliser notre terminologie, et forts de ce constat, nous avons étudié les données quantitatives relatives aux vocabulaires de base en L2. Il s'avère que les chiffres avancés pour l'anglais (Laufer 1996 ; Adolphs & Schmitt 2003) et pour le français (Milton 2009) sont quasiment identiques. En effet, selon ces études, le niveau seuil de connaissances lexicales se situe, pour un apprenant étranger, à environ 3 000 vocables qui assurent une « couverture textuelle » à hauteur de 95%. Autrement dit, un apprenant avec un bagage lexical de 3 000 vocables est en mesure de comprendre un texte non technique en L2 (Laufer 1996). La compréhension de textes oraux n'est pas moins exigeante. En s'appuyant sur les travaux de Adolphs & Schmitt (2003), Milton (2009) constate que la communication orale requiert également un vocabulaire de 3 000 de vocables.

3. Des « familles de mots » aux dérivés sémantiques proches

Dans cette section, nous proposons une alternative plus rigoureuse et intuitivement plus satisfaisante aux familles de mots : les associations de dérivés sémantiques proches. Ces associations jouent un rôle central dans notre stratégie d'élaboration d'une nomenclature de base pour la construction d'une ressource lexicale comme le RLF, ainsi que pour implanter une stratégie d'enseignement du lexique.

3.1 Retour sur la dérivation

Commençons par effectuer un bref retour sur la dérivation en langue. Après avoir apporté quelques précisions sur les notions de dérivation morphologique, synchronique et diachronique, nous serons en mesure d'explicitier la notion clé de dérivation sémantique, à partir de laquelle sera construite celle de dérivé sémantique proche.

3.1.1 Dérivation morphologique et modèles lexicaux

La dérivation en langue est habituellement comprise comme une *dérivation morphologique*. Cette dernière peut être caractérisée très schématiquement de la façon suivante : on considère qu'il existe un lien de dérivation morphologique entre deux lexies L_1 (source de la dérivation) et L_2 (cible) – lien noté $L_1 \xrightarrow{m} L_2$ – lorsque L_2 peut être décrite comme (i) sémantiquement et (ii) morphologiquement construite à partir de L_1 . Cette caractérisation est volontairement très générale, presque simpliste, afin d'être compatible avec les différentes approches théoriques de la dérivation morpho-

logique, notion qu'il ne nous appartient pas d'examiner en détail ici⁸.

On notera que la dérivation morphologique, en tant que mécanisme langagier, peut être conçue de deux façons.

1. Il peut s'agir d'une **dérivation synchronique**. C'est alors un mécanisme morphologique d'application systématique, que le locuteur peut utiliser afin de construire une lexie « à la demande », comme dans le cas de la dérivation française par préfixation en RE- 'de nouveau' : DÉCHIRER ^m → REDÉCHIRER, DESSINER ^m → REDESSINER, etc.
2. Il peut s'agir d'une **dérivation diachronique**. C'est alors un mécanisme qui permet de créer des néologismes, qui vont éventuellement se lexicaliser dans la diachronie. Ainsi, la dérivation morphologique par suffixation mise en jeu dans BIGLER ^m → BIGLEUX n'existe que dans la diachronie du français. Un locuteur qui dit *Ce type est bigleux* n'a pas appliqué un mécanisme de dérivation par suffixation en -EUX pour construire l'adjectif BIGLEUX. Il a simplement utilisé un adjectif déjà présent dans le lexique de la langue. Si le locuteur applique la règle de dérivation en question selon son bon vouloir, il crée un néologisme, comme dans le cas d'une dérivation LOUCHER ^m → *LOUCHEUX.

Il est important de souligner que, lorsqu'on étudie la dérivation en lexicologie française, on s'intéresse en général à la dérivation morphologique – telle que nous venons de la décrire –, et qu'il s'agit dans l'immense majorité des cas de dérivations diachroniques : en français, les règles de dérivation morphologique sont pour la plupart d'application non systématique et, lorsqu'appliquées, servent à créer des néologismes, c'est-à-dire des lexies qui vont être perçues comme des créations du locuteur⁹.

Il est bien évident qu'un modèle lexicographique doit rendre compte des dérivés diachroniques, lexicalisés dans la langue, mais qu'il ne lui appartient pas de prendre en charge la dérivation synchronique. Cette dernière est prise en charge par la composante morphologique de la grammaire de la langue. Ce n'est que lorsqu'une dérivation morphologique normalement synchronique se lexicalise, c'est-à-dire lorsque la lexie dérivée se trouve associée à des particularités qui ne peuvent être calculées par la règle grammaticale de dérivation, qu'il convient de la prendre en compte dans le modèle lexical. Les particularités en question sont principalement de deux ordres.

1. Individualisation sémantique : le sens de la lexie dérivée n'est pas directement calculable à partir (i) du sens de la source de la dérivation et (ii) du sens du signe morphologique utilisé pour opérer la dérivation. Par exemple, le sens de RELIRE [*Je cherche quelqu'un pour relire mon article.*] est différent de 'lire de nouveau'. Dans un tel cas, on doit considérer que la lexie véritablement dérivée de LIRE – *J'ai lu son article, et je l'ai même relu deux fois* – a donné lieu à une expansion polysémique, qui impose la modélisation d'un vocable RELIRE comportant au moins deux acceptions : RELIRE I 'lire qqch. de nouveau' (morphologiquement transparent) et RELIRE II 'lire un texte pour le corriger ou l'améliorer'.
2. Individualisation sur le plan de la combinatoire : la lexie dérivée possède une combinatoire propre, qui ne peut être complètement expliquée par la combinatoire de la source de la dérivation et par la règle qui l'a produite. Par exemple, la forme participiale *connu* dans *C'est quelqu'un de très connu* possède une combinatoire lexicale propre : elle contrôle notamment des collocations du type *notoirement connu* ou *connu comme le loup blanc*. Il s'agit donc

8 Soulignons, de plus, que nous raisonnons sur une dérivation morphologique qui s'appliquerait aux lexies, alors que, strictement parlant, il s'agit d'un mécanisme mettant en jeux des entités **morphologiques** – les mots-formes et leurs morphèmes constitutifs. Pour un examen de la dérivation morphologique en langue, voir notamment Mel'čuk (1994 : Chapitre IV).

9 Cela n'est pas du tout le cas d'une lexie comme REDÉCHIRER, que l'on ne qualifiera pas de néologisme : sans être présente dans les dictionnaires, elle est de façon virtuelle dans le lexique dans la langue, du fait du caractère systématique de la dérivation qui la construit.

bien d'une lexie adjectivale (et non d'un « emploi adjectival d'un participe »), qu'il convient de décrire dans un article lexicographique consacré à la lexie en question.

Ces faits sont bien connus des lexicographes, qui sont nécessairement confrontés, lors de la construction de la nomenclature de leurs modèles lexicaux, au problème du diagnostic du caractère lexicalisé ou non des dérivés morphologiques. Les décisions doivent être prises au cas par cas, en menant une analyse de chaque dérivé avant son insertion éventuelle dans la nomenclature.

3.1.2 Dérivation sémantique

Dans son ouvrage sur l'enseignement du vocabulaire (Picoche 2007), J. Picoche fait le commentaire suivant à propos de la dérivation :

« Le fait que les dérivés opèrent une sélection parmi les sens du mot de base a pour conséquence que certains sens de ce mot restent sans dérivé. C'est, à côté de beaucoup de circonstances fortuites, une des raisons pour lesquelles il existe de nombreux mots qui, sans être morphologiquement des dérivés, fonctionnent pourtant comme des dérivés : Si le nom d'action correspondant à *retomber* est *retombée*, le nom d'action correspondant à *tomber* est *chute* ; *Max gifle Jean* se transforme en *Max donne une gifle à Jean* mais *Max frappe Jean* en *Max donne un coup (de poing) à Jean* ; *manger* peut être transformé en *faire un repas* ; *dormir* en *faire un somme* ; *se tromper* en *faire une erreur*. Il existe donc des familles de mots non morphologiques, mais purement sémantiques. À côté de la série totalement homogène *fin, finir, final, finalement*, on peut aligner une série morphologiquement hétéroclite mais sémantiquement tout aussi homogène : *terme, achever, dernier, enfin*¹⁰. »

Ce que dit ici J. Picoche est tout à fait limpide et a des implications tellement évidentes sur la façon dont il convient d'envisager l'enseignement et l'apprentissage de la langue que l'on peut s'étonner que la lexicologie et la didactique du français continuent de se focaliser sur une dérivation strictement morphologique. Il est pourtant de bon sens que cette dernière doit être intégrée dans un cadre plus vaste. La Lexicologie Explicative et Combinatoire (Mel'čuk et coll. 1995) concorde entièrement avec la perspective de J. Picoche sur la dérivation – pour des raisons aussi bien théoriques que pédagogiques – et complète la dérivation morphologique par une notion plus large, à caractère véritablement lexicologique : la **dérivation sémantique**, que nous noterons $L_1^s \rightarrow L_2$ pour la distinguer de la dérivation morphologique (notée $L_1^m \rightarrow L_2$)¹¹.

On trouve dans Mel'čuk & Polguère (2007 : 18–19) une présentation élémentaire de la dérivation sémantique. Trois conditions y sont spécifiées pour qu'existe entre deux lexies L_1 et L_2 une relation de dérivation sémantique $L_1^s \rightarrow L_2$, conditions que nous reproduisons ci-dessous avec de légères adaptations :

1. L_2 est en relation sémantique forte avec L_1 ; dans le cas le plus standard, L_2 se définit en termes de L_1 .
2. La relation sémantique entre L_1 et L_2 est récurrente dans la langue.
3. La relation entre L_1 et L_2 peut s'exprimer de façon morphologique dans la langue.

On voit que la dérivation sémantique possède une couverture plus large que la dérivation morphologique puisqu'il n'est pas nécessaire qu'existe une relation morphologique entre la source L_1 et la cible L_2 . Cependant, les liens de dérivation sémantique couvrent un spectre sémantique relativement étroit : il s'agit du sous-ensemble des liens sémantiques qui sont récurrents en langue et tendent à se morphologiser. Ainsi, il ne suffit pas qu'une lexie L_2 se définisse en termes d'une autre lexie L_1 pour qu'elle soit considérée comme son dérivé sémantique. Il faut aussi que le lien sémantique soit suffi-

10 Chapitre 1, section 3 « Les familles ont une histoire », sous-section « Les simples relations ».

11 Soulignons que J. Picoche fait elle-même usage du terme *dérivé sémantique*.

samment « standard » pour que la langue ait développé des outils morphologiques permettant son expression. Illustrons cela avec la paire BOIRE [*Elle a bu sa vodka d'un trait.*] ~ VERRE [*Son verre était à moitié vide.*]. Il est justifié de dire que la lexie VERRE est un dérivé sémantique de la lexie BOIRE, car : (i) 'verre' s'analyse approximativement en 'petit récipient servant à boire' ; (ii) le sens 'objet servant à faire L₁' est très souvent lexicalisé ; (iii) cette lexicalisation peut très naturellement s'effectuer de façon morphologique – par exemple, (S')ABREUVER ^{s/m} → ABREUVOIR, HACHER ^{s/m} → HACHOIR, TONDRE ^{s/m} → TONDEUSE, etc.

L'écart sémantique entre deux lexies liées par une dérivation sémantique peut tout à fait être (presque) nul, leur différence se ramenant à un changement de partie du discours. Par exemple, la lexie ODEUR [*Je ne supporte pas l'odeur de ce savon.*] est un dérivé sémantique de la lexie SENTIR [*Ce savon sent mauvais.*] car : (i) 'odeur de X' s'analyse approximativement en 'ce que X sent' ; (ii) l'écart sémantique entre les deux lexies se ramène en fait à un simple changement de partie du discours, la nominalisation d'un sens verbal étant un phénomène très courant ; (iii) la nominalisation s'encode fréquemment de façon morphologique – par exemple, MARCHANDER ^{s/m} → MARCHANDAGE, POLLUER ^{s/m} → POLLUTION, etc. La synonymie exacte ou la conversivité – qui relèvent aussi de l'équivalence sémantique – peuvent tout autant être considérées comme des cas de dérivations sémantiques, car il s'agit de relations basiques entre lexies, pour lesquelles les langues offrent des outils morphologiques ; par exemple :

- dérivation synonymique par troncation : APPARTEMENT ^{s/m} → APPART ;
- dérivation converse par forme réfléchi : ÉTONNER [*Ma question étonne Éva.*] ^{s/m} → S'ÉTONNER [*Éva s'étonne de ma question.*].

Remarquons que, dans le cas du réfléchi, si la lexie S'ÉTONNER est morphologiquement dérivée de ÉTONNER (elle est morphologiquement plus complexe), elle ne se distingue sémantiquement de cette dernière que par une permutation de structure actancielle : *X étonne Y* et *Y s'étonne de X* dénotent exactement la même situation. Il est donc légitime de considérer aussi bien une dérivation sémantique ÉTONNER ^s → S'ÉTONNER qu'une dérivation sémantique inverse S'ÉTONNER ^s → ÉTONNER.

Nous allons maintenant expliquer comment la notion de dérivation sémantique peut être utilisée pour construire une alternative à la notion de famille de mots (examinée plus haut, en 2.2) sur des bases qui nous semblent plus rigoureuses, théoriquement et pédagogiquement.

3.2 Dérivés sémantiques proches

3.2.1 Gradation dans la proximité sémantique des dérivés

À partir d'une lexie donnée L, on peut identifier une association de lexies {L} _{dsém}, constituée de L et de tous ses dérivés sémantiques. Une association de ce type, qui repose sur la notion bien circonscrite de dérivation sémantique, nous semble plus rigoureusement spécifiée qu'une famille de mots, qui relève en fin de compte d'une notion assez floue et dont la pertinence interlinguistique reste problématique au vu des différences considérables que les langues manifestent sur le plan morphologique.

Une association construite strictement en fonction de la dérivation sémantique est cependant potentiellement trop vaste pour constituer ce que l'on pourrait voir comme étant une « molécule » de connaissance lexicale. Ainsi, BEURRE est un dérivé sémantique de LAIT, mais on ne considérera pas, dans un contexte d'apprentissage du vocabulaire, qu'il appartient à la « famille lexicale immédiate » de LAIT, au même titre que LAITIER par exemple ; et cela, pour des raisons qui ne sont pas uniquement morphologiques. On doit donc aller encore plus loin dans la caractérisation d'une notion se substituant à celle de famille de mots, en restreignant le spectre des liens sémantiques qui est cou-

vert par les associations de dérivations sémantiques. Telle est la finalité de l'élaboration de la notion de dérivation sémantique proche. Pour introduire celle-ci, commençons par souligner une caractéristique importante de la dérivation sémantique : le caractère graduel de la proximité sémantique dont elle rend compte.

La dérivation sémantique recouvre un spectre assez large de relations sémantiques : des relations correspondant à un écart sémantique nul à celles pour lesquelles l'écart est suffisamment important pour que la présence même d'un lien de dérivation sémantique puisse être débattue. Ainsi, à une extrémité de l'échelle de gradation en termes de proximité sémantique, se trouve la proximité « extrême » qu'est l'identité de sens – c'est-à-dire la synonymie exacte : VOITURE^s → BAGNOLE. À l'autre extrémité, se trouvent des proximités « faibles » représentées, par exemple, par une dérivation sémantique adjectivale non vide et non morphologique comme S'ENDORMIR^s → SOMNOLENT.

La théorie Sens-Texte (cf. section 1.2) est, à notre connaissance, la seule approche théorique qui propose un outil formel permettant de décrire exhaustivement les différents types de liens de dérivation sémantique présents dans les langues, ainsi qu'une intégration de la description de ces **liens lexicaux paradigmatiques** avec celle des **liens syntagmatiques** instanciés dans les syntagmes phraséologiques du type base + collocatif appelés **collocations** (Hausmann 1989 ; Mel'čuk 1995), comme :

- *pluie + forte, grosse...* ; angl. *rain + driving, heavy, strong...* ;
- *clou + planter, enfoncer [un clou]* ; angl. *nail + to drive [a nail]* ;
- etc.

L'outil descriptif « unificateur » pour les liens paradigmatiques et syntagmatiques proposé par la théorie Sens-Texte est le système des **fonctions lexicales standard**. Les fonctions lexicales Sens-Texte forment un système extrêmement riche, que nous ne pouvons présenter ici, nous contentant de renvoyer le lecteur à l'abondante littérature traitant de la question¹². Il nous faut cependant faire usage ici des fonctions lexicales, ainsi que des formalisations qui leur sont associées, puisque la notion de dérivé sémantique proche que nous allons maintenant introduire repose entièrement sur la spécification d'un sous ensemble des fonctions lexicales paradigmatiques standard. Nous prendrons donc soin d'explicitier la valeur sémantico-syntaxique de chaque fonction lexicale que nous serons amenés à introduire.

3.2.2 Dérivés sémantiques proches

Le système des fonctions lexicales standard de la théorie Sens-Texte nous a permis de caractériser de façon précise un ensemble de treize liens privilégiés de dérivation sémantique connectant des lexies manifestant une très grande proximité sémantique : ce que nous appellerons désormais la **dérivation sémantique proche**. Nous identifions une relation de dérivation sémantique proche entre deux lexies L_1 et L_2 , relation notée $L_1 \text{dsp} \rightarrow L_2$, si :

- soit L_2 est un synonyme exact de L_1 ;
- soit L_2 est une « transformée » sémantique minimale de L_1 .

Bien entendu, dire qu'une lexie est une transformée sémantique minimale d'une autre ne relève aucunement d'un critère objectif. Une part importante de subjectivité entre dans la sélection d'un ensemble donné de liens lexicaux qui, seuls, seront considérés comme correspondant à de la dérivation sémantique proche. Une fois cette sélection établie, cependant, l'inclusion ou l'exclusion d'un lien donné lexical $L_1 \rightarrow L_2$ se fera sur une base tout à fait objective : il ne s'agira de dérivation sé-

¹² Voir Mel'čuk (1996), pour une présentation générale, et Polguère (2007), pour une étude du caractère standard vs non standard des fonctions lexicales.

mantique proche que si le lien en question correspond à une des treize relations préalablement identifiées.

La liste que nous proposons a été tout d'abord le produit d'une réflexion menée à partir de l'étude du système des fonctions lexicales paradigmatiques standard et de son utilisation dans des projets lexicographiques tels que le *Dictionnaire explicatif et combinatoire du français contemporain* (Mel'čuk et coll. 1984, 1988, 1992, 1999), la base lexicale *DiCo* (Mel'čuk & Polguère 2006), etc. La pratique lexicographique montre notamment que certains liens standard ont une fréquence d'apparition dans les articles lexicographiques beaucoup plus élevée que d'autres et que, dans tous les cas, il s'agit des liens manifestant un écart sémantique très faible entre la lexie source et la lexie cible. Sans pouvoir entrer ici dans des considérations très théoriques, mentionnons que ce phénomène est tout à fait en accord avec la nature fonctionnelle profonde des langues, qui sont organisées comme des générateurs de paraphrases (Milićević 2007).

Notre liste initiale de dérivations sémantiques proches a été légèrement révisée de façon empirique, suite au travail effectué sur la nomenclature dite *d'amorçage* du RLF, qui est présentée plus bas dans la section 4. Nous donnons maintenant la liste à laquelle nous sommes finalement parvenus. Dans l'énumération qui suit, chaque dérivation sémantique proche est introduite par la fonction lexicale paradigmatique standard correspondante, suivi d'une brève description de sa valeur sémantico-syntaxique, ainsi que d'une ou plusieurs illustrations et, éventuellement, d'un commentaire. Le symbole « L » désigne ici l'*argument* de la fonction lexicale, c'est-à-dire la source de la dérivation sémantique proche.

1. **Syn** : synonyme exact de L ; par ex. VOITURE → AUTOMOBILE_N.
2. **Anti** : antonyme exact de L ; par ex., LÉGAL → ILLÉGAL.
3. **Syn_{sexe}** : quasi-synonyme (plus précisément, synonyme à intersection) de L qui dénote le même type d'individu que L, mais de sexe opposé ; par ex., ACTEUR → ACTRICE, COUSIN → COUSINE. Ce lien de fonction lexicale est parfaitement symétrique ; cependant, la lexie du genre masculin est souvent en même temps un générique, en français. D'un point de vue formel, **Syn_{sexe}** est équivalent, dans le langage formel des fonctions lexicales, à une fonction dite *semi-standard* encodée ainsi : **de sexe opposé Syn** .
4. **Fem** : lexie dénotant un être vivant femelle correspondant à l'être vivant mâle dénoté par L ; par ex., COQ → POULE. La fonction lexicale **Fem** est un « néologisme » dans le système des fonctions lexicales, introduit dans le cadre du projet de construction du RLF. Strictement parlant, elle ne fait pas partie du système des fonctions lexicales standard de la théorie Sens-Texte. Elle est fonctionnellement peu éloignée de la fonction lexicale précédente, mais s'en distingue tout de même par le fait qu'elle ne relève pas de la synonymie véritable. Elle serait dans les faits beaucoup plus proche d'une relation de contrastivité (fonction lexicale standard **Contr**).
5. **Masc** : lexie dénotant un être animé mâle correspondant à l'être animé femelle dénoté par L ; par ex., POULE → COQ. Il s'agit donc de la relation exactement symétrique de la précédente.
6. **V₀** : verbalisation de la lexie non verbale L ; par ex., BAGARRE → SE BAGARRER.
7. **S₀** : substantivisation vide de la lexie non nominale L ; par ex., PARTIR → DÉPART.
8. **Adj₀** : adjectivisation vide de la lexie non adjectivale L ; par ex., ÉCOLE → SCOLAIRE.
9. **Adv₀** : adverbialisation vide de la lexie non adverbiale (normalement, adjectivale) L ; par ex., LENT → LENTEMENT.
10. **Si** : nom dénotant le i^e (1^{er}, 2^e, 3^e...) actant de L ; par ex., CONDUIRE ('X conduit

Y') → CONDUCTEUR [= **S**₁].

11. **A_i** : adjectif dénotant le fait d'être le i^e (1^{er}, 2^e, 3^e...) actant de L ; par ex., FATIGUE ('fatigue de X causée par Y') → FATIGUÉ [= **A**₁]. On admettra aussi comme dérivation sémantique proche les cas où l'adjectif dérivé ajoute un sens de perfectivité par rapport au sens de la source de la dérivation, comme dans le cas de DÉCOURAGER ('X décourage Y') → DÉCOURAGÉ (qui signifie '[Y] qui a été découragé par X'). Strictement parlant, il s'agit ici de la fonction lexicale standard complexe **A₂Perf**.
12. **Able_i** : adjectif dénotant le fait d'être susceptible d'être le i^e (1^{er}, 2^e, 3^e...) actant de L ; par ex., SE FIER ('X se fie à Y') → FIABLE [= **Able**₂].
13. **Mult** exact : nom collectif qui dénote spécifiquement un ensemble de L ; par ex., FEUILLE → FEUILLAGE. On notera que le collectif doit être « exact » ; il doit donc contenir le sens de L dans sa définition et signifier 'ensemble de L'. Ainsi, BANC [*de poissons, d'huitres...*] n'est pas un dérivé sémantique proche de POISSON, car c'est une lexie trop vague.

La remarque faite à propos de **Mult** doit être généralisée à toutes les autres fonctions lexicales mentionnées ci-dessus : elles sont prises dans leur « sens étroit », tel que spécifié dans les descriptions qui les accompagnent. Par exemple, strictement parlant, VICTIME [*d'un meurtre*] est un **S**₂ valide de MEURTRE [*par l'individu X de l'individu Y*], mais VICTIME ne devrait pas être considéré comme étant un dérivé sémantique proche de MEURTRE, car son sens est beaucoup plus vague que 'individu Y tué par meurtre' (*[un] meurtré).

Grâce à la liste donnée ci-dessus, nous sommes en mesure de substituer aux « familles de mots » les associations lexicales induites par la dérivation sémantique proche. Ces dernières sont des petits ensembles de lexies sémantiquement très proches, en relation de (quasi-) paraphrase, souvent, mais non nécessairement, liées morphologiquement. Nous allons maintenant voir comment les associations de dérivations sémantiques proches nous permettent de gérer la croissance de la nomenclature d'une ressource lexicale telle que le RLF.

4. Nomenclature du RLF

4.1 Construction d'une nomenclature en trois volets

Rappelons que le RLF a pour vocation, non seulement de décrire les lexies du français, mais aussi de rendre compte de la structuration même du lexique en tant que réseau de lexies. Il s'agit de modéliser le positionnement des lexies au sein d'un réseau de relations paradigmatisées et syntagmatisées. Une double piste de réflexion, à la fois didactique et lexicologique, nous a amenés vers l'hypothèse que le noyau dur du système lexical du français que doit cibler le RLF correspond à son vocabulaire de base tel qu'on le conçoit dans un contexte d'enseignement/apprentissage du lexique. La description, dans une ressource lexicale, des acceptions des vocables qui constituent un vocabulaire de base permet en effet de procéder dans un ordre qui, contrairement au cheminement alphabétique traditionnellement appliqué pour la construction des dictionnaires, est déterminé par la structure intrinsèque de la langue : on entre dans le lexique par un noyau lexical minimal servant à amorcer le processus de croissance d'une nomenclature ; on s'attache ensuite aux dérivés sémantiques proches issus de ce noyau lexical ; puis on « s'aventure » dans l'ensemble du lexique selon la logique des liens sémantiques et de combinatoire issus du vocabulaire acquis. Ce *modus operandi* présente l'avantage d'assurer une croissance de la nomenclature qui se fait d'une manière à la fois aléatoire et contrôlée par le système de liens lexicaux privilégiés que nous venons de discuter (section 3).

Il résulte de tout cela que le processus de croissance de la nomenclature du RLF s'articule en trois volets :

1. identification d'une **nomenclature d'amorçage (NA)**, qui réunit les vocables constituant le noyau du vocabulaire de base du français et qui est injectée dans la base lexicale du RLF dès sa création pour « amorcer » sa nomenclature et son processus de croissance¹³ ;
2. génération systématique d'une **nomenclature directement induite (NDI)** de la NA, qui ajoute aux vocables de la NA les dérivés sémantiques proches de leur lexie de base¹⁴ ;
3. génération aléatoire, mais logiquement maîtrisée, d'une **nomenclature indirectement induite (NII)**, qui comprend, entre autres, les vocables dont les acceptions sont des cibles de liens de fonctions lexicales issus des entrées de la NA et de la NDI.

Nous allons tout d'abord traiter du problème de l'identification de la NA (4.2), pour ensuite passer à celui de l'implantation et du contrôle de sa croissance (4.3).

4.2 Identification de la nomenclature d'amorçage (NA)

De nombreuses ressources permettent aujourd'hui d'accéder à un vocabulaire de base du français. En construisant la NA, nous avons systématiquement pris en compte les données des trois listes suivantes :

1. la nomenclature du *Robert Benjamin* (Collectif Robert 2009), un dictionnaire pédagogique destiné au cycle des apprentissages fondamentaux de l'école primaire et comportant 6 500 entrées ;
2. la liste Eduscol, disponible sur le site du Ministère de l'Éducation nationale français¹⁵ ; cette ressource repose sur des études de fréquence menées sur le *Trésor de la Langue française* par Brunet (1981) et comprend 1 463 entrées, qui doivent être travaillées tout au long du cycle des approfondissements (8 à 10 ans) ;
3. un vocabulaire du Ministère de l'Éducation, du Loisir et du Sport (MELS) du Gouvernement du Québec, élaboré à l'Université de Montréal par une équipe de didacticiens (Lefrançois et coll. 2011) à partir des études de fréquence menées sur un corpus de manuels scolaires et de littérature jeunesse ; ce vocabulaire, à l'usage des enseignants intervenant dans les premières classes du cycle secondaire, comporte 4 548 vocables¹⁶.

Une liste de 6 784 vocables a ainsi été arrêtée, avec cependant la nécessité de la soumettre à un examen critique conformément à la démarche proposée par McCarthy (1999). Il s'agissait en effet, d'une part, d'éliminer des éléments qui ne relèvent pas de vocabulaire de base et de l'autre, de rationaliser une liste de vocabulaire issue des ressources s'inscrivant dans des cadres méthodologiques bien différents. Nous avons ainsi été amenés à éliminer des vocables « parasites » tels que ANORAK – issu de la nomenclature du *Robert Benjamin* –, qui est certes fréquent dans le discours des écoliers et de leurs parents, mais dont le potentiel communicatif est peu élevé. Par ailleurs, nous avons dû introduire dans la future NA des vocables tels que SE COMPORTER, DÉCOURAGER ou MANIFESTER, absents des listes consultées, mais dont l'utilité nous apparaissait incontestable.

Pour transformer une liste de vocabulaire en nomenclature d'amorçage, il a été nécessaire d'émettre

13 Lors de l'injection de la NA dans la base RLF, un vocable a été créé automatiquement pour chaque élément de la liste, avec une acception par vocable créé. Immédiatement après l'injection, les articles lexicographiques des acceptions sont vides, si ce n'est pour l'indication d'une partie du discours.

14 Rappel : la notion de lexie de base d'un vocable a été introduite à la section 1.2.

15 Cf. <http://eduscol.education.fr/pid23250-cid50486/vocabulaire.html> (mise à jour le 9 février 2010, consultée le 25 mai 2011).

16 Nous remercions Pascale Lefrançois (Université de Montréal) et Ophélie Tremblay (Université du Québec à Montréal) de nous avoir donné accès à leurs données de recherche. Nous remercions également Caroline Bégin (MELS) et Hélène Cajolet-Laganière (Université de Sherbrooke) d'avoir autorisé la diffusion de l'information contenue dans le rapport de recherche Lefrançois et coll. (2011).

continuellement des hypothèses sur la structuration des sous-réseaux lexicaux. Ainsi, face aux quatre vocables CRÉATION, CRÉER, CRÉATEUR et CRÉATRICE, issus de la compilation des listes de référence, nous avons considéré que la lexie de base du verbe CRÉER permettait d'accéder par des liens de dérivation sémantique proche aux lexies :

- CRÉATION en tant que dérivé nominal de type S_0 ;
- CRÉATEUR et CRÉATRICE en tant que dérivés du type S_1 (noms du premier actant).

Nous avons donc sélectionné le verbe CRÉER comme pivot de l'association lexicale de dérivés sémantiques proches à insérer dans la NA. Cette démarche nous a conduits à une NA comptant 3 734 vocables.

4.3 Modèle d'expansion de la nomenclature d'amorçage

Les fonctions lexicales décrites dans la section 3 permettent d'étendre la NA du RLF en ajoutant à la ressource lexicale les vocables dont une acception est la cible d'un lien de dérivation sémantique proche issu de la lexie de base d'un vocable de la NA. Cette dernière est ainsi complétée par une nomenclature directement induite (NDI). Ce mode de croissance a permis d'atteindre une nomenclature NA+NDI d'environ 10 000 vocables.

Notons que la NDI inclut, en plus des vocables introduits du fait de liens de dérivation sémantique proche, tous les vocables possédant une des caractéristiques suivantes.

- Ce sont des verbes pronominaux correspondant aux verbes non pronominaux de la NA (et vice versa).
- Ce sont des homonymes de vocables de la NA.
- Ce sont des locutions formées à partir des lexèmes de la NA¹⁷. Ainsi, la NA comprend les vocables RETOURNER, TAILLER et VESTE ; ceux-ci permettent d'induire les locutions très courantes du français RETOURNER SA VESTE et TAILLER UNE VESTE, qui doivent être intégrées dans la NDI du RLF.

La nomenclature NA+NDI constitue une bonne hypothèse pour le noyau du lexique français. Le processus d'expansion de la nomenclature du RLF ne s'arrête cependant pas là. La croissance systématique déterminée par la dérivation sémantique proche se poursuit par l'ajout d'une nomenclature dite *indirectement induite* (NII). Celle-ci comprend des vocables dont les acceptions sont des cibles (i) de liens paradigmatiques plus riches que la dérivation sémantique proche et (ii) de liens syntagmatiques. Ainsi, le nom CÉLIBATAIRE, présent dans la NA, entraînera l'insertion dans la NII de l'adjectif ENDURCI, qui est un des collocatifs intensificateurs de CÉLIBATAIRE (fonction lexicale syntagmatique **Magn**). Même si cet adjectif n'est bien évidemment pas au cœur du lexique français, son insertion en phase 3 de croissance de la nomenclature est justifiée par le fait qu'il est un collocatif privilégié d'une lexie qui, elle, a été identifiée comme appartenant au lexique de base. La croissance de la nomenclature du RLF se poursuit ainsi graduellement, au fur et à mesure que les lexicographes sont amenés à tisser de nouveaux liens lexicaux issus des vocables déjà présents dans la nomenclature¹⁸.

La Figure 1 ci-dessous visualise une partie du processus d'expansion de la NA à partir de la lexie de base du vocable DÉTRUIRE.

17 La NA du RLF, comme la plupart des vocabulaires de base, est exclusivement constituée de lexèmes (cf. section 1.2). Les locutions étant des lexies à part entière, elles doivent apparaître dans la nomenclature du RLF et il faut donc prévoir un processus permettant leur insertion dans la base lexicale. C'est ce processus que nous caractérisons brièvement ici.

18 Pour une présentation de la méthodologie de description des vocables du RLF, voir Polguère (2011).

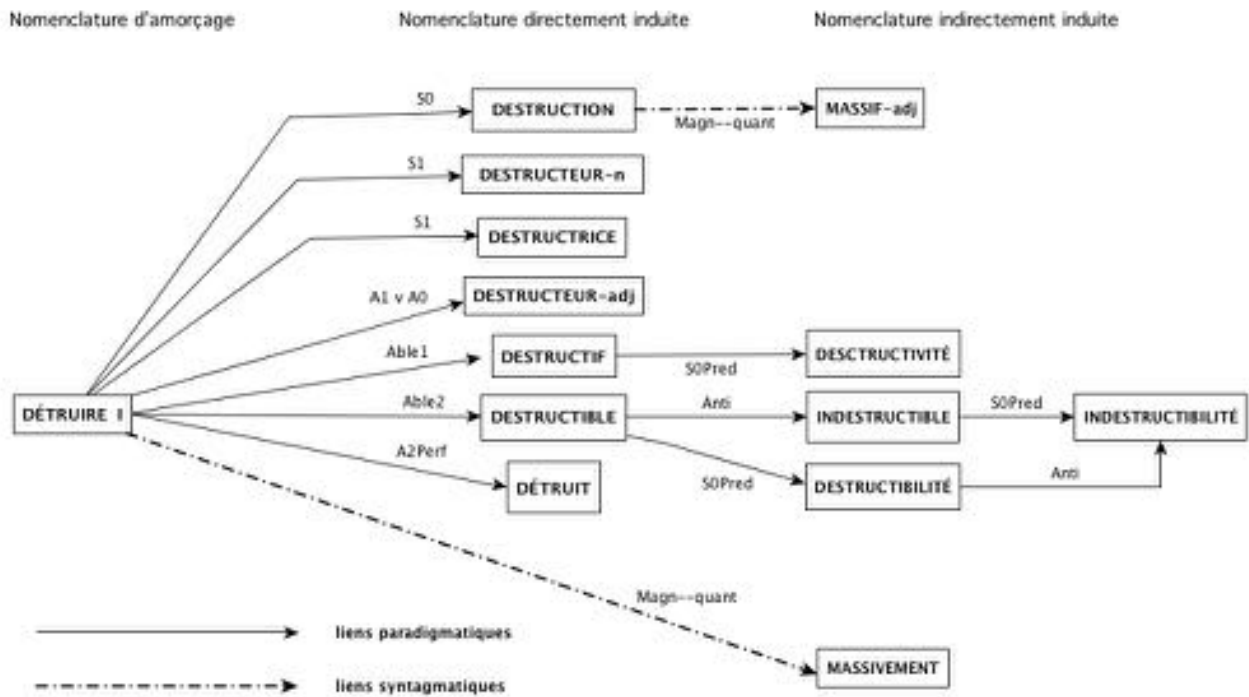


Figure 1 : Schéma de croissance de la nomenclature du RLF

Une fois la NA injectée dans le RLF, le tissage des liens de dérivation sémantique proche a été effectué de façon systématique sur toute cette nomenclature initiale ; puis s'est engagé le processus d'expansion quasi organique de la base, conditionné par le travail lexicographique proprement dit de description des vocables. Le graphe ci-dessous, Figure 2, résume l'évolution de la couverture lexicale du RLF, du début (ou presque) du tissage des liens de fonctions lexicales jusqu'au moment de la rédaction du présent texte. La croissance lente, mais régulière, du taux de polysémie – nombre de vocables divisé par le nombre de lexies (acceptions) – correspond au dégagement progressif de la structure polysémique des vocables de la base. La décroissance ponctuelle que l'on peut constater sur plusieurs mois a coïncidé en grande partie avec la phase de tissage des liens de dérivation sémantique proche à partir de la NA, qui a mené à l'introduction de la NDI et donc a généré un apport soudain de nouveaux vocables dans la base.

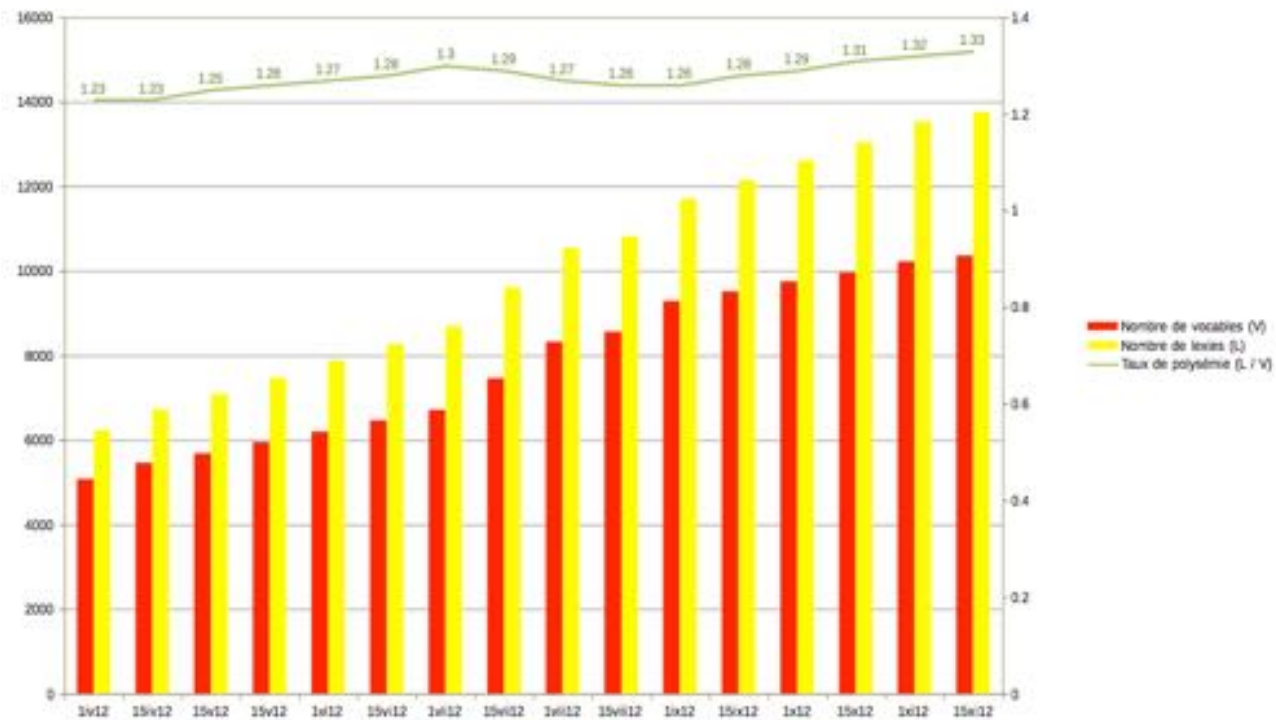


Figure 2 : Statistiques de l'expansion de la couverture lexicale du RLF

5. Conclusions

Le modèle d'identification et d'expansion de la nomenclature du RLF que nous venons d'explicitier nous semble particulièrement intéressant d'un point de vue pédagogique pour au moins deux raisons.

Tout d'abord, il s'inspire directement des réflexions menées par les spécialistes de l'enseignement du vocabulaire. Nous avons cherché à rationaliser et consolider la notion trop floue de « famille de mots », en lui substituant celle d'association lexicale fondée sur la dérivation sémantique proche. Nous avons de plus pris en compte plusieurs listes de vocabaires établies dans une perspective pédagogique pour synthétiser la NA du RLF. La couverture lexicale initiale du RLF résulte donc, en premier lieu, d'une réflexion pédagogique sur les vocabulaires.

Ensuite, le processus de croissance de la couverture lexicale du RLF pourrait, selon nous, inspirer les didacticiens intéressés à construire une méthode d'exploration et d'acquisition du lexique. En effet, ce processus présente certaines analogies avec la façon dont peut s'effectuer le développement « naturel » du vocabulaire d'un apprenant. La nomenclature du RLF croît de façon à la fois aléatoire – il n'y a pas de programmation stricte de l'ordre d'insertion de nouveaux vocabaires dans la base – et systématique – la croissance se fait en suivant, à l'intérieur de champs lexicaux bien identifiés (Polguère à paraître), les liens sémantiques et de combinatoire qui tissent le réseau lexical de la langue. Or, il ne fait aucun doute que la croissance du vocabulaire d'un individu est à la fois en grande partie aléatoire et en même temps conditionné par les liens forts qu'entretiennent les lexies de la langue.

Notons pour conclure que le RLF sera entre les mains de ses utilisateurs un outil pédagogique aux applications multiples. Deux aspects du potentiel pédagogique du RLF méritent d'être soulignés.

Premièrement, sa logique de structuration est transposable d'une langue à une autre, un grand nombre de fonctions lexicales ayant un caractère standard. Apprendre à naviguer dans le système lexical de sa langue maternelle, à comprendre, puis à prévoir les relations de sens prépare le locu-

teur à se lancer dans l'étude d'une L2, en lui fournissant les outils conceptuels nécessaires à cette navigation.

Deuxièmement, le RLF est destiné à devenir une ressource favorisant la structuration et la consolidation des connaissances lexicales et par là même le développement du vocabulaire actif. En effet, plus les liens rattachant une lexie à ses congénères sont nombreux, plus son ancrage dans le vocabulaire d'un individu est solide (Meara 2007). C'est cette approche qui avait déjà été utilisée dans le cadre de la rédaction du *Lexique actif du français* (Mel'čuk & Polguère 2007), ouvrage à finalité pédagogique qui décrit un échantillon du lexique français en termes des dérivations sémantiques et collocations. Le RLF est destiné à aller bien au-delà, non seulement en visant une couverture lexicale beaucoup plus vaste, mais aussi en s'affranchissant du support dictionnaire grâce à sa structure de graphe lexical¹⁹.

Bibliographie

Adolphs, S., Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics*, 24 (4), 425–438.

Aitchison, J. (2003). *Words in the Mind : An Introduction to the Mental Lexicon*. 3^e édition, Oxford UK : Blackwell.

Atkins, B. T. S. (1996). Bilingual Dictionaries : Past, Present and Future. In M. Gellerstam, J. Järberg, S.-G. Malmgren, K. Norén, L. Rogström & C. R. Pappmehl (dir.) : *Euralex'96 Proceedings*, Göteborg : Université de Göteborg, 515–590.

Brunet, É. (1981). *Le Vocabulaire français de 1789 à nos jours d'après les données du « Trésor de la Langue Française »*, Genève/Paris : Slatkine/Champion.

Chetouani, L. (1997). *Vocabulaire général d'enseignement scientifique*. Paris : L'Harmattan.

Collectif Robert (2009). *Le Robert Benjamin*. Paris : Le Robert.

Cortier, C. & Bouchard, R. (dir.) (2008). « Quel oral enseigner, cinquante ans après le Français fondamental ? », *Le Français dans le monde. Recherches et applications*, 43, Cle-International.

Déro, M. (1998). *Inventaire du vocabulaire et développement des connaissances du Cours Préparatoire à la Sixième de Collège*. Thèse de doctorat, Université de Haute Bretagne – Rennes 2.

Gader, Nabil, Lux-Pogodalla, V. & Polguère, A. (2012). Hand-Crafting a Lexical Network With a Knowledge-Based Graph Editor. In *Proceedings of CogALex III*, Mumbai.

Galisson, R. (1971) *Inventaire thématique et syntagmatique du français fondamental*. Paris : Librairies Hachette et Larousse.

Gougenheim, G., Michéa, R., Rivenc, P. & Sauvageot, A. (1959). *L'élaboration du français fondamental, 1^{er} degré*. Paris : Didier.

Hausmann, F. J. (1989). Le dictionnaire de collocations. In F. J. Hausmann, O. Reichmann, H. E. Wiegand & L. Zgusta (dir.) : *Wörterbücher : ein internationales Handbuch zur Lexicographie / Dictionaries: An International Encyclopaedia of Lexicography / Dictionnaires : Encyclopédie internationale de lexicographie*. Handbücher zur Sprach- und Kommunikationswissenschaft, Band 5.1, Berlin/New York NY : Walter de Gruyter, 1010–1019.

Laufer, B. (1996). The lexical threshold of L2 reading : where it is and how it relates to L1 reading ability. In K. Sajavaara & C. Fairweather (dir.), « Approaches to Second Language Acquisition »,

19 Voir à ce propos l'approche de la lexicographie des dictionnaires virtuels, mentionnée dans la section 1.3 et présentée dans Polguère (2012b).

Jyvaskyla Cross Language Studies, 17, 55–62.

Lefrançois, P., Tremblay, O., Lombard, V. (2011). *Constitution de listes de mots pour l'apprentissage de l'orthographe et du lexique au primaire et au début du secondaire*. Rapport de recherche pour le Ministère de l'Éducation, du Loisir et du Sport du Québec (MELS), Université de Montréal, Montréal, 14 février 2011.

Lo Bianco, J., Scull, J. & Ives, D. (2008). *The Words Children Write 2008 : Research Summary of the Oxford Wordlist*. Melbourne : Oxford University Press.

Lux-Pogodalla, V., Polguère, A. (2011). Construction of a French Lexical Network. Methodological Issues. *Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011. An ESSLLI 2011 Workshop*, Ljubljana, 54–61.

McCarthy, M. (1999) What constitutes a basic vocabulary for spoken communication ? *English Language and Literature*, 1, 233–249.

Meara, P. (2007). Towards a new approach to modelling vocabulary acquisition. In N. schmitt & M. McCarthy, Michael (dir.) : *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge UK : Cambridge University Press, 109-121.

Mel'čuk, I. (1994). *Cours de morphologie générale (théorique et descriptive) Vol. 2. Deuxième partie : Significations morphologiques*. Montréal & Paris : Les Presses de l'Université de Montréal & CNRS Éditions.

Mel'čuk, I. (1995). Phrasemes in Language and Phraseology in Linguistics. In M. Everaert. E.-J. van der Linden, A. Schenk & R. Schreuder (dir.) : *Idioms : Structural and Psychological Perspectives*, Hillsdale N.J./Hove UK : Laurence Erlbaum Associates, 167–232.

Mel'čuk, I. (1996). Lexical Functions : A Tool for the Description of Lexical Relations in the Lexicon. In L. Wanner (dir.) : *Lexical Functions in Lexicography and Natural Language Processing*, Language Companion Series 31, Amsterdam/Philadelphie : John Benjamins, 37–102.

Mel'čuk, I. (1997) *Vers une linguistique Sens-Texte*. Leçon inaugurale, Paris : Collège de France.

Mel'čuk, I. et coll. (1984, 1988, 1992, 1999). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I-IV*, Montréal : Les Presses de l'Université de Montréal.

Mel'čuk, I., Clas, A., Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve : Duculot.

Mel'čuk, I., Polguère, A. (2006). Dérivations sémantiques et collocations dans le DiCo/LAF. *Langue française*, 50, 66–83.

Mel'čuk, I., Polguère, A. (2007). *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. Collection « Champs linguistiques », Bruxelles : De Boeck & Larcier.

Milićević, J. (2006) A Short Guide to the Meaning-Text Linguistic Theory. *Journal of Koralex*, 8, 187–233.

Milićević, J. (2007). *La paraphrase. Modélisation de la paraphrase langagière*. Berne : Peter Lang.

Milton, J. (2009) *Measuring Second Language Acquisition*. Bristol UK/Buffalo NY/Toronto : Multilingual Matters.

Nation, I. S. P. (1982). Beginning to learn foreign language vocabulary : a review of research. *RELC Journal*, 13, 61–68.

Odgen, C. K. (1930). *Basic English : A General Introduction with Rules and Grammar*. London : Paul Treber & Co.

Picoche, J. (1995) L'Enseignement du vocabulaire en français langue maternelle au niveau de l'enseignement secondaire. In N. Andrieux-Reix & G. Hasenhor (dir.), *Études de lexicologie et dialectologie*, Paris : Conseil international de la langue française, 365–377.

Picoche, J. (2007). *Enseigner le vocabulaire, la théorie et la pratique*. Champigny-sur-Marne : Éditions Allouche, publication électronique sur CD-ROM [version révisée de l'ouvrage paru en 1993 sous le titre *Didactique du vocabulaire français*, collection « fac.linguistique », Paris : Nathan].

Polguère, A. (2007). Lexical function standardness. In L. Wanner (dir.) : *Selected Lexical and Grammatical Issues in the Meaning-Text Theory. In Honour of Igor Mel'čuk*, Language Companion Series 84, Amsterdam/Philadelphie : John Benjamins, 43–95.

Polguère, A. (2008). *Lexicologie et sémantique lexicale. Notions fondamentales*. Collection « Paramètres », Montréal : Les Presses de l'Université de Montréal.

Polguère, A. (2011). *Mémo RLF. Étapes dans la rédaction d'une entrée de vocable*. Projet RELIEF/RLF, Nancy : ATILF CNRS (dernière mise à jour : 20 février 2012).

Polguère, A. (2012a). Lexicographie des dictionnaires virtuels. In Y. Apresjan, I. Boguslavsky, M.-C. L'Homme, L. Iomdin, J. Milićević, A. Polguère & L. Wanner, L. (dir.) : *Meanings, Texts, and Other Exciting Things. A Festschrift to Commemorate the 80th Anniversary of Professor Igor Alexandrovič Mel'čuk*, Studia Philologica, Moscou : Jazyki slavjanskoj kultury Publishers, 509–523.

Polguère, A. (2012). Like a Lexicographer Weaving Her Lexical Network. In *Proceedings of CogA-Lex III*, Mumbai.

Polguère, A. (à paraître). Les petits soucis ne poussent plus dans le champ lexical des sentiments. In F. Baider & G. Cislaru (dir.) : *Cartographie des émotions*, Villeneuve-d'Ascq : Presses Universitaires du Septentrion.

Brève histoire d'une erreur lexicale

Polysémie et liens lexicaux dans l'enseignement du vocabulaire

Sikora, Dorota

Université du Littoral-Côte d'Opale & HLLI (EA 3040)
Dorota.Sikora@univ-littoral.fr

1 Introduction¹

Partant de quelques erreurs produites par des apprenants de français langue seconde (FLS) de niveau B1, cet article propose une réflexion sur l'objet même des apprentissages lexicaux. L'analyse des erreurs en question, ainsi que les considérations plus générales qu'elles inspirent sont menées dans une double perspective, à la fois didactique et lexicologique.

Les erreurs sont toujours intéressantes pour un enseignant, car elles constituent une source de renseignements précieux, permettant de réajuster la démarche pédagogique. Celles qui nous occuperont ici le sont d'autant plus qu'elles surgissent dans un vocabulaire considéré comme facile. Les vocables VISITE et VISITER, travaillés en classe de langue dès le niveau A1, étaient perçus par nos étudiants B1 comme connus et ne présentant pas de difficulté majeure. Nous nous efforcerons donc de comprendre pourquoi ils ont donné lieu à des emplois inappropriés. Nous chercherons une réponse à cette question dans le lexique, en observant les entités qui le composent et leur fonctionnement. Il apparaîtra au fur et à mesure que l'impression de connaître un « mot » peut être trompeuse : tout en semblant familier, il s'avère mal positionné et de ce fait, peu intégré dans le stock lexical de l'apprenant.

Au-delà d'un diagnostic portant sur un ensemble d'erreurs, notre réflexion nous amènera à poser la question plus générale de l'objet de l'enseignement et de l'apprentissage du vocabulaire. Faut-il apprendre des mots (au sens courant du terme), des familles des mots, des lexies ou des réseaux lexicaux ? Une fois de plus, nous chercherons des réponses du côté de lexique, en admettant que la démarche pédagogique et le processus d'apprentissage doivent porter sur ce dont le lexique d'une langue est fait : les lexies (ou unités lexicales) et les relations qui les connectent dans le système lexical.

Notre réflexion est largement inspirée d'une part par une expérience didactique d'enseignement de FLS et de l'autre par la pratique lexicographique dans le développement d'une ressource lexicale qui offre de nombreuses applications pédagogiques. La base de données lexicales *Réseau Lexical du Français* (RLF), en cours de construction actuellement au laboratoire ATILF – CNRS, est une ressource de type dictionnaire virtuel qui non seulement assure une large couverture du lexique français, décrit de manière rigoureuse et structurée, mais qui permet de comprendre son organisation et son fonctionnement. Précisons que le RLF est élaboré dans le cadre théorique et méthodologique de la Lexicologie Explicative et Combinatoire (désormais la LEC).

Dans la section 2, nous présentons l'activité qui avait été proposée à nos étudiants B1 et les erreurs qui ont surgi dans leurs productions. C'est dans cette section également que nous poserons la question de savoir ce qu'est l'erreur lexicale et ce que l'on entend par connaissances lexicales, notamment dans les travaux de didacticiens de langue seconde, quelle qu'elle soit. Nous y procéderons également à une mise au point terminologique, indispensable pour notre double perspective didactique et lexicologique. La section 3 est consacrée à l'analyse des sens des vocables VISITE, VISITER et VISITE MÉDICALE et des relations lexicales qui les situent dans le réseau que constitue le lexique du français. La quatrième et dernière partie sera destinée à réunir, en guise de conclusions, les renseignements que le cas particulier des erreurs examinées permet de tirer au sujet de la pédagogie du vocabulaire.

2 Une erreur lexicale dans une double perspective didactique et lexicologique

2.1 Présentation de l'activité proposée aux étudiants

Après la lecture et la discussion du règlement interne d'une entreprise, le travail s'était focalisé sur l'expression de modalités en français avec l'objectif de maîtriser les opérateurs de modalité déontique (ou modalité du faire, Kronning 1996). Dans cette optique, les étudiants devaient insérer un même contenu propositionnel Y dans la portée des expressions modalisantes suivantes :

L'individu X est prié de faire Y / L'individu X doit/ne doit pas faire Y / L'individu X est obligé de faire Y / Il faut faire Y / Y est obligatoire / Il est obligatoire de faire Y / Y est interdit / Il est interdit de faire Y

Parmi les dix phrases soumises à ces procédés de paraphrase, l'une s'est montrée particulièrement difficile pour les étudiants. En effet, l'énoncé proposé sous la forme *Les visites médicales sont obligatoires pour les salariés* a donné lieu aux reformulations suivantes :

1. *Les salariés doivent visiter chez le médecin.
2. *Les salariés sont priés de visiter chez le médecin.
3. *Le salarié est obligé de visiter chez le médecin.
4. *Il faut visiter chez le médecin.
5. *Il est obligatoire de visiter chez le médecin

La question qu'il convient de poser face à ces difficultés lexico-syntaxiques est celle de savoir pourquoi, tout en maîtrisant l'expression de modalité du faire, c'est-à-dire un contenu sémantique complexe tant du point de vue conceptuel que sur le plan de son expression linguistique, des étudiants avancés rencontrent des difficultés à manier des lexies qui le sont nettement moins ?

Selon Milićević (2007), la paraphrase consiste à former des phrases équivalentes, c'est-à-dire des phrases qui peuvent être considérées comme synonymes ou quasi-synonymes. Dans la mesure où il s'agit d'opérer des modifications à la fois lexicales et syntaxiques, la compétence de paraphrase suppose chez le locuteur, qu'il soit de langue maternelle française ou non, la maîtrise des relations de sens (liens paradigmatiques d'hyponymie, de synonymie, d'antonymie, conversions nominales, verbales, adjectivales et adverbiales, etc) et de combinatoire syntaxique. Pour produire les deux phrases ci-dessous :

6. On peut déposer son courrier jusqu'à 18 heures.
7. Le dépôt du courrier est possible jusqu'à 18 heures.

il est nécessaire, entre autres, de nominaliser le verbe DÉPOSER et de lui associer le nom COURRIER en tant que complément. Dans la section suivante, nous passerons en revue les types d'erreurs lexicales pour identifier le ou les mécanismes qui ont conduit nos étudiants à produire les phrases (1) à (5).

2.2 Qu'est-ce qu'une erreur lexicale ?

Depuis les travaux de Porquier (1977) et de Perdue (1980), les erreurs d'apprenants sont attentivement analysées en tant que sources d'informations précieuses pour l'enseignement². Cette approche constructiviste de l'erreur a pris un nouvel élan dans les années 1990, lorsque l'exploitation des corpus d'apprenants a permis de développer des études systématiques des productions écrites et orales des étudiants de FLS. Le projet French Interlanguage Database lancé en 1998 a permis de compiler un corpus de 450 000 mots en provenance de productions écrites d'apprenants FLS (Granger et coll., 2001). Les erreurs repérées dans ce corpus ont été réparties selon le domaine dont elles relèvent (l'orthographe, le lexique, la grammaire, la syntaxe) et selon les caractéristiques grammaticales (par exemple, le genre a été pris en compte en ce qui concerne les noms). D'autres études s'intéressent plus spécifiquement à certains aspects linguistiques de l'erreur. Ainsi, Le Bot et Richard (2011) proposent plus spécifiquement une analyse morphologique des écrits d'apprenants de FLS. Il est également fréquent d'étudier les erreurs

récurrentes dans les productions d'étudiants d'une même langue maternelle (Masseron et Luste-Chaa, 2008).

Pour notre propos, nous porterons un intérêt tout particulier à l'étude de Hamel et Milićević (2007), qui se penchent sur les erreurs lexicales, c'est-à-dire celles qui touchent aux emplois inappropriés des unités lexicales (lexies). Les auteures définissent l'erreur lexicale comme :

Tout emploi inapproprié d'une lexie L ayant comme cause des connaissances insuffisantes de ses propriétés sémantiques, formelles et/ou de co-occurrence. Un emploi est jugé inapproprié s'il mène à l'agrammaticalité, mais aussi s'il résulte en une maladresse. (Hamel et Milićević, 2007 : 29)

Cette définition ne fait aucune référence à une norme, qui serait nécessairement liée à l'usage. Une telle approche semble préférable pour deux raisons. Tout d'abord, *norme* renvoie à un critère extralinguistique, souvent de nature sociale. Ensuite, par son sens lexical, l'emploi du terme *norme* véhicule un contenu évaluatif, puisqu'un écart par rapport à une norme déclenche plus ou moins ouvertement un jugement négatif. La définition de Hamel et Milićević (2007) s'appuie, au contraire, sur des éléments internes à la langue telles que les propriétés sémantiques, formelles et combinatoires. L'erreur lexicale est mesurée à l'aune de son inappropriation, de son agrammaticalité ou tout simplement de l'impression de maladresse qu'elle produit. Ainsi, pour qualifier d'erronée une production d'apprenant, on se remet à l'intuition du locuteur natif et à son jugement d'acceptabilité.

Dans cette optique, Hamel et Milićević (2007) répartissent les erreurs lexicales en trois catégories :

1. Les erreurs de signifiant qui se manifestent dans une restitution inadéquate, à l'oral et/ou à l'écrit, de L.
2. Les erreurs liées au signifié de la lexie L employée de manière inappropriée.
3. Les erreurs de combinatoire qui apparaissent suite à un choix inapproprié de cooccurrents, c'est-à-dire lorsque L se voit associer des lexies qui ne sont pas celles qu'elle sélectionne habituellement. On connaît ainsi les problèmes que pose aux apprenants le choix de collocatifs appropriés (cf., à titre d'exemple, Bolly, 2011).

Conformément à cette typologie, les problèmes constatés dans les phrases (1) à (5) relèvent à la fois du deuxième et du troisième type : le verbe VISITER a été doté d'un sens qu'il n'a pas et de ce fait, mal combiné avec les autres éléments de ces phrases.

Si les typologies permettent d'anticiper certaines difficultés susceptibles d'apparaître dans les productions d'apprenants et de prendre les devants, le propre de l'erreur lexicale est néanmoins son caractère inattendu. Tout en prévenant certains problèmes, l'enseignant en voit apparaître d'autres là où il ne s'attendait pas à les avoir. Il en est ainsi de l'exercice de paraphrase que nous avons proposé à nos étudiants B1 : les structures ciblées dans cette activité, c'est-à-dire les opérateurs de modalité déontique, n'ont pas posé de problème majeur, alors que les erreurs ont surgi là où il a fallu manier VISITE, VISITER et VISITE MÉDICALE. Les deux premiers sont pourtant des « mots » bien connus dès le niveau A1, puisque le champ sémantique des 'loisirs' est habituellement abordé dès les premières étapes d'apprentissage, pour ressurgir ensuite à l'occasion des thèmes tels que 'vacances', 'tourisme', etc. De plus, il s'agit d'éléments lexicaux qui ne présentent aucun des facteurs susceptibles de rendre leur apprentissage difficile (Laufer, 2007, Milton, 2009). VISITE et VISITER font partie du vocabulaire de base (McCarthy, 1999) du français (à ce propos, voir également Picoche, 2007, Polguère et Sikora, 2013).

Quant à VISITE MÉDICALE, la situation semble plus complexe dans la mesure où il s'agit d'une locution dont le sens ne peut pas être obtenu par un simple calcul sémantique des apports respectifs de ses composantes. Malgré son caractère non compositionnel, cette locution reste transparente, notamment pour un étudiant de niveau B1. Nous reprendrons cette question dans la section 3.2.

Il est possible d'ores et déjà d'avancer une hypothèse sur l'origine des erreurs dans les phrases (1) à (5). Elles ne semblent pas liées à la compétence syntaxique, puisque la forme SN₁ V V_{inf} Prep SN₂ de la phrase (1) *Les salariés doivent visiter chez le médecin* est l'un des patrons de surface réalisé par de nombreux énoncés phrastiques en français. Elle n'appellerait aucune critique si, au lieu de VISITER, on

trouvait le verbe ALLER en position de V_{inf} : la phrase *Les salariés doivent aller chez le médecin* serait alors une réalisation du même schéma.

Le problème réside dans une relation de sens établie entre VISITE, VISITER et VISITE MÉDICALE. L'erreur a donc un caractère lexical, et elle est triple. En effet, dans un premier temps, les étudiants semblent associer VISITE et VISITE MÉDICALE, en considérant qu'il s'agit d'une même unité lexicale, *modulo* la modification adjectivale dans le second cas. Deuxièmement, ils font une hypothèse erronée sur les relations paradigmatiques entre VISITE et VISITER, en admettant que le second est une traduction verbale du premier pris dans le sens de *visite chez mes grands-parents*. Il s'ensuit un transfert du régime prépositionnel de cette acception de VISITE sur le verbe VISITER doté d'un sens inexistant.

2.3 Qu'est-ce que connaître un mot ?

En didactique de langues, il est admis depuis Nation (1982), de mesurer le stock lexical de l'apprenant d'une L2 en « mots », tout en entendant par là des familles de mots. Une famille de mots comprend toutes les formes fléchies, ainsi que les dérivés forgés à l'aide des affixes les plus fréquents. Ainsi VISITE, VISITER, VISITEUR, VISITEUSE, VISITATION forment une famille de mots qui, dans une évaluation quantitative des acquisitions lexicales des apprenants, compte pour un seul mot. C'est en s'appuyant sur cette notion que Nation (2001) définit les trois dimensions de la connaissance lexicale. Connaître un mot, c'est :

1. Maîtriser la forme du mot à l'écrit comme à l'oral, tout en identifiant ses éléments constitutifs, notamment les morphèmes qui y figurent.
2. Associer le sens à une forme et, inversement, être en mesure de trouver un mot pour le sens qu'on veut exprimer. Il s'agit donc de le maîtriser tant sur l'axe sémasiologique qu'onomasiologique. La capacité de comprendre les connotations qui s'y attachent s'inscrit également partie dans cette dimension de la compétence lexicale.
3. Enfin, être capable d'utiliser un mot suppose la connaissance de ses caractéristiques grammaticales et de son paradigme flexionnel, ainsi que l'aptitude à le combiner avec d'autres éléments sur le plan phrastique.

Le terme de *famille de mots* a montré son utilité pour l'évaluation des connaissances lexicales : il fournit une unité de mesure pour vérifier à la fois leur étendue et leur profondeur. Cependant, il n'est pas sans poser de sérieux problèmes, lorsque l'on prend en compte la phase d'apprentissage qui précède nécessairement celle d'évaluation.

En effet, que pourra-t-on dire du degré de connaissance de la famille de mots VISITER³, en analysant les productions présentées dans les exemples (1) à (5) ? Il semblerait que les étudiants, tout en maîtrisant la forme, ont des difficultés à en systématiser les sens et, par voie de conséquence, les propriétés combinatoires.

Par ailleurs, les travaux sur le lexique mental (Aitchison, 2003) tendent à montrer que le stock lexical des locuteurs d'une langue est organisé en réseau dont les différents éléments sont reliés par une multitude de liens. Une telle structuration caractérise également le vocabulaire acquis ou en cours d'acquisition d'un apprenant de L2. Meara (2007) en propose le modèle explicatif suivant :

Unknown words are words that have no connections of any kind to the learner's lexicon. Known words are words that are connected, but the number of these connections may vary. That gives us a natural mechanism for talking about words which are 'known' to a greater or lesser extent. Poorly-known words are words with few connections, while better-known words are simply words with many connections. (Meara, 2007 : 119)

Le schéma présenté ci-dessous (Figure 1) permet de visualiser le modèle de Meara, en illustrant le cas d'un mot (W) qui n'est que très faiblement associé au système lexical (L) en cours de construction chez l'apprenant.

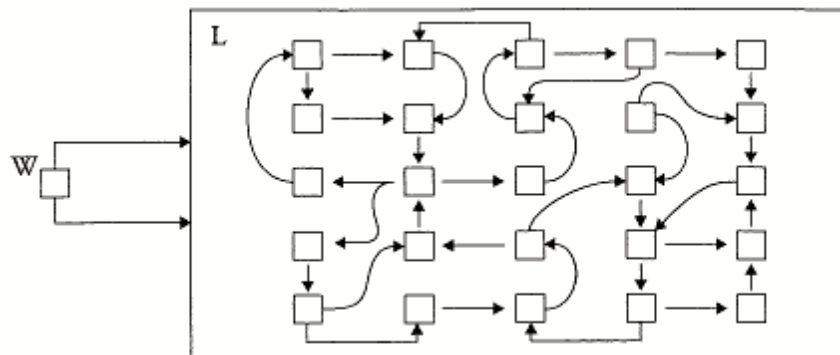


Figure 1 : Structure d'un réseau lexical d'apprenant de L2 dans Meara (2007 : 120)

La description que propose Nation (2001) de ce qu'est la compétence lexicale d'une part et, de l'autre, les travaux sur le lexique mental (Aitchison, 2003) parmi lesquels nous situons le modèle de Meara (2007) indiquent qu'il est indispensable d'affiner la notion de famille de mots. En effet, il semble difficile de soutenir que l'acquisition d'un mot (au sens de famille de mots) et son intégration dans le vocabulaire de l'apprenant consiste à le rattacher par de liens multiples à d'autres mots (c'est-à-dire à d'autres familles de mots). Ainsi, dans l'exemple (8), le nom VISITE peut très bien être associé, par un lien de quasi-synonymie, à DÉCOUVERTE, ce qui n'est aucunement le cas en (10).

8. La visite du Louvre a pris toute une journée.
9. La découverte du Louvre a pris toute une journée.
10. La visite de la secrétaire d'état en France a duré deux heures.
11. *La découverte de la secrétaire d'état en France a duré deux heures.

La raison en est à chercher dans la structure du lexique lui-même. En effet, ce que l'on entend par famille de mots n'a pas de réalité sur le plan lexical : il n'existe pas d'entité linguistique correspondante.

C'est du côté du lexique d'une langue, en l'occurrence celui du français, que nous nous tournons à présent pour rechercher la solution des difficultés que nous venons de signaler. Il convient tout d'abord d'explicitier la nature des entités lexicales qui, dans le schéma de Meara (2007), correspondent à W, c'est-à-dire aux éléments qui constituent son réseau L. La question a deux facettes d'importance cruciale pour notre propos, car il s'agit de préciser à la fois de quoi est fait le lexique d'une langue et quel est l'objet des apprentissages lexicaux.

2.4 Lexique et ses réseaux de lexies

Les notions élaborées dans le cadre de la LEC permettent d'identifier avec précision les entités constitutives du système lexical (Mel'čuk et al, 1995). Le lexique d'une langue, c'est l'ensemble de toutes les lexies (ou unités lexicales). Une lexie est une association d'un signifiant (au sens saussurien du terme), d'un signifié et de propriétés combinatoires. Dans nos exemples (8) et (10), le signifiant VISITE recouvre donc en réalité deux lexies différentes. On distingue deux types majeurs de lexies : des lexèmes (VISITE, VISITER) et des locutions (ou phrasèmes complets) tels que CASSER LES PIEDS ou VISITE MÉDICALE⁴.

Il existe cependant des phrasèmes qui ne sont pas des locutions, *ipse eo* des unités lexicales. Les collocations sont des phrasèmes semi-contraints, qui comprennent une lexie, dite base de la collocation, choisie librement par le locuteur en fonction du sens qu'il veut transmettre, et un collocatif, c'est-à-dire une lexie dont la sélection est contrainte par celle de la base. Ainsi, pour intensifier l'importance d'une BLESSURE (base de la collocation), le locuteur doit choisir comme collocatif la lexie adjectivale GRAVE, et pas le lexème SÉRIEUX.

Mais le lexique d'une langue est plus qu'un ensemble de lexies, puisque celles-ci sont reliées entre elles par de nombreuses relations. Elles forment un véritable réseau structuré par des liens multiples. Plus précisément, une lexie – lexème ou locution – fonctionne dans son propre réseau lexical, qui s'intègre dans un système plus large. La polysémie étant largement présente dans le lexique, la plupart des lexies sont tout d'abord reliées par des liens de sens à au-moins un copolysème. On appelle *copolysèmes* des lexies qui, comme VISITE dans les exemples (8) et (10), ont le même signifiant avec des sens qui, bien que différents, possèdent des composantes sémantiques communes. Précisons que la LEC décrit la polysémie uniquement d'un point de vue synchronique : les liens entre les copolysèmes, dits *ponts sémantiques*, doivent être perçus par les locuteurs contemporains, sans qu'il soit nécessaire de remonter vers un étymon commun. S'il n'est pas possible de construire un pont sémantique en synchronie, c'est que nous sommes face à des homonymes.

Dans une ressource dictionnaire construite conformément aux principes de la LEC, les lexies copolysèmes sont groupées en vocables et organisées de manière à faire apparaître les liens de sens qui existent entre elles, tels qu'ils sont conceptualisés par un locuteur contemporain. Il s'agit, bien évidemment, d'une hypothèse qu'on construit après avoir identifié l'acception de base d'un vocable, c'est-à-dire celle qui permet d'organiser les relations entre les copolysèmes. Le vocable est donc un regroupement sémantiquement structuré des lexies qui ont le même signifiant, mais dont les signifiés présentent un écart plus ou moins important.

Conformément aux conventions typologiques de la LEC, nous transcrivons les noms de vocables en petites capitales. Nous utiliserons le même format pour les noms de leurs acceptions (lexies), tout en les accompagnant des numéros qui facilitent l'identification de leur place dans la structure polysémique du vocable. Ainsi, les chiffres romains indiquent un écart métaphorique ou métonymique, alors que les chiffres arabes et les petites lettres marquent une différence sémantique plus fine, par exemple une spécialisation ou une extension à l'intérieur d'une famille métaphorique ou métonymique.

3 Polysémie et liens lexicaux

3.1 VISITE, VISITER et leurs structures polysémiques

Notre hypothèse de travail consiste à admettre que pour construire et consolider le réseau lexical de l'apprenant, il est indispensable de maîtriser et d'exploiter les liens qui structurent le lexique lui-même. Nous nous tournons donc tout d'abord vers les structures polysémiques des vocables VISITE et VISITER. Ils regroupent en effet des lexies qui sont de nature prédicative. La LEC définit un prédicat comme un sens liant d'autres sens, qui sont ses arguments. En réalité, les lexies appartenant aux quatre classes majeures peuvent être des prédicats sémantiques : les verbes, les adjectifs et les adverbes impliquent toujours au moins un argument sémantique. Une majorité des noms ont également cette caractéristique. Ainsi, la lexie VISITE employée dans l'exemple (8) est un prédicat nominal⁵ (ou quasi-prédicat, cf. Polguère, 1992) avec deux actants sémantiques : l'individu X qui rend visite et l'individu Y qui la reçoit. Cette structure sémantique, exprimée par la forme propositionnelle fait partie de la description de l'unité lexicale. Il est indispensable de la connaître pour maîtriser les relations paradigmatiques et syntagmatiques qui situent la lexie dans le réseau lexical de la langue et que l'on exploitera pour l'intégrer dans le stock lexical que l'apprenant doit construire.

Plutôt que de proposer une définition lexicographique pour chaque lexie, nous en caractériserons le sens en utilisant le système d'étiquettes sémantiques proposé dans Polguère (2012). Une étiquette sémantique est une reformulation minimale normalisée d'un contenu sémantique, qui permet de rattacher une lexie à une classe de sens⁶. Cela revient donc à identifier et à nommer le type ontologique de ce sens.

Le choix de l'étiquette sémantique pour une lexie n'est pas anodin dans la mesure où celle-ci situe le sens concerné dans un paradigme particulier et laisse pressentir le type d'actants susceptibles de lui être attachés. Notre objectif étant de réfléchir à la nature des erreurs que nous avons rapportées dans les exemples (1) à (5) et, dans une perspective plus large, à l'objet des démarches pédagogiques visant

l'apprentissage du vocabulaire, nous nous limiterons ici à cette description minimale. Il est clair cependant qu'en situation de classe, on ne saurait faire l'économie d'une paraphrase définitoire explicitant le sens d'une lexie.

3.1.1 Structures polysémiques des vocables VISITE et VISITER

Il nous est impossible, dans le présent cadre, de proposer une analyse sémantique détaillée des structures polysémiques des vocables VISITE et VISITER, pas plus que nous ne pouvons exposer le raisonnement qui nous a conduit à distinguer les copolysèmes que nous présentons dans les Tableaux 1 et 2⁷. Notre objectif est d'identifier la source des erreurs apparues dans les phrases (1) à (5). Ce faisant, nous tenons à montrer également que parvenir aux connaissances lexicales dont l'étendue correspond à celle postulée par Nation (2001) passe par l'identification et les apprentissages successifs non pas de mots, ni de familles de mots, mais des lexies.

Rappelons que l'organisation du réseau de copolysèmes que nous proposons pour ces deux vocables est faite dans une perspective synchronique. Il s'agit d'hypothèses qui modélisent les connaissances lexicales d'un locuteur francophone contemporain, capable d'utiliser ces lexies sans être nécessairement au courant de leurs cheminements en diachronie.

| Lexie | Forme propositionnelle | Étiquette sémantique classifiante | Exemple |
|---------------|--|--|--|
| VISITE I | ~ de l'individu X chez l'individu Y | 'réunion d'individus' | Hier, j'ai rendu visite à Ernst Jünger, qui réside à Wilflingen, un petit village de la Haute-Souabe. (Frantext) |
| VISITE II | Individu X, qui est une ~ de l'individu Y | 'individu qui est dans une certaine situation' | Il faut bien se tenir aux repas, quand il y a des visites . Yves est toujours en bout de table, avec ordre de ne rien dire et de « disparaître après le dessert ». (Frantext) |
| VISITE III.1 | ~ de l'individu X dans le lieu Y | 'localisation' | C'était sa première visite à Paris et elle a trouvé que l'opéra était molto caricato. (Frantext) |
| VISITE III.2a | ~ par l'individu X de l'endroit Y | 'activité d'un certain type' | Si pendant la visite du musée, vous avez une envie de vous désaltérer ou discuter avec vos amis, le Panoramique Café vous propose sa terrasse ensoleillée à tout moment de la journée. (Internet) |
| VISITE III.2b | ~ de l'endroit Y, à laquelle participent les individus X | 'événement social' | Outre la balade en petit train, le visiteur bénéficie d'une visite guidée qui lui fera découvrir l'un des joyaux lorrains de la fortification cuirassée du XIX ^e siècle. (L'Est Républicain) |
| VISITE IV | ~ par l'individu X de Y | 'action d'un certain type' | Lors de chaque visite d'un site Internet, des informations courantes sont collectées et échangées. (Internet) |

Tableau 1 : Lexies copolysèmes du vocable VISITE.

| Lexie | Forme propositionnelle | Étiquette sémantique classifiante | Exemple |
|-------------|--|---------------------------------------|---|
| VISITER I.1 | L'individu X ~ l'endroit Y | 'avoir une certaine activité' | Une partie du groupe fut même autorisée à visiter le fabuleux Temple du ciel, visite que notre trio des chefs dédaigna. (Frantext) |
| VISITER I.2 | L'individu X ~ le lieu Y | 'localisation' | Willem n' avait jamais visité les lieux en personne ; le trafic de ces gens-là avec la terre ferme se faisait plus loin, au nord-est du Zuiderzee. (Frantext) |
| VISITER II | L'individu X ~ le contenu informationnel Y | 'action d'utiliser quelque chose' | N'hésitez pas à m'envoyer des messages et à visiter mon blog. (FrWac) |
| VISITER III | L'individu X ~ l'individu Y' | 'situation relationnelle' | Laissant maman se reposer des fatigues de Plombières, il profite de ses vacances pour visiter les frères et les beaux-frères qu'il a dans la région. (Frantext) |
| VISITER IV | L'état psychique X ~ l'individu Y | 'être dans un certain état psychique' | Il arrive toutefois qu'un doute le visite , Hanjure, je l'ai vu dans ses yeux. (Frantext) |

Tableau 2 : Lexies copolysèmes du vocable VISITER.

Un premier bilan comparatif fait ressortir une différence quantitative : le vocable VISITE compte six acceptions, une de plus que VISITER. On constate ensuite que prises en synchronie, les structures polysémiques de ces vocables s'organisent différemment. Si, pour un locuteur contemporain du français, le sens VISITER I.1, dénotant une activité touristique, semble la première acception du vocable verbal, sa conversion nominale VISITE III.2a n'occupe pas la même place centrale dans la structure polysémique de VISITE, qui s'organise quant à elle à partir de la lexie dénotant une réunion d'individus. Il n'en est pas moins vrai que VISITER I.1 et VISITE III.2a expriment le même contenu sémantique dans un format, respectivement, verbal et nominal. Il s'agit donc d'un cas de conversion pure. On ne peut pas en dire autant de VISITE I et de VISITER III, qui ne peuvent pas produire de paraphrase exacte (cf. la section 3.1.2 ci-dessous). Enfin, VISITE II, VISITE III.1 n'ont pas d'équivalent verbal, et vice versa : VISITER IV ne peut pas du tout être nominalisé.

Les copolysèmes des vocables VISITE et VISITER forment ainsi deux réseaux qui ne sont pas des reflets en miroirs l'un de l'autre. Chacun d'entre eux comprend des lexies qui lui sont propres. Nous porterons désormais une attention toute particulière à VISITE I et VISITER III qui semblent directement impliquées dans les problèmes de paraphrase dont témoignent les exemples (1) à (5).

3.1.2 Dérivation sémantique et conversion verbale

Une fois les lexies des vocables VISITE et VISITER identifiées, nous sommes en mesure d'aborder le problème des leurs conversions nominales et verbales. Il s'agit de vérifier si celles-ci sont systématiques, c'est-à-dire s'il est toujours possible de construire une paraphrase exacte, en utilisant le nom et le verbe correspondants. Pour ce faire, il nous faudra aller au-delà des relations de forme construites sur le plan morphologique, pour nous intéresser aux liens de sens.

La notion de *dérivation sémantique* (cf. Mel'čuk et coll., 1984, 1988, 1992, 1998, Mel'čuk et Polguère 2006, 2007)⁸ s'appuie sur des relations de synonymie plus ou moins exacte : on considère que la lexie L₂

est un dérivé sémantique de la lexie L_1 , si le sens de L_2 est construit sur celui de L_1 . Autrement dit, le sens de L_2 se définit à partir de celui de L_1 . Prenons l'exemple de VISITER II et de ses liens de dérivation sémantique qui nous amènent tout d'abord vers la lexie nominale VISITE IV. Intuitivement, il semble judicieux d'admettre que la lexie VISITE IV est un dérivé sémantique proche (c'est-à-dire, présentant un degré de synonymie très élevé) du verbe VISITER II, puisque c'est lui qui – une fois défini – servira à expliquer le contenu sémantique du nom. Pour expliquer le sens d'un dérivé sémantique, que ce soit dans un contexte lexicographique ou pédagogique, on procédera donc en deux étapes :

1. Premièrement, on proposera une définiton de la lexie source de dérivation sémantique, en précisant, dans la forme propositionnelle entre parenthèses, quels en sont les actants sémantiques.

VISITER II ($X \sim Y$) = 'X regarde un site internet Y pour y trouver des informations'

2. Deuxièmement, en nous appuyant sur l'étape 1, nous pourrions expliquer le sens du dérivé sémantique⁹ sans courir le risque de circularité, si souvent reprochée aux dictionnaires traditionnels :

VISITE IV (~ par l'individu X de Y) = 'action de VISITER II Y, réalisée par X'

VISITER II et VISITE IV ont un sens identique, c'est-à-dire ils expriment la même structure sémantique sous une forme respectivement verbale et nominale. Ils sont ainsi liés par une relation de conversion pure, qui permet de construire des paraphrases exactes. Mises à part les structures thématiques et informationnelles différentes, les phrases (12) et (13) expriment le même sens.

12. Vous pouvez visiter notre site internet 24 heures sur 24.

13. Les visites de notre site internet sont possibles 24 heures sur 24.

Or, l'existence d'une lexie verbale et d'une lexie nominale sémantiquement liées ne signifie pas qu'il s'agit toujours d'une conversion pure. Rappelons les exemples de Frantext illustrant les sens VISITE I et VISITER III dans les Tableaux 1 et 2 ci-dessus.

14. Hier, j'ai rendu visite à Ernst Jünger, qui réside à Wilflingen, un petit village de la Haute-Souabe.

15. Laissant maman se reposer des fatigues de Plombières, il profite de ses vacances pour visiter les frères et les beaux-frères qu'il a dans la région.

On remarque une certaine assymétrie dans les relations sémantiques entre ces deux lexies. Il est, en effet, possible de reformuler l'exemple (15), en utilisant VISITE I, alors qu'une paraphrase de (14) au moyen du verbe VISITER III ne peut pas se faire sans que le sens de la phrase soit modifié.

16. ? J'ai visité Ernst Jünger, qui réside à Wilflingen, un petit village de la Haute-Souabe.

17. Laissant maman se reposer des fatigues de Plombières, il profite de ses vacances pour rendre visite aux frères et aux beaux-frères qu'il a dans la région.

Les deux lexies, nominale et verbale, impliquent deux actants humains, mais le VISITER III véhicule une information supplémentaire, absente de l'exemple (14), sur le deuxième actant (*Ernst Jünger*). Elle le présente comme celui qui ne peut pas, lui, rendre des visites : il est seulement en mesure de les recevoir. Il s'agit donc, pour le premier actant *je*, de rendre visite au sens I, mais dans une configuration très particulière, dans laquelle le second actant est présenté moins comme son bénéficiaire, que comme un patient¹⁰. Le sens de VISITER III contient ainsi des éléments absents de celui de VISITE I. La lexie verbale a donc une dénotation plus spécifique, *ipse eo* plus riche que la lexie nominale dont elle est sémantiquement dérivée. On comprend dès lors pourquoi on peut remplacer VISITER III par VISITE I : on passe alors d'un sens plus spécifique à un sens plus général. En revanche, substituer une lexie sémantiquement plus riche, comme VISITER III, à VISITE I, dont le sens est plus général, n'est possible que sous certaines conditions, dans des contextes bien définis.

3.1.3 VISITE I et son réseau de liens syntagmatiques

La mise en parallèle des structures polysémiques des vocables VISITE et VISITER a déjà permis, à la fin de la section 3.1.1, d'affiner notre diagnostic initial. Il apparaît que les erreurs produites en (1)-(5) sont dues au lien paradigmatique de conversion, tissé à tort par l'apprenant, entre les lexies VISITE I et VISITER III.

Dans la mesure où il n'est pas possible, en français contemporain, d'utiliser VISITER III pour rendre fidèlement le sens que VISITE I sous une forme verbale, il existe de nombreuses collocations qui permettent d'exprimer la participation des actants X et Y dans l'événement que constitue une visite. La lexie VISITE I se caractérise ainsi par un réseau de liens syntagmatiques nettement plus dense que ses copolysèmes. Ce réseau comprend notamment les verbes RENDRE II et FAIRE II.1 qui fournissent le moyen d'exprimer la participation du premier actant X, alors que AVOIR III.1 et RECEVOIR IV verbalisent celle de l'actant Y. La place importante qu'occupent les collocatifs verbaux dans le réseau lexical de VISITE I distingue donc clairement ce sens de ses copolysèmes.

Par ailleurs, VISITE I possède de nombreux collocatifs adjectivaux. Il est habituel de qualifier une visite de COURTE III.2, BREVE II, LONGUE III.1, PETITE III.2, de préciser son caractère AMICAL I, FAMILIAL, etc. Il est dès lors légitime de se demander si, parmi les collocatifs adjectivaux de VISITE I, ne se trouve pas l'adjectif MÉDICAL II ('relatif à la santé'). Autrement dit, il s'agit de savoir si le syntagme VISITE MÉDICALE qui est à l'origine des problèmes de paraphrase présentés dans les exemples initiaux, est une collocation de la lexie VISITE I. La question est d'autant plus judicieuse qu'il n'est pas rare de trouver dans des corpus des phrases telles que (18) :

18. Je vais rendre une petite visite à mon ophtalmo pour vérifier quelques petites choses. (FrWac)

Si MÉDICAL II était un collocatif de VISITE I, comme le sont les adjectifs PETIT III.2 ou AMICAL I, il serait possible de l'insérer auprès de VISITE I dans la phrase (18). Or, un tel ajout produit l'énoncé (19), pour le moins étrange, qui montre que MÉDICAL II n'est pas un collocatif de VISITE I.

19. ? Je vais rendre une petite visite médicale à mon ophtalmo pour vérifier quelques petites choses.

Le caractère étrange de la phrase (19) conduit à penser que MÉDICAL II n'est pas un simple collocatif de VISITE I. Il convient dès lors d'abandonner nos tentatives d'analyse compositionnelle.

3.2 Statut lexical de VISITE MÉDICALE

La question de savoir s'il faut voir dans le syntagme VISITE MÉDICALE une collocation de VISITE I est loin d'être triviale, car il y va en réalité de son statut lexical, et donc de sa structure conceptuelle et sémantique sous-jacente. Il s'agit, dans un premier temps, de comprendre son fonctionnement et sa place dans le lexique français, pour ensuite envisager une stratégie efficace pour l'apprendre.

Dans la sous-section 3.2.1, nous analyserons la structure sémantique de la locution VISITE MÉDICALE pour montrer qu'elle ne se laisse pas réduire à celle de VISITE I. En effet, elle se trouve au centre d'un réseau de relations paradigmatiques bien particulières. En 3.2.2, nous nous intéresserons à ses liens syntagmatiques, qui fournissent également de bonnes raisons de penser que VISITE MÉDICALE n'est pas simplement une collocation de VISITE I, ni d'ailleurs d'aucun autre sens du vocable VISITE, mais une locution dite *faible*. Contrairement à une locution forte telle que CASSER LES PIEDS, qui ne contient le sens ni de CASSER ni de PIED, dans une locution faible, on retrouve le contenu sémantique de ses composantes, en l'occurrence de VISITE I et de l'adjectif MÉDICAL II, mais elles s'articulent autour d'un élément supplémentaire qui occupe la place centrale dans la définition (Mel'čuk, 2013, Polguère, à paraître).

3.2.1 Structure sémantique et liens paradigmatiques

Malgré sa transparence, VISITE MÉDICALE n'est pas un phrasème compositionnel. Pour comprendre son sens, il ne suffit pas de combiner VISITE I ('rencontre des individus X et Y, qui a lieu parce que X va chez Y', où Y serait un médecin) et celui de l'adjectif MÉDICAL II ('relatif à la santé'). Son sens comporte,

certes, celui de ces éléments, mais il existe des indices linguistiques forts qui conduisent à penser qu'une visite médicale est conceptualisée avant tout comme un acte officiel.

L'exemple (20) offre une illustration intéressante du sens que véhicule la locution VISITE MÉDICALE, dans la mesure où le locuteur y sépare explicitement le suivi médical à proprement parler (*surveillance médicale*) de la visite médicale qui, comme il se doit dans le cas d'une formalité à accomplir, a un caractère officiel et obligatoire. En effet, sont soumis à cette obligation les individus qui, par ailleurs, bénéficient d'un suivi médical ayant déjà permis d'établir un diagnostic de leur état de santé.

20. À noter que si l'intéressé est soumis à une surveillance médicale spéciale (travailleur handicapé, personne affectée à certains travaux, travailleur de nuit, femme enceinte...), la visite médicale doit avoir lieu avant l'embauche. (FrWac)

À la différence des lexies qui forment la structure polysémique du vocable VISITE, la locution VISITE MÉDICALE est un quasi-prédicat liant trois actants sémantiques : l'individu X qui la passe, la personne ou l'établissement Y qui l'accueille et la raison Z pour laquelle elle a lieu. On parle en effet de visite médicale *d'embauche, de reprise, du permis de conduire*, etc.

On pourra dès lors paraphraser¹¹ le sens de la locution VISITE MÉDICALE (de X auprès de Y, en vue d'obtenir Z) de la manière suivante :

Acte officiel que l'individu X doit réaliser en se rendant auprès de la personne Y, pendant lequel Y contrôle l'état de santé de X en vue de Z.

Comme il se doit pour une locution faible (cf. supra), on y retrouve les contenus sémantiques des deux lexies qui la forment. L'apport de VISITE I est rendu par la composante 'en se rendant auprès de la personne Y', alors que 'l'état de santé' renvoie à l'adjectif MÉDICAL II. Néanmoins, notre définition s'organise autour de l'élément central, que nous proposons de classer à l'aide de l'étiquette sémantique 'acte officiel'. Il s'agit donc d'une structure sémantique particulière qui traduit linguistiquement un concept différent de celui de VISITE I.

Parmi les liens paradigmatiques qui relie VISITE MÉDICALE à d'autres lexies, nous ne trouvons pas les quasi-synonymes de VISITE I, tels que RENCONTRE I ou ENTREVUE. En revanche, sa relation d'hyponymie avec la lexie FORMALITÉ est largement attestée :

21. Lorsqu'un employeur embauche un salarié, il doit accomplir certaines formalités dont la visite médicale d'embauche fait partie. (Internet)

Notons également que dans le réseau des relations paradigmatiques qui relie VISITE MÉDICALE à d'autres lexies, on remarque l'absence du verbe VISITER et des noms d'actants. En effet, son premier actant sémantique X – par exemple, un(e) salarié(e) – ne peut pas être désigné(e) en tant que VISITEUR ou VISITEUSE.

3.2.2 VISITE MÉDICALE et ses liens syntagmatiques

Si la locution VISITE MÉDICALE n'a pas de dérivés sémantiques proches, elle est au centre d'un réseau de liens syntagmatiques relativement dense. En effet, elle constitue la base de nombreuses collocations, avec notamment des collocatifs verbaux. Les différentes étapes de la participation du premier actant X à une visite médicale peuvent être verbalisées à l'aide de lexies telles que SE PRÉSENTER, AVOIR II, EFFECTUER, PASSER III.1, SUBIR. C'est un réseau complètement différent de celui qui s'organise autour de VISITE I, puisque même le verbe AVOIR ne leur est pas commun. En effet, les deux lexies – VISITE MÉDICALE et VISITE I – pointent vers deux acceptions différentes du vocable AVOIR : la locution sélectionne son acception AVOIR II, alors que dans le réseau syntagmatique de la lexie nominale, on trouve AVOIR III.1, quasi-synonyme du verbe RECEVOIR IV, permettant d'exprimer, sous une forme verbale, la participation du deuxième actant.

De même, VISITE MÉDICALE résiste aux modificateurs adjectivaux qui caractérisent les différentes acceptions du vocable nominal VISITE. Elle accepte seulement une qualification adjectivale qui porte sur les aspects

fonctionnels de sont déroulement. Elle sélectionne de ce fait des collocatifs adjectivaux tels que : ANNUEL, OBLIGATOIRE, etc.

3.2.3 Une locution (faible) dans le réseau lexical

Pourquoi est-il important d'identifier le statut lexical d'un phrasème tel que VISITE MÉDICALE ? Une première raison découle d'un constat pratique : le lexique français contient une multitude de locutions qui sont, rappelons-le, des unités lexicales. Une deuxième raison est directement liée à cette omniprésence : les apprenants sont constamment exposés à ces lexies multilexémiques. Ils sont obligés d'abord de les comprendre, puis de les employer. Cela signifie la nécessité de saisir leurs sens et leurs structures sémantiques. Or, quelle que soit la L2 étudiée (voir, à titre d'exemple, Hobeika-Chakroun, 2013), les enseignants remarquent que face à ces unités multilexémiques, le premier réflexe consiste à adopter une démarche décompositionnelle, en repérant et en additionnant les éléments connus. Cette stratégie est compréhensible dans la mesure où elle présente un côté rassurant pour l'apprenant. Mais elle n'en est pas moins trompeuse, car – comme nous venons de le voir – une locution est une structure sémantique indépendante de ses composantes lexicales. Elle se trouve au centre d'un réseau structuré par des relations paradigmatiques et syntagmatiques qui la situent dans le lexique. Les lexies qui forment ce réseau, et qu'il faut maîtriser pour manier une locution comme VISITE MÉDICALE, ne sont pas celles qui se rattachent à VISITE I : il suffit de rappeler ici les collocatifs verbaux RENDRE II et RECEVOIR IV dont on ne peut pas se servir pour construire des collocations de VISITE MÉDICALE.

Les réflexions inspirées par les erreurs dont témoignent les exemples (1) à (5) nous conduisent ainsi bien au-delà de ce cas particulier. La question qui se pose en effet est celle de savoir quelles en sont les implications pédagogiques.

4 En guise de conclusion : vers un apprentissage d'unités lexicales

Les erreurs lexicales apparues dans les paraphrases (1) à (5) sont d'autant plus intéressantes qu'on ne saurait prétendre que VISITE et VISITER soient inconnus aux étudiants de FLS de niveau B1. Bien au contraire, il s'agit d'un vocabulaire perçu comme familier. Cette familiarité s'est pourtant avérée trompeuse, car si les copolysèmes des vocables VISITE et VISITER sont ressentis par les apprenants comme connus, ils sont, de toute évidence, mal positionnés dans leur stock lexical. Ce positionnement inapproprié explique la triple origine des erreurs que nous venons d'étudier.

1. Tout d'abord, il apparaît que l'entité lexicale à traiter n'a pas été clairement déterminée. L'acquisition s'est faite sur une base formelle donnant priorité au signifiant, et non pas à l'unité lexicale dans ses trois dimensions : signifiant, signifié et propriétés combinatoires.
2. Le statut lexical du phrasème VISITE MÉDICALE n'a pas été identifié par les apprenants. Ils y ont vu une collocation de la lexie nominale VISITE I, alors qu'en réalité il s'agit d'une locution.
3. Le rapprochement entre le nom VISITE et le verbe VISITER a été fait sur une base morphologique, et non pas sémantique. Or, comme l'indique Picoche (2007), la dérivation morphologique opère une modification formelle de la base, mais elle est déterminée sémantiquement. Autrement dit, la dérivation morphologique – que ce soit par l'adjonction d'un suffixe ou par troncation – est secondaire par rapport à la dérivation sémantique, c'est-à-dire par rapport à la relation entre deux sens qui se matérialise grâce à l'adjonction ou à la suppression d'un morphème. Nos explorations des réseaux lexicaux des différents copolysèmes des vocables VISITE et VISITER ont montré que le lien de conversion verbale n'est pas systématique. Là où il est possible, il peut s'accompagner – comme cela a lieu pour VISITE I et VISITER III d'une modification sémantique qui empêche la paraphrase exacte.

Ce diagnostic touche en fait à un seul et même problème, celui du type d'entités sur lesquelles doit se focaliser la démarche pédagogique visant l'apprentissage du vocabulaire. Elle doit en effet prendre pour objet la lexie, que celle-ci soit un lexème comme le sont VISITE I, VISITER I.1 et les autres acceptions des

vocables VISITE et VISITER, ou une locution telle que VISITE MÉDICALE. Nous tenons ainsi la réponse à la question que nous avons posée dans la section 2.3 – celle de savoir quelle est la nature de l'entité W dans le schéma proposé par Meara (2007) : W devrait être identifié à la lexie. Autrement dit, les nœuds du réseau lexical que construit l'apprenant, sont des lexèmes ou des locutions, et non pas des vocables ou des familles de mots.

Si pour être bien connu et solidement intégré dans le vocabulaire actif de l'apprenant, un W doit s'attacher par de nombreux liens à d'autres éléments lexicaux, W est en réalité une L : une lexie ou unité lexicale, c'est-à-dire un lexème ou une locution. C'est la maîtrise de sa structure sémantique et de ses liens lexicaux qui permettra d'introduire et d'ancrer W/L dans le stock lexical de l'apprenant.

Un regard lexicologique sur les erreurs (1)-(5) nous a ainsi permis de définir l'objet des apprentissages lexicaux. Il nous est impossible, dans le présent article, de discuter des stratégies particulières à adopter en classe de langue (cf. Schmitt, 2007). Elles varieront nécessairement selon le niveau, voire selon les attitudes des apprenants, que chaque enseignant est le plus à même d'évaluer et d'exploiter. Sachant en revanche que l'acquisition du lexique est un processus qui se poursuit tout au long d'une vie, il nous semble important de soumettre à la discussion quelques orientations majeures pour un enseignement du vocabulaire à la fois efficace et ouvert vers une démarche autonome.

Tout d'abord, il est essentiel de donner aux apprenants les moyens d'isoler une lexie, en les sensibilisant aux faits de polysémie et à leurs conséquences. Un travail intensif ayant pour objectif de faire apparaître les liens de copolysémie entre les différentes acceptions d'un vocable est le premier pas dans un apprentissage lexical. Il est indispensable dès le niveau A1 pour au moins deux raisons. Premièrement, identifier un copolysème permettra de construire autour de lui un réseau des lexies auxquelles il est lié par des liens paradigmatiques (de dérivation sémantique, de synonymie, d'antonymie, d'hyperonymie) et syntagmatiques. À court terme, une telle démarche optimise l'apprentissage, puisqu'au lieu d'une seule lexie, on en découvre toute une grappe bien organisée par le lexique lui-même (en plus de VISITE MÉDICALE, l'apprenant accède à SE PRÉSENTER, AVOIR II, EFFECTUER, PASSER III.1, SUBIR, ce qui évite l'impasse de la « vache fondamentale », cf. Galisson, 1971). À moyen et long termes, elle met en place et développe des habitudes de travail précieuses dans un contexte de L2 quelle qu'elle soit. Car – deuxièmement – les relations de sens qui structurent la polysémie (métaphore, métonymie, spécialisation, extension) sont récurrentes d'une langue à l'autre. La capacité de les repérer, voire de les anticiper et de les rechercher, permettra à l'apprenant d'optimiser l'apprentissage lexical et, à terme, de développer et de structurer son stock lexical en autonomie.

L'expérience de classe montre qu'une exploration des structures polysémiques est possible dès le niveau A1. Nous l'avons proposé aux étudiants en réponse à une question portant sur le sens de BUREAU dans la phrase *Il arrive au bureau à huit heures*. Sans présenter toute la structure polysémique du vocable, nous avons néanmoins tenu à montrer qu'il en existe plus d'un sens. Nous avons donc choisi de présenter aux étudiants deux lexies reliées par une relation de métonymie, avec une poignée de collectifs indispensables pour les utiliser :

BUREAU I = 'meuble utilisé par l'individu X pour travailler'
BUREAU II = 'local où l'individu X travaille'

Une deuxième dimension de l'apprentissage lexical nous semble tout aussi importante à souligner : si chaque lexie s'inscrit dans un réseau lexical, elle doit être présentée de manière à permettre aux apprenants d'accéder à ce réseau. Concrètement, cela signifie avant tout de prendre systématiquement en compte tous les actants associés à la lexie : l'invité(e) X et son hôte (Y) dans le cas VISITE I, la personne convoquée, le médecin ou le centre agréé et l'objectif de la démarche pour la VISITE MÉDICALE. La prise en compte des participants de situations dénotées par les lexies étudiées stimule le développement du stock lexical de l'apprenant à la fois sur le plan quantitatif (le réseau s'enrichit de nouveaux nœuds W/L) et qualitatif (les nœuds s'en trouvent plus solidement rattachés aux autres éléments du réseau). Conjugué avec un travail systématique de la polysémie que nous venons d'évoquer, cette démarche conduit à repérer les spécificités du réseau qui s'organise autour de chaque lexie, par exemple l'absence de conversion verbale et nominale pour certaines acceptions (cf. la section 3.1.2 ci-dessus).

Troisièmement, des activités systématiques de paraphrase, consistant à transmettre un même contenu sémantique dans plusieurs phrases synonymes (cf. les exemples (6) et (7) dans la section 2.1, (12) et (13) dans la section 3.1.2) conduisent l'apprenant à consolider son réseau lexical. S'entraîner à « dire la même chose autrement », c'est se donner le moyen de manipuler les lexies conformément aux règles de grammaire. Ce type d'exercices permet donc d'articuler l'apprentissage du lexique et une mise en pratique des connaissances grammaticales.

Certaines des orientations exposées ci-dessus ont déjà été discutées, notamment dans Tréville et Duquette (1996), Picoche (2007) dans des cadres théoriques différents. Nous espérons que le diagnostic des erreurs illustrées par les exemples (1) à (5) et la réflexion que nous avons menée en nous appuyant sur la métaphore des réseaux que sont d'une part le lexique d'une langue et de l'autre, le stock lexical acquis par les apprenants apporte de nouveaux arguments en faveur de la pertinence de ces propositions pédagogiques.

Ces propositions se heurtent cependant à une difficulté majeure d'ordre pratique. Un cursus de langue se déroule toujours dans les limites d'un certain volume horaire que l'enseignant s'emploie à gérer de manière à assurer un équilibre entre tous les aspects linguistiques et culturels indispensables à l'apprenant pour progresser dans la maîtrise de la langue cible. En effet, on ne saurait se passer d'informations grammaticales dans leurs différentes dimensions (morphologie, syntaxe, structuration du discours, etc.), pas plus qu'il n'est concevable de négliger une perspective interculturelle de l'apprentissage. Ajoutons à cela les incontournables exercices visant la réutilisation du matériau linguistique étudié dans des tâches communicatives et nous verrons clairement que – sans mettre en question l'importance de l'enseignement du lexique – celui-ci est soumis à des contraintes de temps. Est-il possible de proposer aux apprenants une exploration systématique du vocabulaire ne serait-ce que d'un texte de manuel, dans un cours de deux heures hebdomadaire en présentiel ? Est-il simplement réaliste de proposer un tel programme ?

Notre réponse est affirmative, même si – comme toute activité pédagogique – la démarche que nous proposons doit être constamment adaptée par l'enseignant en charge de cursus, notamment en fonction du public et des contraintes horaires. L'une des stratégies possibles consiste alors à identifier un nombre restreint d'unités lexicales dont la maîtrise est essentielle pour l'apprenant. Il peut s'agir ainsi des lexies qui appartiennent au vocabulaire de base (McCarthy, 1999) ou qui sont particulièrement importantes pour le champ sémantique que l'enseignant propose aux apprenants d'explorer. C'est sur elles que se focaliseront alors les activités destinées à guider l'apprenant à travers les différentes étapes de l'exploration lexicale : mise en relation avec un ou plusieurs copolysèmes, analyse des structures actanciennes, identification de liens paradigmatiques et syntagmatiques. Car l'enjeu essentiel pour une didactique du vocabulaire, c'est adopter une façon de penser le lexique en tant que réseau interconnecté et de mettre en place les habitudes et les démarches individuelles de chacun. Une lexie n'est jamais seule, elle est liée à d'autres unités lexicales. En repérant et en suivant ces liens, on s'équipe d'un moyen efficace pour construire son propre réseau lexical en langue seconde.

Références bibliographiques

- Aitchison, J. (2003). *Words in the Mind: An Introduction to the Mental Lexicon*. 3^e édition, Oxford UK: Blackwell.
- Bolly, C. (2011). *Phraséologie et collocations. Approche sur corpus en français L1 et L2*. Bruxelles, Bern, Berlin, Frankfurt/Main, New York, Oxford, Wien : Peter Lang.
- Galisson, R. (1971). *Inventaire thématique et syntagmatique du français fondamental*. Paris, Hachette : Larousse.
- Granger S., Vandeventer A., Hamel M.-J. (2001). Analyse des corpus d'apprenants pour l'ELAO base sur le TAL, *Traitement automatique des langues*, 42, 2, 2001, 609-621.
- Gross, M. (1981). Les bases empiriques de la notion de prédicat sémantique. *Langages*, 63, 7-53.
- Hamel, M.-J., Miličević, J. (2007). Analyse d'erreurs lexicales d'apprenants du FLS : démarche empirique pour l'élaboration d'un dictionnaire d'apprentissage. *Revue canadienne de linguistique appliquée*, 10(1), 25-45.
- Hobeika-Chakroun, F. (2013). Expressions discursives des expressions idiomatiques dans une perspective didactique. Le cas de l'arabe. Conférence aux « 1^{ères} rencontres phraséologiques : des collocations aux séquences figées », Grenoble, 13–15 novembre 2013.

- Le Bot, M.-C., Richard, E. (2011). Pour une analyse morphologique des productions écrites d'élèves FLS. *Revue Tranel (Travaux neuchâtelois de linguistique)*, 54, 159-172.
- Masseron, C., Luste-Chaa, O. (2008). Typologie d'erreurs lexicales : difficultés et enjeux. Durand J., Habert B., Laks B. (éds.) *Congrès Mondial de Linguistique Française – CMLF'08*, Paris : Institut de Linguistique Française, 519-531.
- McCarthy, M. (1999). What constitutes a basic vocabulary for spoken communication? *English Language and Literature*, 1, 233-249.
- Meara, P. (2007). Towards a new approach to modelling vocabulary acquisition. Schmitt, N., McCarthy, M. (dir.) *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge UK: Cambridge University Press, 109-121.
- Mel'čuk, I. et coll. (1984, 1988, 1992, 1999). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I-IV*, Montréal : Les Presses de l'Université de Montréal.
- Mel'čuk, I., Clas, A., Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve : Duculot.
- Mel'čuk, I., Polguère, A. (2006). Dérivations sémantiques et collocations dans le DiCo/LAF. *Langue française*, 50, 66-83.
- Mel'čuk, I., Polguère, A. (2007). *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. Bruxelles : De Boeck et Larcier.
- Miličević, J. (2007). Paraphrase as a Tool for Achieving Lexical Competence in L2. *Complexity, Accuracy and Fluency in Second Language Use, Learning and Teaching*, Actes du colloque. Brussels: Royal Belgium Academy of Sciences, 124-134.
- Milton, J. (2009). *Measuring Second Language Acquisition*. Bristol UK/Buffalo NY/Toronto: Multilingual Matters.
- Nation, I. S. P. (1982). Beginning to learn foreign language vocabulary: a review of research. *RELC Journal*, 13, 61-68.
- Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Perdue, C. (1980). L'analyse des erreurs : un bilan pratique. *Langages*, 57, 87-94.
- Picoche, J. (2007). *Enseigner le vocabulaire, la théorie et la pratique*. Champigny-sur-Marne : Éditions Allouche, publication électronique sur CD-ROM [version révisée de l'ouvrage paru en 1993 sous le titre *Didactique du vocabulaire français*, collection « fac.linguistique », Paris : Nathan].
- Polguère, A. (1992). *Remarques sur les réseaux sémantiques Sens-Texte*. A.Clas (réd.), *Le mot, les mots, les bons mots*, Montréal, PUM, 109-148.
- Polguère, A. (2011). Classification sémantique des lexies fondée sur le paraphrasage. *Cahiers de lexicologie*, 98-1, 197-211.
- Polguère, A., Sikora, D. (2013). Modèle lexicographique de croissance de vocabulaire fondé sur un processus aléatoire, mais systématique. Garcia-Debanc, C., Masseron, C., Ronveaux, Ch. (éds) *Enseigner le lexique*. Namur : Presses universitaires de Namur, 35-63.
- Porquier, R. (1977). L'analyse des erreurs. Problèmes et perspectives. *Études de Linguistique Appliquée*, 25, 23-43.
- Schmitt, N. (2007). Vocabulary learning strategies. Schmitt, N., McCarthy, M. (dir.) *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge UK: Cambridge University Press, 199-227.
- Sikora, D. (2012). Définir le sens dans un réseau lexical. *Studia Romanica Posnaniensia*, 39/3, 63-79.
- Sikora, D., Polguère, A. (2013). Comment tisser son réseau phraséologique dans l'apprentissage d'une langue seconde. Conférence aux « *1^{ères} rencontres phraséologiques : des collocations aux séquences figées* », Grenoble, 13-15 novembre 2013.
- Tréville, M.-C., Duquette, L. (1996). *Enseigner le vocabulaire en classe de langue*. Paris : Hachette.

Ressources en ligne :

Base textuelle Frantext : www.frantext.fr

Corpus français du Web (FrWac) : <http://nl.ijs.si/>

Corpus de l'Est Républicain : <http://www.cnrtl.fr/corpus/estrepubicain/>

¹ Les commentaires des relecteurs anonymes m'ont beaucoup aidée à améliorer le texte et je tiens à les en remercier.

² Pour une discussion des enjeux liés à l'étude des erreurs d'apprenants, voir entre autres Masseron et Luste-Chaa (2008).

³ Signalons au passage une autre difficulté qui apparaît lorsque l'on essaie de donner un nom à la famille : faut-il adopter le point de vue morphologique et considérer qu'il s'agit du mot VISITE, ou bien choisir le critère sémantique et admettre que c'est le mot VISITER ?

⁴ Il existe trois types de locutions dont l'exposé et une discussion sont proposés dans Mel'čuk (2013) et dans Polguère (*à paraître*). CASSER LA CROUTE est une locution forte, puisque le sens d'aucune des lexies qui la composent n'est présent dans celui de la locution. VISITE MÉDICALE est, au contraire, une locution faible : on retrouve dans son sens celui de l'une des lexies des vocables VISITE et MÉDICAL, mais d'autres composantes sémantiques contribuent à la construction de son signifiant. Une analyse plus détaillée du statut lexical de la locution VISITE MÉDICALE est proposée dans la section 3.2 ci-dessous.

⁵ La notion de prédicat nominal est largement employée également dans les approches Lexique-Grammaire (cf. Gross, 1981).

⁶ La définition d'étiquette sémantique que nous proposons est directement inspirée de celle, plus formelle, formulée par Polguère (2012 : 202).

⁷ Pour une présentation des procédures de discrimination des sens, on pourra se reporter, entre autres, à Mel'čuk et coll. (1995).

⁸ Pour une présentation de l'application de ce concept pour la construction d'une ressource lexicale à usage pédagogique, voir également Polguère et Sikora (2013).

⁹ La lexie nominale VISITE IV n'est pas le seul dérivé sémantique de VISITER II. Il y a également les noms du premier actants VISITEUR I.2 et VISITEUSE I.2, dont le sens est construit sur celui de VISITER II :

VISITEUR I.2 (X, qui est un ~ de Y) = 'individu X, qui VISITE II Y'
VISITEUSE I.2 (X, qui est une ~ de Y) = 'femme X, qui VISITE II Y'

¹⁰ Sans pouvoir discuter cette question ici, signalons néanmoins que la différence sémantique dont nous faisons se traduit par une relation syntaxique différente entre les deux actants de VISITE I et de VISITER I.2.

¹¹ Pour une discussion détaillée de la structuration des définitions lexicographique à usage pédagogique, voir Sikora (2012).

ACTIVITÉ(S) DE DÉFINITION DANS L'APPRENTISSAGE D'UNE LANGUE SECONDE

À Józef Sypnicki
– *in memoriam*

Résumé :

À quoi sert-il d'apprendre à définir dans un cours de langue étrangère ? Réputée difficile à cause de son caractère métalinguistique, et de ce fait détachée des échanges quotidiens, l'activité consistant à définir le sens lexical est peu, voire pas du tout valorisée dans ces enseignements. En s'appuyant sur le modèle des compétences linguistiques proposé par Bialystok (1991), le présent article rassemble tout d'abord des arguments en faveur des activités définitionnelles intégrées dans les apprentissages de langue étrangère dès les niveaux peu avancés. Le travail de définition est en fait à envisager dans une perspective à moyen et long terme, dans un objectif de structuration de compétences linguistiques, notamment lexicales. Dans la seconde partie, l'article présente et discute le déroulement d'une séquence pédagogique en classe de français langue étrangère proposée à un public de niveau A2.

Mots clés : définition, didactique des langues, lexicologie

Abstract :

Why should foreign language students learn to define lexical meanings ? Such a metalinguistic, decontextualized activity is on the one hand thought of as highly difficult and on the other poorly considered from a communicative angle. Against the background of Bialystok's model (1991) of linguistic knowledge, the present paper argues in favor of an early integration of definitional activities in foreign language teaching. Such a pedagogical task brings benefits in a middle and long term perspective, whereby the learner's linguistic, and specially lexical knowledge can be structured. With this in mind, the second part of the article details and discusses a series of definitional activities for A2 learners of French as foreign language.

Keywords : definition, second language teaching, lexicology

1. Introduction : questions et méthodes

Il est facile d'expliquer à quoi sert une définition, lorsque l'on apprend une langue étrangère : son utilité principale est de faire comprendre un sens jusque là inconnu. Le présent article pose une question autrement plus complexe : À quoi bon

apprendre à définir soi-même les unités lexicales d'une langue seconde (désormais L2)¹ que l'on s'efforce de maîtriser, alors que par ailleurs les ressources dictionnairiques sont légion ? La réflexion qui nous amènera à la réponse part, dans la section 2, d'un croisement de chemins didactique et lexicologique. Nous y discutons la place qu'accorde aux définitions le *Cadre européen commun de référence pour les langues*, en comparant son approche actionnelle au modèle de développement de compétences langagières proposé par Bialystok (1991). Dans la seconde partie de cette même section, nous comparons deux types de définitions lexicographiques susceptibles de servir de modèles aux étudiants dans les activités de classe présentées dans la partie suivante. En effet, la section 3 décrit et discute le déroulement d'une séquence pédagogique réalisée auprès d'un public étudiant de niveau A2 en français langue étrangère. Les activités successives ont été conçues en référence au modèle de Bialystok (1991). Elles montrent – comme nous le soulignons en concluant notre réflexion dans la section 4 – que les activités définitionnelles, réputées à juste titre difficiles, peuvent être adaptées à des niveaux de compétences relativement peu avancés. Nous espérons convaincre le lecteur qu'elles ont toute leur place dans une classe de langue, dès lors qu'un apprentissage est envisagé dans une perspective de structuration de connaissances linguistiques.

2. Au croisement des chemins

2.1 DÉFINITIONS et activités définitionnelles

Définition ? – on, le sait, il y en a plusieurs. Le vocable DÉFINITION² est polysémique et quand on aborde la question d'un point de vue lexicologique et lexicographique, il convient de distinguer au moins deux acceptions, c'est-à-dire deux copolysèmes liés, dans le cas qui nous intéresse, par une relation de métonymie. La première acception DÉFINITION 1, illustrée dans l'exemple (1) ci-dessous, dénote, sous un format nominal, le même contenu sémantique que le verbe DÉFINIR dans son sens de base : l'action de construire, dans un acte de communication, une paraphrase définitoire – DÉFINITION 2 ; cf. l'exemple (2).

(1) Le projet de loi prévoit des mesures permettant de faciliter et d'accélérer l'arrivée des médicaments génériques sur le marché. Il procède à la définition de la notion de médicaments biologiques similaires et précise celle du médicament homéopathique.

(Internet)

(2) D'après la définition de la densité, la densité de l'eau est égale à 1.

¹ En parlant de langue maternelle (langue première, L1) et de langue seconde (L2), nous nous situons dans une perspective d'acquisition, sans prendre en compte le statut qu'un parler peut avoir sur un territoire particulier. Nous faisons nôtre le point de vue de Klein (1995 : 3) : « ... a language is 'first' – and so is its acquisition – if no other language was acquired before; otherwise, it is second ».

² Nous adoptons les distinctions notionnelles et les conventions typographiques de la Lexicologie Explicative et Combinatoire : un vocable est un groupement structuré de lexies (unités lexicales, acceptions) copolysèmes. En cas de monosémie, un vocable ne compte, bien évidemment, qu'une seule lexie. Aussi bien les vocables que les lexies sont transcrits en petites capitales, mais les acceptions portent des numéros signalant leur place dans la structure polysémique du vocable.

(Internet)

En bref, les deux lexies copolysèmes qui nous intéressent ici dénotent respectivement l'action (DÉFINITION 1) dont l'objectif est d'élaborer une paraphrase définitoire, c'est-à-dire une DÉFINITION 2.

Loin d'être l'affaire exclusive des dictionnaires, cette polysémie se traduit dans la pratique pédagogique, notamment dans le contexte d'apprentissage de langue maternelle (L1), ainsi que dans les enseignements disciplinaires à l'école. Pour un didacticien cependant, une troisième notion, celle d'*activité définitionnelle*, vient compléter le tableau. L'acte de définir, c'est-à-dire DÉFINITION 1, présente de nombreuses difficultés (cf. *infra*, sous-section 2.2.2) pour les élèves, lorsqu'il s'agit de leur L1, et les problèmes rencontrés se multiplient encore en L2. C'est pour y répondre et pour structurer le travail d'élaboration de paraphrases définitoires, que les enseignants proposent en amont des activités définitionnelles : « un ensemble d'activités scolaires ou plus modestement d'opérations consistant à sélectionner un mot de façon à en observer certains aspects » (Plane, 2005 : 124)³.

Pour notre part, nous empruntons le terme d'*activité définitionnelle*, en l'adaptant au contexte de l'enseignement de L2. Nous entendons par là des activités, faites notamment d'exercices construits ou choisis pour l'occasion, dont l'objectif consiste à apprendre à définir. Notre conception d'activités définitionnelles est ainsi plus large que celle de Plane (2005) dans la mesure où elle englobe aussi bien les informations linguistiques que les apprenants doivent repérer pour construire une paraphrase, que les compétences nécessaires pour maîtriser la définition en tant qu'acte de parole. En simplifiant : les activités définitionnelles doivent préparer le terrain, en fournissant aux étudiants les balises et les outils indispensables pour maîtriser la définition dans ses deux dimensions : celle de construction de paraphrases et celle d'énoncé linguistique structuré d'une certaine façon.

2.2 Le chemin didactique : définition(s) en didactique de langue seconde

En classe de langue, toute séquence pédagogique s'organise en fonction des objectifs que le formateur considère comme légitimes. Il est essentiel dès lors de déterminer en quoi la capacité à définir le sens d'une lexie constitue un objectif pédagogique valide pour un apprenant de langue étrangère. Autrement dit, la question est de savoir à quoi sert de définir.

2.2.1 Définir pour accomplir une tâche sociale : perspective du CECRL

La publication, par le Conseil de l'Europe en 2001, du *Cadre européen commun de référence pour les langues : Apprendre, enseigner, évaluer* (désormais le CECRL) a marqué un tournant important dans la didactique des langues. La seconde partie du titre indique en effet que l'objectif des auteurs a été de saisir le processus didactique dans toutes ses dimensions, en incluant la démarche pédagogique (*enseigner*), l'appropriation et le développement des compétences

³ Voir aussi Plane et Lafourcade (2004).

(*apprendre*), ainsi que les résultats visés (*évaluer*). Les programmes de formation en langues sont aujourd'hui élaborés conformément aux prescriptions du CECRL.

L'utilité d'un cadre de référence commun est incontestable, notamment en ce qui concerne la détermination claire des niveaux de compétences qu'il établit et qui conduit vers une évaluation précise dans un système de certification transparent, lisible et indépendant, quel que soit le cadre de formation. Sans nier ses apports, il ne faut pas oublier que le CERCLE est un document adopté dans un contexte social et politique particulier : celui de la poursuite de l'intégration européenne. Au fur et à mesure que ce processus se poursuivra, les apprenants seront régulièrement amenés à vivre et à travailler avec des personnes de langues et de cultures différentes, à réaliser avec elles des projets communs. C'est ainsi que le chapitre 2 (p. 15)⁴ explique les positions de ses auteurs :

La perspective privilégiée ici [DS : dans le CECRL] est, très généralement aussi, de type actionnel en ce qu'elle considère avant tout l'utilisateur et l'apprenant d'une langue comme des acteurs sociaux ayant à accomplir des tâches (qui ne sont pas seulement langagières) dans des circonstances et un environnement donnés, à l'intérieur d'un domaine d'action particulier.

Une langue seconde est ainsi considérée essentiellement en tant qu'outil permettant au locuteur de réaliser des tâches successives pour assumer son rôle d'acteur social. Et les auteurs de préciser à la page suivante :

Est définie comme tâche toute visée actionnelle que l'acteur se représente comme devant parvenir à un résultat donné en fonction d'un problème à résoudre, d'une obligation à remplir, d'un but qu'on s'est fixé. Il peut s'agir tout aussi bien, suivant cette définition, de déplacer une armoire, d'écrire un livre, d'emporter la décision dans la négociation d'un contrat, de faire une partie de cartes, de commander un repas dans un restaurant, de traduire un texte en langue étrangère ou de préparer en groupe un journal de classe.

Si les auteurs du CECRL se gardent bien d'adopter des positions théotiques tranchées, l'approche actionnelle⁵ qu'ils préconisent attire des objections et une discussion approfondie de ses principes mériterait sans conteste toute notre attention. Faute de place, nous tenons à soulever uniquement les points directement liés aux problèmes qui nous intéressent ici. Il est tout à fait légitime par exemple de se demander si la focalisation assumée sur l'action sociale n'est pas trop réductrice dans sa façon de présenter la langue, en occurrence la L2, uniquement dans une dimension utilitaire, servant à accomplir des tâches pratiques. La définition dans les deux sens étudiés ici peut-elle trouver une place dans des approches actionnelles ? Faut-il admettre que DÉFINITION 1, l'action de définir une lexie, possède une « visée actionnelle » et peut, de ce fait, être considérée comme une tâche ? Le résultat auquel elle aboutit, c'est-à-dire une paraphrase définitoire (DÉFINITION 2), est-il à

⁴ La version sur laquelle nous nous appuyons est disponible en ligne dans un format .pdf. Les numéros des pages que nous citons renvoient à ce fichier électronique (cf. Bibliographie pour la référence complète).

⁵ Pour une discussion approfondie de la perspective actionnelle voir Perrichon (2008). Puren (2014) la présente par ailleurs dans la continuité et dans ses ruptures par rapport aux approches communicatives.

envisager « en fonction d'un problème à résoudre, d'une obligation à remplir, d'un but qu'on s'est fixé » ? Quelle est la place des activités définitionnelles dans l'enseignement de L2 et avec quel niveau de compétences faut-il les associer ?

Le chapitre 4 du CERCL, intitulé *L'utilisation de la langue et l'apprenant utilisateur*, prévoit qu'un étudiant qui atteint un niveau de compétences B1 (le niveau seuil⁶) « [p]eut définir les caractéristiques de quelque chose de concret dont le nom lui échappe » et « [p]eut exprimer le sens d'un mot en en donnant un autre, signifiant quelque chose de semblable (par exemple, 'un camion pour voyageurs' pour 'bus') ». Autrement dit, il s'agit de paraphraser une lexie inconnue ou oubliée pour se faire comprendre dans une situation d'énonciation particulière. On se demande d'ailleurs si « en définissant les caractéristiques de quelque chose de concret dont le nom lui échappe », l'apprenant se lance réellement dans un acte de définition ou bien s'il entreprend simplement de décrire l'objet en question.

L'exemple (3) ci-dessous permet de répondre par l'affirmative aux trois premières questions. Oui, il s'agit bien d'une tâche à visée actionnelle : le besoin de définir une lexie se manifeste régulièrement bien avant le niveau seuil, comme en témoigne cet échange avec une apprenante coréenne de niveau A2, connu également comme celui de survie des utilisateurs élémentaires.

(3)

- Qu'est-ce que vous avez fait le week-end dernier ?
- J'ai surfé sur Internet pour trouver des informations.
- Des informations sur quoi ?
- Sur l'*allongement.
- Pardon ?
- Le lieu où on reste pendant le voyage.

26 novembre 2014

La confusion entre *allongement* employé en (3) et *hébergement* visé par l'étudiante est caractéristique des apprenants peu avancés dont les connaissances lexicales restent limitées et ne sont pas encore structurées en un réseau valide. Il s'agit d'un clang : association de deux lexies sans lien paradigmatique ni syntagmatique, dont la connexion repose sur un rapprochement phonétique. En effet, les deux noms ont un début vocalique ([a] *versus* [e]), ainsi qu'une consonance [-ʒem-] et une assonance [-ã] dans les deux dernières syllabes. Construire une paraphrase définitoire, qui a permis à l'acte de communication d'aboutir, est bel et bien une tâche à visée actionnelle.

Nous tenons à montrer cependant que la visée de l'acte de définir et l'utilité des paraphrases dépassent largement le cadre actionnel, purement social, fixé par le CECRL, en permettant à l'apprenant d'adopter un regard analytique sur ce qui est

⁶ Rappelons le descriptif général des compétences en B1 : l'apprenant de ce niveau « [p]eut comprendre les points essentiels quand un langage clair et standard est utilisé et s'il s'agit de choses familières dans le travail, à l'école, dans les loisirs, etc. Peut se débrouiller dans la plupart des situations rencontrées en voyage dans une région où la langue cible est parlée. Peut produire un discours simple et cohérent sur des sujets familiers et dans ses domaines d'intérêt. Peut raconter un événement, une expérience ou un rêve, décrire un espoir ou un but et exposer brièvement des raisons ou explications pour un projet ou une idée. » (CECRL, chapitre 3, p. 23)

l'objet premier de ses apprentissages, c'est-à-dire la langue, en l'occurrence le français, et de progresser de manière de plus en plus autonome dans son acquisition.

2.2.2 Définir en L1 et en L2 : bref répertoire des difficultés

Dans la pratique de classe, on constate régulièrement les difficultés auxquelles se heurtent les élèves, lorsqu'ils sont amenés à construire une définition dans leur langue maternelle. Plusieurs études expérimentales (Bialystok, 1991, Snow et al., 1991, Pajooheh, 2014) confirment ces observations de terrain. Pour un apprenant de L2, la définition représente un défi encore plus complexe, puisqu'à la liste des problèmes connus en L1 s'ajoutent les moyens linguistiques souvent trop modestes pour gérer cette tâche.

Tout d'abord, il n'est pas sûr que, ayant suivi des enseignements disciplinaires dans sa langue et appris, par exemple, la définition de REELLE ZAHL pour un germanophone (ou selon la langue, REAL NUMBER pour un anglophone, LICZBY RZECZYWISTEJ dans le cas d'un polonophone, etc.), ce même élève soit en mesure de paraphraser le sens des lexies telles que AUTO, KÜHLSCHRANK, FREUDE (respectivement, CAR, FRIDGE, JOY; SAMOCHÓD, LODÓWKA, RADOŚĆ, etc.) qu'il utilise par ailleurs couramment dans des échanges quotidiens en langue maternelle.

Il s'agit en effet de deux types de lexies, au fonctionnement linguistique différent, qu'il faut bien se garder de confondre. REELLE ZAHL et ses équivalents dans les autres langues correspondent à ce que Vygotsky (1986) qualifie de concepts scientifiques, auxquels on accède par une formation systématique et qui se détachent de l'univers du discours du locuteur. Ils font objet d'un usage conscient, rendu possible par une analyse de leur contenu et par un apprentissage. Les autres lexies que nous venons de citer en exemples correspondent, au contraire, à des *concepts spontanés* (Vygotsky, 1986), qui se développent par l'expérience et par la pratique langagière à l'intérieur d'une communauté linguistique, sans qu'une phase de conscientisation soit nécessaire.

Deuxièmement, la définition des lexies non spécialisées est une activité métalinguistique⁷ et, de ce fait, décontextualisée. Il s'agit, pour le locuteur, d'aller au-delà des emplois particuliers, qui se réalisent dans des situations de discours précises, pour saisir ce qui leur est commun indépendamment de celles-ci. Pour ce faire, il doit disposer des connaissances très développées et bien structurées de sa langue maternelle, notamment en ce qui concerne les relations lexicales. Ce qui implique une réflexion et une analyse linguistiques qui ont pour objet la L1 en tant que système, sinon dans son ensemble, du moins dans certaines de ses dimensions.

⁷ Rappelons que, sans être nommée, la définition est présentée par Jakobson (1963 : 217-218) comme une réalisation par excellence de la fonction métalinguistique de la langue : « Qu'on imagine un dialogue aussi exaspérant que celui-ci : "Le sophomore s'est fait coller." "Mais qu'est-ce que *se faire coller* ?" "Se faire coller veut dire la même chose que *sécher*." "Et *sécher* ?" "Sécher, c'est *échouer à un examen*." "Et qu'est-ce qu'un *sophomore* ?" insiste l'interrogateur ignorant du vocabulaire étudiantin. "Un *sophomore* est (ou signifie) un étudiant de seconde année." L'information que fournissent toutes ces phrases équationnelles porte uniquement sur le code lexical du français : leur fonction est strictement métalinguistique. »

Pour un apprenant de L2, l'acte de définir exige la mise en œuvre des connaissances lexicales et grammaticales⁸ dans une langue qui est en cours d'acquisition. On pourrait ainsi s'attendre à ce que des apprenants B2, qui de surcroît ont souvent un bagage scolaire, voire universitaire dans leur langue maternelle, produisent plus facilement des paraphrases définitives en L2 que ceux de niveau B1, puisqu'ils disposent de connaissances linguistiques plus vastes. Or, il n'en est rien, comme le remarque entre autres Pajooohesh (2014), et cela pour plusieurs raisons souvent cumulées. Les apprenants peuvent très bien intégrer des lexies nouvellement apprises dans leur discours, sans pour autant être capables d'expliquer leur sens dans une paraphrase définitive. En L2, comme en L1, une bonne maîtrise de relations lexicales, de structures grammaticales appropriées d'une part, et de l'autre, l'appropriation d'une forme de discours normée sont nécessaires. On comprend donc aisément pourquoi la compétence définitionnelle que l'apprenant peut avoir dans sa langue maternelle ne se transpose pas automatiquement sur celle qu'il développe en L2.

2.2.3 Définir pour construire ses compétences : modèle de Bialystok (1991)

[L]anguage proficiency is not simply a skill or procedure which is mastered independently of other forms of knowledge and other types of abilities. Language is also a logical symbolic system, capable of being known itself and capable of guiding and shaping other aspects of cognition.

(Bialystok, 1991 : 113)

L'acte de définir est souvent envisagé au-delà de sa dimension linguistique pour impliquer la cognition et le système conceptuel sous-jacent. En effet, comme l'indique Pajooohesh (2014 : 80), les tâches métalinguistiques telles que la définition ne peuvent se faire sans recours à un système conceptuel, qui s'organise pourtant en relation avec des formes et des structures linguistiques. Il nous est impossible, dans le cadre du présent article, de discuter des relations entre la langue et la cognition. Il est en revanche indispensable, pour situer notre propos, d'expliquer l'optique dans laquelle nous proposons d'exploiter le modèle de Bialystok pour introduire des activités définitionnelles dans une classe de L2.

La dimension métalinguistique de l'acte de définir se traduit à la fois par l'orientation de cette tâche vers la langue elle-même, par les compétences de l'apprenant qu'elle implique et par le caractère conscient de sa démarche. Bialystok propose un modèle qui permet de comprendre le fonctionnement des mécanismes définitionnels et l'importance de cette activité pour les apprenants⁹.

Selon Bialystok (1991), toute activité langagière, c'est-à-dire toute production de la parole, repose sur deux mécanismes qui doivent fonctionner en amont de l'acte de communication.

⁸ Nous reviendrons sur la structure linguistique des énoncés définitives dans la section suivante.

⁹ Si le modèle a été élaboré à partir de l'observation d'enfants bilingues, il peut être étendu, comme l'auteur l'indique, aux apprenants de L2. Certaines thèses de Bialystok nous paraissent néanmoins trop déterministes : « Representations of language in a broad sense [i.e. knowledge and its structure – (...)] are organized around meanings and uses. In this way, language is a set of semantic relations that organize our knowledge of the world. » Sans nier le lien entre la langue et la structure des connaissances, le déterminisme strict de cette thèse nous semble aller à l'encontre de l'expérience quotidienne de la langue.

1. Avant d'être en mesure de se lancer dans un acte de communication en L2, l'apprenant doit avoir effectué une analyse des données linguistiques (*analysis of linguistic knowledge*, Bialystok 1991 : 116) qui lui sont parvenues auparavant soit dans un cours, soit dans une situation d'échanges quotidiens. En effet, les informations dont il dispose nécessitent un traitement de la part de l'apprenant pour former une représentation valide et opérationnelle du code qu'il cherche à maîtriser. Sur le plan lexical, ce mécanisme d'analyse de données linguistiques, indispensable pour la structuration des connaissances, conduit entre autres à établir des connexions à la fois hiérarchiques comme celle d'hyponymie entre par exemple ANIMAL et CHIEN, ANIMAL et CHAT, et non-hiérarchiques de co-hyponymie qui se tissent entre CHIEN et CHAT. Notons que si le modèle de Bialystok est construit dans une perspective sémantique, la compétence d'analyse s'étend également sur les aspects formels et structurels de la langue, notamment sur les règles de grammaire.

Comme le souligne Bialystok (1991 : 118), c'est grâce à l'analyse que l'ensemble des représentations de la langue, *ipse eo* l'ensemble de connaissances linguistiques, se construit et évolue par des restructurations constantes : les différents aspects du système linguistique sont traités et re-traités à différentes étapes d'apprentissage. Les capacités d'analyse vont en s'améliorant avec l'âge, l'expérience et la formation reçue, notamment lorsque celle-ci fournit à l'apprenant des méthodes et des outils adaptés.

Une tâche métalinguistique se distingue ainsi des autres activités langagières par deux caractéristiques indissociablement liées à la compétence d'analyse. D'une part, elle nécessite le recours à des connaissances linguistiques structurées, et de l'autre, elle oblige le locuteur, en l'occurrence un apprenant de L2, à prendre le *definiendum*, c'est-à-dire une lexie, dans sa dimension de signe linguistique détaché de son contexte immédiat. Il s'agit ainsi de la saisir telle qu'elle se présente en langue, et non plus seulement en discours.

2. Le second mécanisme en jeu, c'est le contrôle de la production langagière qui conduit l'apprenant à sélectionner, d'une manière consciente ou non, les éléments nécessaires pour accomplir une tâche. Ce processus requiert donc d'une part des stratégies d'attention, permettant de se diriger vers les connaissances linguistiques dont on a besoin : « [L]earners obtain higher level of control when they have mastered the appropriate attentional strategies for more difficult problems » Bialystok (1991 : 120), et de l'autre une structure des connaissances linguistiques valide et opérationnelle en ce qu'elle facilite la sélection.

Il est important d'insister, comme le fait Bialystok, sur deux caractéristiques de ces processus. Premièrement, aussi bien l'analyse des connaissances langagières que le contrôle de la production se développent progressivement chez l'apprenant tout au long de son apprentissage. Ils doivent de ce fait se mesurer par rapport à un continuum plutôt que de manière binaire acquis/non acquis.

Deuxièmement, les différentes tâches langagières requièrent un degré plus ou moins élevé à la fois d'analyse et de contrôle. Ainsi, les échanges quotidiens à l'oral n'impliquent que dans une très faible mesure l'analyse de connaissances et le contrôle de leur traitement, à l'opposé de toute activité à caractère métalinguistique, telle que la définition, qui exige un degré très élevé de maîtrise des deux processus.

2.2.4 Une perspective didactique dans un objectif à long terme

Au-delà des besoins purement communicatifs illustrés plus haut (exemple 3 dans la sous-section 2.2.1), le travail d'élaboration de paraphrase définitoire des lexies courantes – celles qui correspondent aux concepts spontanés (Vygotsky, 1986) – conduit l'apprenant à développer des compétences et à maîtriser des mécanismes permettant d'aller progressivement vers des tâches de plus en plus complexes. La définition apparaît ainsi en tant que démarche stimulant le processus d'analyse, puisque pour le définir, il faut situer le sens étudié par rapport à d'autres, auxquels il est lié par différents liens paradigmatiques. L'étudiant est ainsi amené à traiter tout un réseau de lexies qu'il ne pourra maîtriser qu'en le structurant.

Parallèlement, les stratégies attentionnelles activées par le contrôle lors de la production doivent conduire à sélectionner des connaissances linguistiques nécessaires.

Le modèle de Bialystok permet ainsi de mesurer la complexité de la tâche que représente l'élaboration de définitions pour les apprenants de L2. Mais cette difficulté même, pourvu qu'elle soit accompagnée d'une pratique pédagogique appropriée, est à envisager comme facteur favorable pour l'apprentissage. En effet, comme le remarque Bogaards (1994), plus une tâche est significative, c'est-à-dire plus elle implique l'apprenant, plus elle est efficace pour les apprentissages. Cuq (2004) souligne que les activités qui requièrent de la part de l'apprenant une conscientisation et un raisonnement, *ipse eo* un haut degré d'implication personnelle, laissent une empreinte mémorielle durable. Elles « visent à constituer une trace structurée et par conséquent plus ancrée dans la mémoire à long terme. L'association du contexte ou des réseaux (...) est l'élément fondamental qui permet la récupération de cette trace » (Cuq, 2004 : 68).

Il semble ainsi fort utile, dans un cursus de L2, de planifier régulièrement des exercices destinés à stimuler à la fois l'analyse et le contrôle de la production langagière. Les activités définitionnelles en font partie et nos expériences en classe de français langue étrangère avec un public de jeunes adultes montrent qu'il est possible de les pratiquer même à des niveaux peu avancés, par exemple en A1/A2 (cf. section 3 ci-dessous). Certes, les connaissances des apprenants restent à ce stade réduites et fragiles, mais elles n'excluent pas des activités destinées à les structurer.

2.3 Le chemin lexicologique et lexicographique : définitions et leurs micro-structures

Dans la section 2.2.3, nous avons indiqué qu'en tant qu'activité métalinguistique, la définition est une façon de traiter le signe linguistique. Définir, c'est – en suivant Jakobson (1963 : 218) – construire une relation équationnelle entre un signifiant et un signifié : « Un *sophomore* est (ou signifie) un étudiant de seconde année ». Riegel (1987, 1990) montre cependant que le choix du prédicat verbal – *être* ou *signifier* dans l'exemple de Jakobson – n'est pas anodin, puisqu'il détermine la nature de la relation qui s'établit entre le *definiendum* et le *definiens*. En effet, le verbe *être* sert à construire des énoncés définitoires copulatifs qui répondent à la question *Qu'est-ce qu'un sophomore ?* et ses variantes (*Un sophomore, c'est quoi ?*, etc.) et à asserter une identité référentielle (cf. Riegel, 1987 : 31). *Signifier* permet de répondre aux questions clairement métalinguistiques

Que signifie sophomore ? Quel est le sens de sophomore ?, en construisant une définition qui « assigne un signifié à un autonyme » (*ibidem*). Cette différence est liée au problème qui se pose invariablement depuis l'Antiquité : celui de savoir si l'on définit des choses, des catégories conceptuelles ou bien des formes lexicalement codées. Dans le premier cas, il s'agirait de décrire les membres d'une catégorie et dans le second le signifié du *definiendum*¹⁰.

L'expérience montre cependant que cette différence se trouve le plus souvent gommée dans le discours de classe. En effet, quelle que soit la question posée, l'objectif des apprenants est d'accéder à un sens inconnu, peu importe qu'on y parvienne *via* le référent du mot ou sa définition. Il se peut également que la démarche s'organise en sens inverse, celui qu'illustre l'exemple (3) ci-dessus, lorsque le recours à la paraphrase définitoire a pour visée de retrouver la forme lexicale appropriée, c'est-à-dire le signifiant dans la langue cible correspondant au *definiendum*. Dans les sous-sections qui suivent, nous discutons les différentes façons de traiter la relation entre le signifiant et le signifié pour repérer celle qui permet non seulement d'accéder au sens lexical, mais qui s'avère la plus adaptée pour structurer les connaissances linguistiques.

2.3.1 Définir pour accéder au sens

Lorsqu'il s'agit de moyens permettant aux apprenants d'accéder au sens des lexies jusque là inconnues, l'enseignant dispose de tout un arsenal de procédés. Il est possible de fournir un équivalent en langue maternelle. Si cette stratégie contribue à résoudre rapidement un problème ponctuel, son usage se limite à des classes uniglottes : il est en effet difficile d'envisager une telle démarche lorsque les apprenants ne partagent pas une même L1¹¹. Des objets concrets tels un meuble, un animal ou une plante peuvent être montrés par un geste ou sur des images projetées au gré des besoins. Les apprenants accèdent alors au sens dénoté de manière immédiate, avec peu d'effort, ne s'impliquant que dans un faible degré dans la démarche, puisque la réponse à la question de savoir ce que signifie la lexie *L* leur est fournie toute prête.

Comparativement, mimer une scène destinée à illustrer par exemple le sens du verbe *aimer* ou bien dessiner un cœur engage l'apprenant nettement plus, puisque sa tâche consiste alors à décoder un ensemble de signes, ce qui est une démarche consciente. De même, l'activation d'un réseau de synonymes, plus souvent de quasi-synonymes ou encore d'antonymes (*détester* veut dire *ne pas aimer*) suppose une réflexion et, par conséquent, un ancrage mémoriel plus solide, puisque les lexies

¹⁰ On notera, à la suite de Riegel (1987) que cette différence se manifeste par la présence d'un déterminant à valeur générique en cas d'identité référentielle construite avec *être*, impossible lorsque l'on recourt à *signifier* : *Un sophomore est un étudiant de deuxième année* versus **Un sophomore signifie un étudiant de deuxième année*.

¹¹ Il va sans dire que le recours à une langue supposée commune à l'ensemble des apprenants (par exemple, à l'anglais) présente des risques élevés de confusion. Si, en dehors de certains noms concrets, il est souvent difficile, voire impossible, de miser sur une équivalence sémantique entre deux lexies *L*₁ et *L*₂ appartenant aux systèmes lexicaux de deux langues, un troisième système de référence conduit à des approximations de compréhension clairement préjudiciables aux résultats des apprentissages.

nouvellement découvertes s'insèrent alors dans un réseau de relations paradigmatiques.

Accéder au sens lexical *via* une définition dans une ressource dictionnaire, puis formuler soi-même des paraphrases définitives engage l'apprenant dans une activité plus exigeante encore. Tout d'abord, elle oblige à décontextualiser la lexie définie, c'est-à-dire à l'envisager en dehors d'emplois situationnels. Il en résulte une certaine mise à distance de l'entité linguistique traitée. Dans un second temps, la définition conduit à construire et à exploiter un réseau de relations lexicales structurant les acquisitions. Les recherches que nous avons brièvement présentées dans la section précédente indiquent que ce travail exige une forte implication de la part de l'apprenant, puisque l'analyse des données linguistiques et le contrôle de production portent sur des éléments d'une grande diversité. L'objectif principal de toute activité définitionnelle que l'on propose aux apprenants est donc de les orienter dans leur démarche, en les amenant progressivement vers les éléments linguistiques indispensables pour réaliser cette tâche.

Comme le souligne Vigner (2012), un travail d'écriture est avant tout un travail de réécriture : pour produire un texte, l'apprenant doit d'abord explorer et comprendre les règles qui permettent de construire le discours d'un certain type, tout particulièrement lorsqu'il s'agit de séquences hautement normalisées telles que les définitions. Ajoutons que les observations de Vigner s'appliquent également à la production orale qui obéit, elle aussi, à des contraintes structurelles, du moins en ce qui concerne la construction des définitions.

C'est la raison pour laquelle, dans la section suivante, nous observerons deux types de définitions que l'apprenant est susceptible de trouver dans des ressources mises à sa disposition. En 2.3.2, nous nous intéressons aux paraphrases classiques héritées des philosophes grecs antiques. La sous-section 2.3.3 est consacrée aux définitions dites populaires (*folk definitions*, cf. Stock, 1988, Lew, 2010) ou phrastiques, car elles prennent la forme de phrases qui, dans certains dictionnaires, fournissent une explication plutôt qu'une définition à proprement parler.

2.3.2 Structure de définition analytique

Dans une définition classique, le *definiens* est construit autour d'une formule précisant le genre prochain, qui s'accompagne de différences spécifiques propres au *definiendum*. Cette structure peut néanmoins se décliner de manière plus ou moins formelle (cf. Frassi, ce volume). La figure 1 reprend la paraphrase du sens premier de CIGARETTE proposée dans la version électronique du *Petit Robert 2011* (désormais, *PR*). Dans la figure (2), nous présentons la partie définitionnelle de la description lexicographique de cette même lexie dans la base de données lexicales *Réseau Lexical du Français (RL-fr)*¹².

1. Petit rouleau de tabac haché et enveloppé dans un papier fin.

Figure 1 : Paraphrase définitive de la lexie CIGARETTE 1 dans *Le Petit Robert 2011*.

¹² Pour une présentation de la base *Réseau Lexical du Français* et de son évolution, le lecteur pourra se reporter par exemple à Lux-Pogodalla et Polguère (2011) et Polguère (2014).

objet 1.1a destiné à être fumé III.a par l'individu X
 de forme allongée
 constitué d'un petit tube de papier très fin rempli de tabac¹ 1.2 haché, que X fait se consumer par une de ses extrémités
 que X tient entre ses lèvres 1.1 ou entre ses doigts 1.1a quand X fume

Figure 2 : Définition de CIGARETTE I dans la base de données lexicales *RL-fr*.

Au-delà des différences formelles¹³ repérables dans la présentation, les deux paraphrases définitoires sont organisées autour d'un pivot qu'est l'expression du genre prochain. Nous emprunterons désormais à la *Lexicologie Explicative et Combinatoire* le terme de *composante centrale* (désormais CC) pour la désigner. Celle-ci s'accompagne de *composantes périphériques* (CP par la suite) indiquant les différences spécifiques.

Les choix des lexicographes sont par ailleurs intéressants à noter. Le PR limite la paraphrase à la description de l'aspect physique d'une cigarette. La CC donne la forme de l'objet. Il s'agit d'une information certes utile, mais sur le plan lexical il semble difficile d'admettre que ROULEAU renvoie réellement le genre prochain de CIGARETTE 1, dont il suffirait de préciser, dans les CP, la taille (*petit*) et les éléments constitutifs (le tabac et le papier fin qui l'enveloppe).

Le *definiens* du *RL-fr* s'organise autour d'une CC, 'objet destiné à être fumé par l'individu X', qui classe CIGARETTE I parmi les lexies dénotant des artefacts. Or, par leur nature, les artefacts sont destinés à l'usage d'un certain type. Par conséquent, un utilisateur est impliqué. D'entrée de jeu, la définition du *RL-fr* rend compte du caractère prédicatif de cette lexie nominale¹⁴. Les CP spécifient ensuite la forme (*de forme allongée*), les éléments constitutifs de l'objet (*constitué d'un petit tube de papier très fin rempli de tabac haché*), le mode de fonctionnement (*que X fait se consumer par une de ses extrémités*) et l'usage qui en est fait (*que X tient entre ses lèvres ou entre ses doigts quand X fume*).

Les définitions construites conformément à un même principe peuvent ainsi révéler d'importantes différences de contenus. Nous avons tenu à le signaler à travers cette comparaison rapide, car il s'agira, dans la pratique pédagogique, de proposer aux apprenants un modèle qu'ils pourront imiter dans leurs productions. Le choix d'une CC n'est en effet pas anodin dans la mesure où la lexie exprimant le genre prochain est une première façon d'inscrire le *definiendum* dans un réseau de relations à la fois conceptuelles et lexicales, qui déterminera par la suite sa place dans le vocabulaire de l'apprenant. Les activités définitionnelles que nous présentons dans la section 3 illustreront notre propos.

Il nous semble intéressant cependant de nous arrêter, ci-dessous, sur les éventuelles applications pédagogiques des définitions phrastiques populaires, qui rompent radicalement avec la tradition aristotélicienne.

¹³ Différences formelles que nous ne faisons que signaler, faute de place nécessaire pour les discuter d'une manière tant soit peu exhaustive. Pour une présentation des définitions élaborées dans le *RL-fr*, voir à titre d'exemple Sikora (2012), et surtout Mel'čuk et Polguère (ce volume).

¹⁴ L'existence de prédicats nominaux, qui comme DÉFINITION I, renvoient à des faits est aujourd'hui largement admise. Lorsqu'il s'agit de nom, qui – comme CIGARETTE I – dénote non pas un fait, mais une entité, tout en sélectionnant des actants (un fumeur, le cas échéant), Mel'čuk et Polguère (2008) parlent de quasi-prédicat (voir également Mel'čuk et Polguère, ce volume, section 2.1.1). Pour notre réflexion, nous retenons simplement le caractère prédicatif de ces lexies nominales.

2.3.3 De l'utilité des définitions phrastiques

Plusieurs dictionnaires pédagogiques de langue anglaise tels que COBUILD¹⁵ ou Cambridge Dictionaries Online (CALD)¹⁶, abandonnent, du moins pour certaines acceptions, les définitions traditionnelles, en optant pour un format inspiré des conversations quotidiennes (Stock, 1988). De tels énoncés, qui se présentent sous forme de phrases complexes à subordonnée conditionnelle ou temporelle¹⁷, ressemblent à des explications que l'on peut obtenir de la part d'un locuteur natif, lorsqu'on le questionne sur le sens d'une lexie. Les exemples (4), (5) et (6) offrent une illustration de ces définitions dites populaires (*folk definitions*), dont la forme phrastique varie souvent quant au degré de sa complexité.

(4)

TICK AWAY/BY

If time ticks away/by, it goes past:

With the final seconds ticking away, Milan scored a goal.

[CALD]

(5)

LOVABLE

If you describe someone as lovable, you mean that they have attractive qualities, and are easy to like.

■ EG: “⇒ His vulnerability makes him even more lovable.”

[COBUILD]

(6)

REVIVAL

When there is a revival of something, it becomes active or popular again.

■ EG: [+ of] “⇒ This return to realism has produced a revival of interest in a number of artists.”

[COBUILD]

D'un point de vue didactique, les définitions phrastiques seraient susceptibles de présenter de nombreux avantages par rapport à l'organisation traditionnelle. Tout d'abord, elles sont conçues pour traduire une perception naïve et spontanée du sens lexical qu'ont les locuteurs d'une langue, capables de l'expliquer, en s'appuyant uniquement sur leur intuition. On peut donc admettre qu'elles sont culturellement marquées et de ce fait, apportent une dimension que n'ont pas les définitions issues d'une analyse décompositionnelle.

Deuxièmement, contrairement aux paraphrases classiques, les définitions phrastiques palient les problèmes liés à la décontextualisation de la lexie définie, puisque celle-ci apparaît dans la séquence définitionnelle non pas en mention, mais en usage, le plus souvent dans une subordonnée hypothétique ou temporelle. La lexie à définir figure dans un environnement syntaxique qui lui est propre. Un

¹⁵ <http://www.collinsdictionary.com/dictionary/english-cobuild-learners>

¹⁶ <http://dictionary.cambridge.org/dictionary/english/>

¹⁷ Cette forme syntaxique permet de faire la différence entre les définitions phrastiques et les phrases équationnelles (cf. ci-dessus, note 6).

étudiant attentif pourrait ainsi y repérer, en cas de lexie prédicative, les expressions qui dénotent ses actants, *ipse eo* identifier sa structure actancielle.

Une définition populaire serait donc à même de fournir à l'apprenant une aide substantielle tant pour ce qui est du sens lexical qu'en ce qui concerne certaines caractéristiques combinatoires. Selon Hanks (1987), le lexicographe fournit ainsi aux utilisateurs, en l'occurrence aux apprenants de l'anglais L2, non seulement des informations sur le sens de la lexie, mais également celles dont on a besoin pour maîtriser son usage.

Un certain nombre de réticences sont cependant à noter (cf. Cumming et al., 1994, Nesi, 1998, voir aussi Milićević, ce volume). Notons pour notre part que les informations lexicales et grammaticales présentes dans une définition populaire ne sont pas fournies directement. En réalité, une telle paraphrase définitoire constitue une source de données brutes à partir desquelles l'apprenant doit formuler des hypothèses tant sur le sens de la lexie qu'il cherche à comprendre que sur sa combinatoire. Le décodage d'une définition populaire constituerait ainsi une tâche exigeant une forte implication de la part de celui qui l'effectue. Une fois accomplie, elle laisserait une trace mémorielle nette et durable.

Les observations menées en classe de langue par Lew et Dziemianko (2006) mettent sérieusement en question le bien-fondé de ces postulats. Ainsi, le caractère culturellement marqué des définitions populaires, loin de faciliter l'accès au sens lexical, a plutôt tendance à égarer l'apprenant : en plus de saisir le sens, il est en effet amené à traiter les aspects situationnels et culturels évoqués dans la proposition principale. Prenons le cas de TICK AWAY/BY. La définition phrastique rapportée en (4) conduit à établir une équivalence entre ce verbe et la collocation *go past*, mais elle ne permet pas de faire le lien avec le bruit produit par les aiguilles d'une horloge mesurant le temps qui passe. Or, cette composante sémantique semble évidente pour un locuteur natif susceptible de fournir une explication telle que (4). Elle ne l'est pas nécessairement pour un apprenant qui pourtant en aura besoin pour inférer la nature sémantique des lexies (TIME, MINUTE, etc.) dénotant le premier actant de TICK AWAY/BY.

Sur le plan grammatical, comme pour la combinatoire, les bénéfiques apparaissent moindres que l'on ne s'y serait attendu. L'expérimentation menée par Lew et Dziemianko (2006) auprès des étudiants polonophones a eu pour objectif de vérifier dans quelle mesure la microstructure des définitions phrastiques permet aux apprenants d'identifier la partie du discours de la lexie définie. Or, il s'avère que le taux de reconnaissance correcte de cette donnée première pour tout un ensemble de propriétés grammaticales est de 33,2 %, soit deux fois moindre que celui enregistré pour les définitions analytiques, qui s'élève à 66,7 %.¹⁸ Des recherches supplémentaires permettraient certainement expliquer les raisons de cette différence frappante. La question qui émerge, et qui nous semble essentielle à étudier, est celle de savoir dans quelle mesure l'apprenant accède réellement à un sens inconnu sans reconnaître la partie du discours de la lexie étudiée.

Le traitement lexicographique de LOVABLE en (5) illustre la complexité de la tâche que l'apprenant doit accomplir pour comprendre le sens de cette lexie. Elle

¹⁸ De plus, force est de constater que l'information syntaxique qu'apportent les exemples (4) à (6) est plutôt modeste, car elle se limite à un seul cotexte.

apparaît en effet non pas dans un cotexte habituel, fait de collocatifs les plus fréquents, mais employée de manière métalexicalement dans une subordonnée hypothétique qui fait référence à l'activité de description (cf. exemple 5 ci-dessus : *If you describe someone as lovable, ...*). Non seulement le *definiendum* n'apparaît pas en usage, mais en plus la subordonnée fait référence à un acte virtuel, et de ce fait hautement abstrait. La proposition principale véhicule, elle aussi, un nombre important de problèmes à traiter. Elle repose notamment sur les supposées intentions du locuteur interpellé directement (*you mean*). Pour terminer ce survol rapide, signalons que l'actant sémantique de LOVABLE exprimé par *someone* dans la subordonnée est anaphoriquement repris par *they* dans la principale. Ce jeu du singulier et du pluriel ne facilite aucunement le travail d'analyse que l'apprenant doit accomplir pour accéder au sens de la lexie.

Enfin, si des activités dans lesquelles l'apprenant s'implique laissent effectivement une trace mémorielle plus marquée, on sait également que celles qui représentent un coût cognitif trop élevé sont vouées à l'échec. La pratique de classe repose sur un dosage fin de difficulté(s) que l'apprenant pourra traiter sans être découragé. Or les définitions populaires que nous avons citées ci-dessus, ainsi que les tests menés par Lew et Dziemianko (2006) autorisent à se demander si elles peuvent être exploitées par un public d'utilisateurs non autonomes (c'est-à-dire de ceux qui n'ont pas encore atteint un niveau de compétences C1 ou C2). Il semblerait que, contrairement aux attentes, les définitions populaires forcent l'apprenant à émettre de trop nombreuses hypothèses de nature diverse – lexicales, grammaticales, pragmatiques et culturelles à la fois – ce qui représente un effort et un coût cognitif trop élevé pour être efficace.

2.4 Bilan d'étape :

Classiques aristotéliennes ou populaires ? Nos réflexions et observations des sections précédentes vont dans le même sens que celles de Lew et Dziemianko (2006) : d'un point de vue didactique, les définitions analytiques semblent présenter des avantages sur les formulations dites populaires.

Dans une ressource telle que le *RL-fr*, la structure des définitions (cf. la figure 2 dans la sous-section 2.3.2) repose sur des liens paradigmatiques. Le recours au genre prochain implique en effet soit l'emploi d'un synonyme au sens plus large que le *definiendum*, soit – notamment dans le cas des lexies nominales – d'un hyperonyme. Or, comme le montrent plusieurs recherches sur la structure du lexique mental (Aitchison, 2004, Meara, 2009), les relations de ce type sont décisives pour l'acquisition du vocabulaire, car elles structurent le réseau que forme le stock lexical du locuteur.

Le choix du modèle qui servira de base aux activités définitionnelles proposées aux étudiants a des conséquences importantes. Comme le souligne Vigner (2012) pour la production écrite, lors d'un apprentissage, tout travail d'écriture consiste d'abord dans une réécriture, c'est-à-dire dans l'imitation d'un document-type. C'est le cas d'une carte postale ou d'un courriel amical que l'on demande de rédiger aux apprenants A1 et d'une lettre de motivation en B2. Que l'on propose des activités écrites ou orales autour de la définition, l'imitation d'un exemple bien choisi est une étape incontournable.

Il reste enfin la question de savoir pourquoi apprendre à définir des lexies que l'on est par ailleurs souvent capable d'utiliser dans les échanges quotidiens. Nous avons vu, avec l'exemple 3, qu'il existe des situations où définir devient une tâche telle que le CERCL l'entend et que l'apprenant doit réaliser en tant qu'acteur social pour répondre à un besoin communicatif.

Cependant, si la définition peut, dans certaines situations de communication, constituer une tâche, elle est surtout un moyen de structurer les connaissances lexicales. S'entraîner à définir, c'est développer à la fois la compétence d'analyse linguistique par l'exploration des relations dans le lexique, et le contrôle de la production langagière par la construction des énoncés fortement contraints dans leur forme. Il s'agit en plus d'un travail conscient d'organisation de ses connaissances linguistiques, condition que nous considérons comme *sine qua non* sur le chemin vers l'autonomie des apprentissages en L2.

3. Définir, définitions et activités définitionnelles : une séquence pédagogique pour un public A2

Dans cette section, nous présentons une séquence pédagogique comprenant plusieurs activités définitionnelles conçues pour amener l'apprenant à élaborer des paraphrases définitoires des lexies courantes. Elle s'adresse à une classe de niveau de compétences A2. C'est à dessein que nous avons choisi de nous tourner vers un public peu avancé : notre intention est de montrer que les activités définitionnelles sont modulables. Elles peuvent (et doivent) s'adapter aux connaissances dont disposent les apprenants et qu'il s'agit de structurer.

Dans cette optique – il est important de le souligner – l'apprenant n'a pas à réaliser un travail de linguiste lexicographe. Les définitions qu'il devra formuler n'ont aucunement besoin d'être aussi élaborées que celles de la ressource dictionnaire utilisée comme modèle. Trois contraintes restent néanmoins de mise :

1. La paraphrase doit rendre, fût-ce de manière incomplète, le sens lexical du *definiendum* et permettre à l'interlocuteur de le comprendre.
2. Sa structure doit être conforme à ce que les apprenants ont acquis, tant du point de vue lexical que du point de vue grammatical.
3. Bien que lexicalement et grammaticalement en lien avec les connaissances linguistiques de l'apprenant, les définitions produites doivent respecter l'organisation classique en genre prochain et différences spécifiques.

3.1 Explorer une définition lexicographique en quête d'informations

L'usage de dictionnaires constitue, bien sûr, l'une de stratégies les plus fréquentes pour la découverte du sens de nouvelles lexies. La démarche que nous proposons se situe cependant dans une perspective différente. Il s'agit d'explorer des paraphrases définitoires dans leur dimension textuelle, c'est-à-dire en tant que textes structurés selon un certain nombre de règles. Un travail de compréhension écrite en somme, dans le but d'identifier les composantes et les règles de leur agencement. Afin d'orienter l'attention des étudiants vers les propriétés structurelles de la

définition elle-même, nous avons choisi de travailler sur CIGARETTE I, lexie dont le sens n'était pas inconnu aux étudiants.

La paraphrase définitoire que nous proposons comme support écrit est celle du *RL-fr* (figure 2 rappelée ci-dessous) : elle est en effet plus élaborée que celle du *Petit Robert 2011*, dans la mesure où elle comprend plus de CP de natures diverses (cf. la discussion dans la section 2.3.2).

objet 1.1a destiné à être **fumé** 3.a par l'individu X
de forme allongée
constitué d'un petit tube de papier très fin rempli de **tabac**¹ 1.2 haché, que X fait se consumer par une de ses extrémités
que X tient entre ses **lèvres** 1.1 ou entre ses **doigts** 1.1a quand X fume

Figure 2 : Définition de CIGARETTE I dans la base de données lexicales *RL-fr*.

Les lexies employées dans le *definiens*, ainsi que les structures grammaticales (enchaînement de propositions relatives, construction causative *faire se consumer*) ne font pas nécessairement partie des compétences productives en A2, mais on peut admettre qu'elles sont accessibles sur le plan réceptif du fait même que la lexie CIGARETTE I est connue.

De plus, cette définition présente en réalité tout un scénario lié à l'utilisation d'une cigarette dans la mesure où la paraphrase mentionne explicitement le fumeur et le mode de son interaction avec l'objet. Ce n'est pas par hasard que nous venons d'utiliser le terme de scénario. Le *definiens* d'une lexie prédicative telle que CIGARETTE I est facilement transposable en une série d'images auxquelles on pourra recourir pour mettre en avant les composantes de la paraphrase définitoire. La figure 3 reproduit une telle transposition à l'aide de dessins¹⁹.

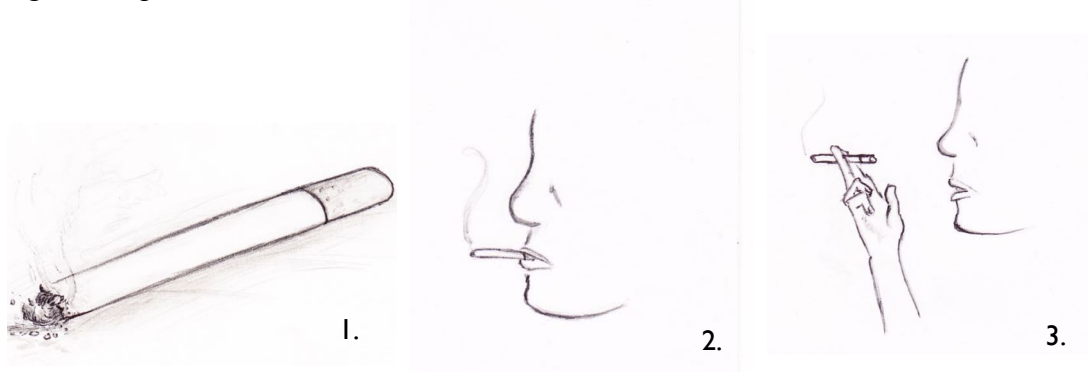


Figure 3 : Images illustrant les composantes de la définition de CIGARETTE I. Dessin : Jan Pouivet.

Une telle visualisation des CP constitue un document déclencheur pour toute une série d'activités définitionnelles proposées avec l'objectif de réfléchir sur les informations que doit comporter la définition d'un certain type d'objet. Une fois identifiées et comprises, elles pourront être formulées en des termes accessibles aux étudiants de niveau A2.

3.2 Organiser des connaissances lexicales

¹⁹ Nous remercions chaleureusement Jan Pouivet pour ces croquis réalisés à notre demande.

L'étape suivante de notre démarche a été organisée autour des moyens lexicaux et grammaticaux à employer dans les tâches de production de paraphrases définitoires envisagées à la fin de cette séquence pédagogique. Précisons qu'il s'agissait de mobiliser des savoirs préalablement acquis, puisque lors des séquences précédentes, nous avons exploré et exploité le matériel didactique proposé dans le chapitre 9 « Lieux de vie » du manuel *Alter Ego*²⁰. Dans les activités définitionnelles qui devaient suivre, nous nous sommes limitée à un seul champ sémantique, celui de la maison et de son équipement.

Notons que cette restriction nous a permis de palier certaines difficultés liées à la décontextualisation : naviguer dans un champ sémantique bien délimité et préalablement exploré revenait à fournir une indication claire sur le sens à définir, ce qui nous a paru particulièrement important en cas de polysémie.

3.2.1 Activité 1 : Matériau lexical pour exprimer la composante centrale

Rappelons que conformément aux niveaux de compétences décrites dans le CECRL, en A2, un apprenant « [p]ossède un répertoire restreint ayant trait à des besoins quotidiens concrets » (CECRL, p. 89). Pour exprimer la CC d'une définition, il a donc été indispensable d'introduire quelques lexies au sens relativement vague, synonymes moins riches des lexies à définir (cf. OBJET dans la définition de CIGARETTE I ci-dessus).

Un tableau de quatre colonnes a été proposé aux étudiants, afin de répartir les lexies du dossier thématique *Lieux de vie*. Rempli d'abord en binômes, il a ensuite été corrigé collectivement au tableau.

Dans le *RL-fr*, ressource sur laquelle nous avons choisi d'appuyer notre démarche, les relations entre les lexies qui fournissent les intitulés des colonnes et celles qui sont réunies sous ces labels sont systématiquement encodées par des fonctions lexicales²¹ de quasi-synonymie ou de hyponymie.

| | | | |
|--------|-------|----------|----------------------------|
| MEUBLE | PIÈCE | APPAREIL | APPAREIL ÉLECTROMÉNAGER |
|--------|-------|----------|----------------------------|

²⁰ Cf. Berthet et al. (2012 : 171-177). Ce support offre en effet un vocabulaire riche, couvrant les différents types de logement, les pièces et les activités qui s'y déroulent, ainsi que l'équipement de maison et son fonctionnement. Un encadré *Aide-mémoire* (p. 173) réunit trois exemples contenant le verbe *servir* (*Le bureau sert de chambre d'amis, Le jardin sert aux jeux des enfants, Le jardin sert à faire la sieste*), particulièrement utile dans les activités de définition que nous avons conçues.

²¹ Pour une présentation du système formel de fonctions lexicales, le lecteur pourra se reporter par exemple à Polguère (2007).

| | | | |
|-------------|---------------|------------|------------------|
| CHAISE | ENTRÉE | RADIATEUR | LAVE-VAISSELLE |
| FAUTEUIL | COULOIR | SÈCHE- | LAVE-LINGE |
| TABLE | SALON | CHEVEUX | MACHINE À LAVER |
| TABLE BASSE | CUISINE | BOUILLOIRE | SÈCHE-LINGE |
| LIT | CHAMBRE | ÉLECTRIQUE | FRIGO |
| BUREAU I | SUITE | | CONGÉLATEUR |
| | PARENTALE | | FOUR MICRO-ONDES |
| | SALLE À | | |
| | MANGER | | |
| | DRESSING | | |
| | SALLE DE BAIN | | |
| | BUREAU II | | |

Tableau 1 : Activité définitionnelle 1 – révision et organisation du vocabulaire.

3.2.2 Activité 2 : Matériau lexical pour exprimer les composantes périphériques

Compte tenu des connaissances lexicales et grammaticales d'un étudiant A2 qui restent, conformément au CECRL, liées essentiellement aux échanges quotidiens, orientés vers des besoins concrets, faits de phrases courtes et simples²², il est clair que les paraphrases définitives élaborées par les apprenants seront plus simples que celles proposées dans les dictionnaires. La granularité de la description du sens est en effet en rapport avec les compétences. La formulation de certaines CP – notamment, celles qui précisent les éléments fonctionnels (*constitué d'un petit tube de papier très fin rempli de tabac haché dans le cas de CIGARETTE I*), le fonctionnement (*que X fait se consumer par l'une de ses extrémités*) et l'usage d'une cigarette (*que X tient entre ses lèvres ou entre ses doigts quand X fume*) – repose sur des compétences du niveau B1, voire B2, plutôt que A2. En revanche, décrire la forme et la finalité pour laquelle un artefact est conçu est entièrement à la portée de notre public.

Ainsi, la mobilisation des connaissances requises pour formuler ces CP était au cœur de la deuxième activité. Elle s'est déroulée en deux étapes. Dans un premier temps, les étudiants devaient associer à chaque objet de la colonne gauche du tableau 2 ci-dessous, l'une des fonctionnalités proposées à droite. Il est intéressant de noter que la correction collective qui a suivi le travail individuel a permis de compléter la liste de droite : par exemple, pour FOUR MICRO-ONDES, les apprenants ont souhaité préciser que cet appareil servait non seulement à réchauffer les plats, mais également à les décongeler.

| | | | |
|---|------------------|---|-----------------------|
| 1 | fauteuil | a | réchauffer les plats |
| 2 | bureau | b | chauffer la pièce |
| 3 | lit | c | s'asseoir |
| 4 | cuisine | d | travailler |
| 5 | four micro-ondes | e | nettoyer la vaisselle |
| 6 | frigo | f | dormir |
| 7 | lave-vaisselle | g | garder la nourriture |

²² Cf. CECRL, chapitre 5, p. 87-89.

| | | | | | | | | |
|---|-----------|---|---|---|---|--------------------|---|--|
| 8 | radiateur | | | | h | préparer les repas | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | |
| c | d | f | h | a | g | e | b | |

Tableau 2 : Exercice d'association et grille de solutions.

Dans un second temps, lors de la mise en commun des résultats, les étudiants ont reçu l'instruction d'aller au-delà d'une simple association des chiffres et des lettres, pour exprimer les liens entre l'objet et sa fonction dans une phrase simple comme celles que nous reproduisons en (7) et (8) :

(7)

Un fauteuil, c'est un meuble pour s'asseoir.

(8)

Un fauteuil, c'est un meuble qui sert à s'asseoir.

Cet exercice oral a permis de réunir, dans une ébauche de paraphrase définitoire, les éléments des tableaux 1 et 2. Dans la mesure où notre objectif premier était de comparer les résultats du jeu d'associations, nous n'avons pas systématiquement repris les erreurs ponctuelles dans la construction des phrases proposées. Privilégier la réflexion sur le contenu propositionnel pour en articuler les différentes composantes nous a paru indispensable avant l'élaboration collective de quelques paraphrases définitoires modèles.

3.2.3 Activité 3 : écriture collective d'une définition modèle

Afin de fixer les repères nécessaires, les étudiants ont procédé à une séance d'écriture collective de définitions susceptibles de servir de modèles pour leurs propres productions spontanées. Celles de CHAISE et de RADIATEUR, citées en (9) et (10) sont issues de ce travail.

(9)

a. Une chaise, c'est un meuble en bois, en métal ou en plastique, avec quatre pieds, un dossier et souvent avec un coussin, qui sert à s'asseoir.

b. Une chaise, c'est un meuble, plutôt petit, en bois, en métal ou en plastique, pour s'asseoir.

(10)

Un radiateur, c'est un appareil en métal, installé sur un mur pour chauffer la pièce.

Comme nous l'avons signalé plus haut, ces paraphrases ne sauraient satisfaire un lexicographe, puisque la description du sens n'est pas complète. Premièrement, le nombre des CP est fort limité : seuls la matière, la taille, les parties les plus saillantes dans le cas de (9a) et l'usage sont mentionnés. De telles concessions s'avèrent indispensables compte tenu des connaissances lexicales et grammaticales d'un apprenant A2.

De plus, le premier actant sémantique de CHAISE et de RADIATEUR, à savoir l'utilisateur, n'est pas explicitement présent dans la paraphrase définitoire, comme il l'est dans celle de CIGARETTE. Là encore, il nous a fallu prendre en compte les

connaissances linguistiques des apprenants. Introduire l'utilisateur X dans le *definiens* aurait imposé entre autres le recours aux pronoms relatifs complexes (*Une chaise, c'est un meuble sur lequel X/on peut s'asseoir*).

Il n'en est pas moins que l'absence d'actant sémantique dans la définition est à nos yeux préjudiciable à long terme pour les apprentissages lexicaux, notamment pour situer la lexie dans un mini-réseau de relations que l'apprenant doit maîtriser pour pouvoir s'en servir dans le discours. Ainsi, même si l'objectif premier des activités proposées était de privilégier les relations d'hypo-, d'hyponymie et de synonymie inclusive plutôt que d'explorer l'ensemble des liens lexicaux, nous avons tenu à sensibiliser les étudiants tout particulièrement aux relations actanciennes, surtout en ce qui concerne le premier actant, c'est-à-dire l'utilisateur de l'artefact dénoté par le *definiendum*. C'est la raison pour laquelle nous avons lancé en interlude un exercice d'expression orale spontanée consistant dans un jeu de questions-réponses qu'illustre l'exemple (11) ci-dessous :

(11)

- a. Avez-vous baissé le radiateur ce matin ?
- b. Qu'est-ce que vous mettez dans le frigo ?
- c. Où est-ce qu'on peut s'asseoir ?

3.2.4 Activité 4 : jeu de questions/réponses en interaction avec le professeur

Cette dernière activité de notre séquence a été par excellence définitionnelle, car elle consistait à construire des paraphrases, en s'appuyant sur les éléments réunis précédemment.

Pour ce faire, nous avons proposé aux étudiants un quizz dans lequel, après avoir posé une question de type *Qu'est-ce que X ?*, chaque participant désignait le ou la camarade qui devait y répondre. Au fur et à mesure que l'on progressait dans cette activité, les étudiants se sont tout naturellement aventurés en dehors du champ sémantique initial, pour s'interroger mutuellement sur le sens des lexies dénotant les objets de leur entourage immédiat (cf. exemple de TROUSSE ci-dessous, dont la polysémie a par ailleurs été prise en compte par l'une des étudiantes). Observons que si les réponses obtenues ne correspondaient pas toujours au sens du *definiendum* (cf. 12b), la structure du *definiens* respectait invariablement l'organisation analytique. Les réponses sont reproduites sans corrections.

(12)

- a. Un radiateur, c'est un appareil qui sert à chauffer la chambre.
- b. Un radiateur, c'est une machine qu'on utilise pour garder les nourriture plus longtemps.
- c. Un radiateur, c'est un appareil qu'on utilise pour chauffer les pièces.
- d. Un lave-vaisselle, c'est un appareil qu'on utilise pour laver les vaiselles.
- e. Une trousse, c'est une meuble qu'on utilise pour mettre des vêtements.
- f. Une trousse, c'est un petit sac qui sert à transporter les stylos.
- g. Une trousse, c'est un objet qui sert à transporter les stylos.

Lors d'une séance ultérieure, afin de remobiliser les connaissances ainsi développées, nous avons proposé un quizz conduisant à adopter une démarche inverse. Il s'agissait alors d'identifier le *definiendum* à partir d'un *definiens*, en répondant à des questions *Comment appelle-t-on/comment est-ce qu'on appelle un grand appareil qui sert à garder la nourriture au frais ?* que les étudiants posaient à tour de rôle.

4. Conclusions

L'échange rapporté dans l'exemple (3) de la section 2.2.1 a suivi de quelques semaines la séquence pédagogique présentée ci-dessus. Il montre qu'une définition bien construite s'insère parfaitement dans une conversation informelle en L2 et qu'elle peut, comme tout autre acte de communication, servir à régler un problème courant.

Cependant, la démarche que nous avons entreprise n'était pas directement liée à ce type de tâches « à visée actionnelle » préconisées par le CERCL. Les activités définitionnelles et la pratique des définitions ont été abordées dans une perspective de développement et de structuration de connaissances linguistiques à moyen et long termes.

Sans en nier la difficulté, nous espérons avoir montré que les activités définitionnelles sont bien modulables : il est possible de les adapter et de les pratiquer même à un niveau de compétences dit de survie. Un tel entraînement nous semble de première importance dès que l'on se situe non pas dans une optique de bénéfices immédiats de réalisation de tâche sociale, mais dans une perspective à plus longue échéance. Tout en améliorant ses performances langagières d'acteur social, l'apprenant est avant tout acteur de ses propres apprentissages : celui qui, en route vers l'autonomie, doit s'équiper d'outils indispensables pour construire un système de connaissances à la fois structuré et dynamique, notamment sur le plan lexical. Les activités définitionnelles en sont un, puisqu'elles constituent un moyen efficace de développer à la fois les mécanismes d'analyse langagière et de contrôle de production (Bialystok, 1991).

Remerciements

Je tiens à remercier chaleureusement Émilie Perrichon et Alain Polguère pour leurs remarques et commentaires sur une première version de cet article.

Bibliographie :

- AITCHISON Jean (2004) : *Words in the Mind. An Introduction to the Mental Lexicon*, Oxford, Blackwell Publishing.
- BERTHET Annie, WAENDENDRIES Monique, HUGOT Catherine, DAILL Emmanuelle et KIZIRIAN, Véronique, M. (2012) *Alter ego + AI*, Paris, Hachette.
- BIALYSTOK Ellen (1991) : « Metalinguistic Dimensions of Bilingual Language Proficiency », in E. BIALYSTOK (ed.) *Language Processing in Bilingual Children*, Cambridge, Cambridge University Press, p. 113-140.

- BOGAARDS Paul (1994) : *Le vocabulaire dans l'apprentissage des langues*, Paris, Didier.
- Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer*. Unité des Politiques Linguistiques, Strasbourg, [en ligne], http://www.coe.int/t/dg4/linguistic/Source/Framework_FR.pdf (consulté le 15/08/2015).
- CUMMING Geoff, CROPP Samantha et SUSSEX Roland (1994) : « On-line Lexical Resources for Language Learners: Assessment of Some Approaches to Word Definitions », *System*, 22(3), p. 369-377.
- CUQ Jean-Pierre (2004) : « Le lexique en situation d'apprentissage guidé : pour une méthodologie d'enseignement interventionniste dans l'enseignement du français langue étrangère », in É. Calaque et D. Jacques (dir.), *Didactique du lexique. Contextes, démarches, supports*, Bruxelles, De Boeck, p. 61-71.
- HANKS Patrick (1987) : « Definitions and Explanations », in J. Sinclair (ed.), *Looking Up: An Account of the COBUILD Project in Lexical Computing*, London-Glasgow, Collins, p. 116-136.
- JAKOBSON Roman (1963[1992]) : *Essais de linguistique générale. Les fondations du langage*, Paris, Éditions de Minuit.
- KLEIN Wolfgang (1995) : *Second Language Acquisition*, Cambridge, Cambridge University Press.
- LEW Robert (2010) : « Multimodal Lexicography: The Representation of Meaning in Electronic Dictionaries », *Lexikos*, 20, p. 290-306
- LEW Robert et DZIEMIANKO Anna (2006) : « A new type of folk-inspired definition in English monolingual learners' dictionaries and its usefulness for conveying syntactic information », *International Journal of Lexicography*, 19-3, p. 225-242.
- LUX-POGODALLA Veronika et POLGUÈRE Alain (2011) : « Construction of a French Lexical Network. Methodological Issues », in : *Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011*, ESSLI 2011, Université de Ljubljana, p. 54-61.
- MEARA Paul (2009) : *Connected words. Word associations and second language vocabulary acquisition*, Amsterdam/Philadelphia, John Benjamins Publishing Company.
- MEL'CUK Igor et POLGUÈRE Alain (2008) : « Prédicats et quasi-prédicats sémantiques dans une perspective lexicographique », *Revue de linguistique et de didactique des langues (Lidil)*, 37, p. 99-114.
- NESI Hilary (1998) : « Defining a Shoehorn: the Success of Learners' Dictionary Entries for Concrete Nouns », in S. Atkins (ed.), *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators*. Tübingen: Max Niemeyer Verlag, p. 159-178.
- PAJOOESH Parto (2014) : « Knowledge of Word Concepts and Referents », in J. Milton et T. Fitzpatrick (eds) *Dimensions of Vocabulary Knowledge*, Palgrave Macmillan, p. 73-91.
- PERRICHON Émilie (2008) : *Agir d'usage et agir d'apprentissage en didactique des langues-cultures étrangères enjeux conceptuels, évolution historique et construction d'une nouvelle perspective actionnelle*. Thèse de doctorat, Université de Saint-Étienne.

- PLANE Sylvie (2005) : « Les activités définitionnelles au service des apprentissages lexicaux », *Pratiques*, 125/126, p. 115-138.
- PLANE Sylvie et LAFOURCADE Bernadette (2004) : « Pour une approche discursive de l'apprentissage du lexique : les activités définitionnelles », in É. Calaque et D. Jacques (éds) *Didactique du lexique. Contextes, démarches, supports*, Bruxelles, De Boeck.
- POLGUÈRE Alain (2007) : « Lexical function standardness », in L. Wanner (dir.) *Selected Lexical and Grammatical Issues in the Meaning-Text Theory. In Honour of Igor Mel'čuk*, John Benjamins, Amsterdam/Philadelphia, p. 43-95
- (2014) : « From Writing Dictionaries to Weaving Lexical Networks ». *International Journal of Lexicography*, 27/4, 396-418.
- PUREN Christian (2014) : Approche communicative et perspective actionnelle, deux organismes méthodologiques génétiquement opposés... et complémentaires, [En ligne], <http://www.christianpuren.com/mes-travaux/> (consulté le 14 juin 2016).
- RIEGEL Martin (1987) : « Définition directe et indirecte dans le langage ordinaire : les énoncés définitoires copulatifs », *Langue française*, 73, p. 29-53.
- (1990) : « La définition, acte du langage ordinaire. De la forme aux interprétations », in J. Chaurand et F. Mazière (éds) *La définition*, Paris, Larousse, p. 97-110.
- SIKORA Dorota (2012) : « Définir le sens dans un réseau lexical », *Studia Romanica Posnaniensia*, 39/3, Poznań, Éditions universitaires UAM, p. 63-79.
- SNOW Catherine, CANCINO Herlinda, DE TEMPLE Jeanne et SCHLEY Sara (1991) : « Giving formal definitions: a linguistic or metalinguistic skill? » in E. Bialystok (ed.), *Language Processing in Bilingual Children*, Cambridge, Cambridge University Press, p. 90-112.
- STOCK Penelope (1988) : « The Structure and Function of Definitions », in M. Snell-Hornby (ed.), *Zurilex'86 Proceedings: Papers Read at the Euralex International Congress*, Tübingen, Francke Verlag, p. 81-89.
- VIGNER Gérard (2012) : « Écrire en FLE : Quel enseignement pour quel apprentissage ? », in R. Bouchard et L. Kadi (éds), *Didactiques de l'écrit et nouvelles pratiques d'écriture. Le Français dans le monde. Recherches et application*, Paris, CLE International, 51, p. 16-33.
- VYGOTSKY Lev S. (1986), *Thought and Language*, Cambridge MA, MIT Press.

Ressources électroniques :

- « Cambridge Dictionaries Online », [en ligne]
<http://dictionary.cambridge.org/dictionary/english/> (consulté le 28/11/2015).
- « COBUILD », [en ligne] <http://www.collinsdictionary.com/dictionary/english-cobuild-learners> (consulté le 28/11/2015).

Dorota SIKORA
 Université du Littoral – Côte d'Opale
 HLLI (EA 4030)
 Dorota.Sikora@univ-littoral.fr

Locutions en apprentissage de langue seconde

Dorota Sikora*

Université Littoral Côte-d'Opale, Centre du Musée, 34 Grande Rue,
BP 751, 62321 Boulogne-sur-Mer Cedex

Résumé. L'article propose une réflexion à la fois linguistique et didactique sur les locutions et leur intégration dans l'enseignement de vocabulaire d'une langue étrangère. Les propriétés formelles, sémantiques et fonctionnelles de ces phrasèmes ont été largement décrites par la lexicologie, qui peut de ce fait apporter des indices précieux aux didacticiens. En nous appuyant sur la conceptualisation des locutions issue de la Lexicologie Explicative et Combinatoire, et en mettant à contribution leur modélisation en tant qu'unités lexicales dans la base de données *Réseau Lexical du Français*, nous nous interrogeons dans un premier temps sur les difficultés que les locutions présentent dans l'apprentissage de vocabulaire. Dans la seconde partie, nous proposons les grandes lignes d'une démarche pédagogique qui tient compte de ces propriétés linguistiques, tout en respectant le rythme de progression et d'acquisition de compétences défini par le *Cadre européen commun de référence pour les langues* (CECRL).

Abstract. The present paper addresses the issue of idioms in foreign language teaching. Their formal, semantic, and functional properties have been largely discussed and described in lexicology. Within the conceptual framework developed by Explanatory and Combinatorial Lexicology, and relying on the data encoded in *French Lexical Network*, the study firstly focalises on linguistic properties of idioms as main source of difficulties in language learning. The question that arises then is that of the possible methods of teaching idioms in a foreign language class. Thus, in a second part, the paper outlines a general design of an idiom pedagogy, in accordance with their linguistic properties on the one hand, and on the other, with acquisitional and organisational constraints of the language curricula defined in line with *the Common European Framework of Reference for Languages*.

1. Introduction¹

Depuis l'élaboration des approches lexicalistes (cf. Nation 1982, 1990, Bogaards 1994 et surtout Lewis 1993, 2008) et le développement de la notion de *phraséodidactique* à partir des travaux de Gréciano (1994), il ne semble plus nécessaire de plaider en faveur d'une intégration des unités phraséologiques (désormais UP) dans un cursus de langue seconde (L2)². On reconnaît en effet que le lexique d'une langue contient une multitude d'unités polylexicales de nature différente, qui se prêtent plus ou moins bien – souvent pas du tout – à un traitement décompositionnel. Elles doivent être saisies et apprises comme un

* Dorota.Sikora@univ-littoral.fr

tout. La question qui se pose actuellement n'est donc pas s'il faut les enseigner, mais de savoir comment le faire de manière efficace.

La réflexion proposée ci-dessous poursuit un objectif bien pratique : il s'agit de montrer ce qu'une modélisation du lexique développée par des linguistes peut apporter à la didactique de vocabulaire, notamment à l'enseignement d'unités phraséologiques. Nous limitons la discussion aux locutions, non seulement parce que ce sont, comme le suggère Gläser (1988), des prototypes de tous les phraséologismes, mais surtout pour des raisons méthodologiques. Il est préférable, à notre avis, de ne pas aborder l'ensemble des UP, car les différents types de phrasèmes (locutions, collocations, pragmatèmes, proverbes, etc.) se caractérisent par des propriétés structurelles et fonctionnelles différentes. Par conséquent, les difficultés qu'ils posent aux apprenants et la façon de les traiter en classe ne sont pas les mêmes.

Les manuels de L2 semblent aujourd'hui quelque peu désarmés face aux locutions. Une étude réalisée par Berthemet (2017) montre qu'elles y sont très peu présentes : dix locutions à peine dans *Version originale 1* (livre et cahier d'activités) contre 353 collocations. Elles y sont traitées en marge, souvent par un renvoi vers les dictionnaires. Or, comme nous aurons l'occasion de le voir dans la section 2, ces ressources n'offrent pas toujours aux apprenants l'assistance nécessaire. Partant de ce constat, nous discuterons des aspects didactiques, en nous appuyant sur des recherches expérimentales, menées sur plusieurs langues, qui révèlent et confirment certaines régularités et difficultés récurrentes dans l'apprentissage de vocabulaire (cf. Rodgers 1969, Laufer 1990, Nation 1990, Bogaards 1994, Milton 2009).

Le volet linguistique de la présente étude s'appuie sur une modélisation du lexique issue de la *Lexicologie Explicative et Combinatoire* (désormais la *LEC*), repensée dans le cours des travaux d'implémentation du *Réseau Lexical du Français – RL-fr* (Polguère 2014, 2016). Cette ressource offre aujourd'hui une description lexicographique de plus de 3200 locutions toute nature confondue, ce qui constitue un ensemble de données assez développé pour autoriser quelques généralisations. Chemin faisant, nous en adoptons également les conventions typographiques, en marquant les unités lexicales – lexèmes et locutions – en petites majuscules. En cas de polysémie, nous reprenons la numérotation des acceptions encodées dans le *RL-fr*.

La section qui suit propose une discussion des propriétés formelles et sémantiques des locutions en tant que sources de difficultés potentielles en apprentissage de vocabulaire. Dans la section 3, nous nous arrêtons sur les caractéristiques que l'on pourrait qualifier de fonctionnelles, voire interactionnelles, et sur le traitement que l'apprenant doit en assurer selon qu'il fait usage de ses compétences réceptives et productives. Dans la quatrième partie, avec un regard attentif sur les grilles de compétences du *Cadre européen commun de référence pour les langues* (CECRL 2001, North et Goodier 2018), nous réfléchissons sur les grandes lignes dont il faut tenir compte en construisant une démarche phraséodidactique visant tout particulièrement les locutions. La section 5 nous amènera à conclure que les lexicologues, les didacticiens et les enseignants ont tout intérêt à mieux se connaître et à coopérer.

2. Propriétés linguistiques des locutions et difficultés en apprentissage

Si nous considérons, en suivant la *LEC* (cf. Mel'čuk *et al.* 1995, Mel'čuk 2013), que les locutions sont des unités lexicales (lexies) et par conséquent, des signes linguistiques, on pourra analyser leur structure selon les trois dimensions – formelle, sémantique et combinatoire – décrites par Mel'čuk *et al.* (1995). On sera ainsi en mesure de localiser avec précision les sources de difficultés que ce type de phrasèmes présente aux apprenants de L2. Dans cette section, nous nous intéresserons d'abord aux caractéristiques des signifiants locutionnels (2.1), puis à celles de leurs signifiés (2.2). Les propriétés combinatoires

apparaîtront dans les deux sous-sections en rapport avec les facteurs de difficultés liés aux facettes formelle et sémantique des locutions.

2.1 Structure syntagmatique des signifiants locutionnels

La forme signifiante d'une locution est un syntagme composé d'au moins deux lexèmes reliés par des relations de dépendance. La figure 1 illustre le cas de FLEUR BLEUE, où FLEUR est gouverneur syntaxique de l'adjectif BLEU. Cette association d'un patron particulier, en l'occurrence N Adj, et de lexèmes qui le remplissent est appelée *structure lexico-syntaxique* ou SLS (cf. Pausé 2017).

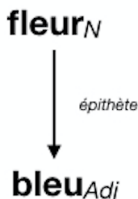


Fig. 1. Structure lexico-syntaxique de la locution FLEUR BLEUE.

Une SLS peut néanmoins être autrement plus complexe, notamment dans le cas de locutions verbales où un réseau de dépendances plus développé requiert, lors de la mise en discours, une solide maîtrise de la flexion verbale et un gestion appropriée de nombreux accords et de reprises anaphoriques. La figure 2 montre l'ensemble de relations entre les lexèmes qui forment la locution RENDRE LA MONNAIE DE SA PIÈCE. On voit dans l'exemple 1, extrait de la base textuelle *Frantext*, que son emploi en discours exige la maîtrise de la flexion verbale, ainsi que le recours aux possessifs et aux clitiques.

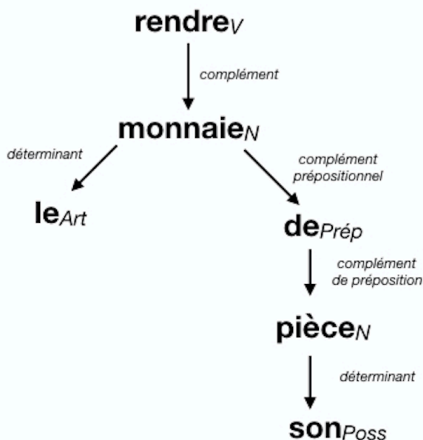


Fig. 2. Structure lexico-syntaxique de la locution RENDRE LA MONNAIE DE SA PIÈCE.

1. Elle a repris de l'aplomb, elle veut me rendre la monnaie de ma pièce. (Frantext)

La structure syntagmatique des signifiants locutionnels est source potentielle de trois types de difficultés souvent repérées dans l'apprentissage du vocabulaire. On observe en effet que la longueur d'une lexie, ainsi que sa partie du discours et ses caractéristiques flexionnelles facilitent ou, au contraire, constituent un obstacle pour la mémorisation et, par

conséquent, pour son réemploi (cf. Laufer 1990). Nous verrons en 2.1.1 et en 2.1.2 comment une forme syntagmatique du signifiant conduit à renforcer ces facteurs de difficulté. Dans la sous-section 2.1.3, nous étudierons le décalage fréquent entre la nature d'une locution, c'est-à-dire sa partie du discours, et les fonctions syntaxiques qu'elle assume dans la phrase.

2.1.1 Longueur

Des tests réalisés sur plusieurs langues indiquent qu'il existe un rapport entre la longueur d'un lexème et la rapidité de son apprentissage. Plus un signifiant est long, plus son apprentissage serait difficile. En commentant ces résultats, Laufer (1990) souligne néanmoins que la longueur devient décisive au-delà d'un seuil défini en nombre de syllabes. Une certaine prudence s'impose dans la réflexion, tout d'abord parce que ces travaux portent essentiellement sur le vocabulaire lexémique, et non pas sur les expressions polylexicales. De plus, ils reposent sur la notion de *mot*, dont on connaît par ailleurs le caractère problématique. Néanmoins, on peut raisonnablement admettre que si la longueur d'un signifiant, mesurée en syllabes, est un facteur qui joue en faveur ou en défaveur d'une mémorisation rapide, elle est à prendre en compte tout autant pour les locutions que pour les lexèmes.

Une structure syntagmatique est, dans l'immense majorité des cas, plus longue qu'un lexème : des dissyllabes comme FLEUR BLEUE sont relativement rares. Or, la longueur a tendance à affecter la mémorisation et l'emploi en discours, notamment pour des raisons phonétiques. Les études réalisées auprès de plusieurs groupes d'apprenants (cf. par exemple : Meara 2009) montrent que les caractéristiques phonologiques et phonétiques jouent un rôle important pour un apprentissage efficace de vocabulaire en L2, tout particulièrement dans sa phase initiale. Une séquence aux caractéristiques articulatoires radicalement différentes de celles qu'on rencontre en L1 s'avère difficile à retenir et, par conséquent, à reproduire. Lorsqu'un apprenant prend la parole, une séquence polylexicale constitue un défi. Elle implique une série de sons à articuler plus complexe que celle d'un lexème et doit en plus s'inscrire dans un contour intonatif approprié.

2.1.2 Complexité flexionnelle

Lorsqu'on analyse les caractéristiques flexionnelles des locutions, il convient de distinguer deux cas de figure. Pour le premier, il n'est pas rare que les marques flexionnelles fassent partie du signifiant locutionnel. C'est le cas notamment de ATOMES CROCHUS, AVOIR LES CHEVILLES QUI ENFLENT, FAIRE DES GORGES CHAUDES et bien d'autres où les éléments de la SLS portent de manière permanente les exposants flexionnels³. Dans ce cas, l'apprenant n'a pas à prendre des décisions pour ce qui est de la flexion à appliquer ; il s'agit des phénomènes de figement, dont nous parlerons plus amplement dans la sous-section 3.2.

Un second cas de figure doit être envisagé sur le plan pédagogique. Les relations de dépendance à l'intérieur du phrasème, qui se manifestent dans ses caractéristiques flexionnelles, peuvent requérir des décisions de la part de l'apprenant. Peut-on, doit-on mettre une locution au pluriel, lorsqu'elle renvoie à plus d'un référent ? Si oui, faut-il fléchir tous les éléments lexicaux ou juste le pivot, c'est-à-dire le plus souvent, le gouverneur syntaxique ? Ainsi, dans le cas d'un syntagme nominal français FLEUR BLEUE, les marques du pluriel s'ajoutent aux deux lexèmes de la SLS. Ailleurs, par exemple pour POLITIQUE DE L'AUTRUCHE, ROND DE JAMBE, TEMPÊTE DANS UN VERRE D'EAU et bien d'autres, seul le gouverneur syntaxique est fléchi, alors que son complément reste invariable. Si le grammème ne se manifeste pas nécessairement à l'oral, le choix de déterminants et de modificateurs adjectivaux ne permet pas d'éviter ces questions. L'exemple 1 ci-dessus a déjà permis d'illustrer le maniement complexe des locutions

verbales : dans un tel cas, la flexion s'accompagne d'un jeu de paramètres déictiques à organiser conformément à la situation d'énonciation.

On note ainsi un rapport entre la longueur d'un phrasème et ses propriétés flexionnelles : plus il y a de lexèmes dans sa SLS, plus les relations de dépendance sont complexes et la flexion compliquée.

2.1.3 Partie du discours : entre nature et fonction

Toute une série de travaux développés depuis plusieurs décennies (cf. Rodgers 1969, Allen et Valette 1972, Phillips 1981) tendent à montrer que, pour l'apprentissage des unités lexémiques, la partie du discours est un facteur susceptible de faciliter l'acquisition ou, au contraire, de la ralentir. Ainsi, les noms et les adjectifs semblent plus faciles à maîtriser que les verbes et les adverbes. Selon les langues et leurs systèmes flexionnels, ces résultats peuvent sans doute être affinés, mais il est certain qu'avec les locutions, la question devient particulièrement épineuse.

Pour l'illustrer, nous prendrons des exemples de locutions gouvernées par un nom, puisque cette partie du discours semble favoriser l'acquisition. De par leur nature, elles devraient être employées dans une phrase en position de sujet ou de complément. C'est effectivement ce qui arrive avec CHEVALIER SERVANT dans les exemples 2 et 3, employé respectivement dans ces deux fonctions.

2. Ton chevalier servant arrive ! (FrWac)
3. Enfin seule pour flirter tout mon soûl au téléphone avec le chevalier servant du moment. (Frantext)

La structure syntaxique de la locution FLEUR BLEUE est identique à celle de PRINCE CHARMANT. En surface, on lui attribue donc la même partie du discours : il s'agit d'une locution nominale. L'exemple 4 la montre cependant modifiée par l'adverbe LÉGÈREMENT et coordonnée avec un adjectif, assumant comme ce dernier une fonction d'épithète habituellement réservée aux adjectifs. Il en est ainsi de la locution nominale MAIN DANS LA MAIN dans l'exemple 5, où elle modifie le nom TRAVAIL, alors que dans 6, elle apparaît comme modificateur adverbial du verbe TRAVAILLER.

4. Cette aubaine s'expliquait en grande partie par l'affection maternelle et légèrement fleur bleue que madame Lamerise portait à son commis, au demeurant fort bon vendeur. (Frantext)
5. Je souhaiterais qu'il y ait un vrai travail main dans la main et plus d'échanges entre nous, explique-t-il. (Internet)
6. On travaille main dans la main avec la paroisse. (Internet)

Cela signifie que l'emploi de nombreuses locutions, tout particulièrement nominales et prépositionnelles, n'est pas celui que l'on pourrait anticiper à partir de leur partie du discours de surface. Ce décalage entre leur nature et leur fonction aboutit à des propriétés combinatoires difficilement prédictibles, ce qui ne facilite pas leur compréhension et encore moins leur réemploi dans le discours.

2.2 Structure sémantique des locutions

Les caractéristiques sémantiques d'un lexème ont une importance toute particulière pour l'apprentissage du vocabulaire, notamment en raison du rôle structurant que jouent les connexions paradigmatiques dans l'organisation de notre lexique mental (cf. Aitchison 2003). De nombreuses recherches montrent en effet que le vocabulaire d'un locuteur, que

ce soit en L1 ou en L2 forme un réseau dont les nœuds sont reliés essentiellement par des relations de sens⁴ (cf. Meara et Wolter 2004, Schur 2007, Meara 2009). Or, certaines particularités sémantiques telles que le caractère abstrait d'un signifié ou sa faible granularité descriptive rendent l'apprentissage plus difficile. Dans cette section, nous nous proposons d'analyser les composants de la structure sémantique complexe que forment les signifiés locutionnels.

2.2.1 Le sens dénotatif des locutions

Tout comme les unités lexémiques, les locutions ont un signifié qu'on peut expliciter dans une paraphrase définitoire. Ainsi, pour expliquer le sens de la locution prépositionnelle AUX ANGES,⁵ on pourra proposer la définition suivante :

AUX ANGES = 'qui éprouve un sentiment de joie intense et une sensation de bien-être'

En tant que lexie pourvue d'un sens dénotatif ou descriptif (cf. Kleiber 1990, 1999), une locution s'inscrit ainsi dans le vaste réseau que forme le lexique, en tissant des relations avec d'autres lexies mono- ou polylexicales. La figure 3 ci-dessous montre le sous-graphe, construit à partir des données encodées dans la base lexicale *RL-fr*, organisé autour de AUX ANGES. Cette lexie est reliée par des relations de quasi-synonymie aux adjectifs RAVI, HEUREUX, CONTENT, etc., ainsi qu'à une autre locution prépositionnelle SUR UN NUAGE. Ces connexions permettent également de connecter AUX ANGES, quoique indirectement, à son quasi-antonyme MALHEUREUX et d'établir les liens avec des lexies nominales telles que BONHEUR, CONTENTEMENT, FÉLICITÉ et SEPTIÈME CIEL, puis avec leur hyperonyme commun SENTIMENT1a⁶. Les couleurs permettent d'identifier des groupements de nœuds (*clusters*) fortement interconnectés.

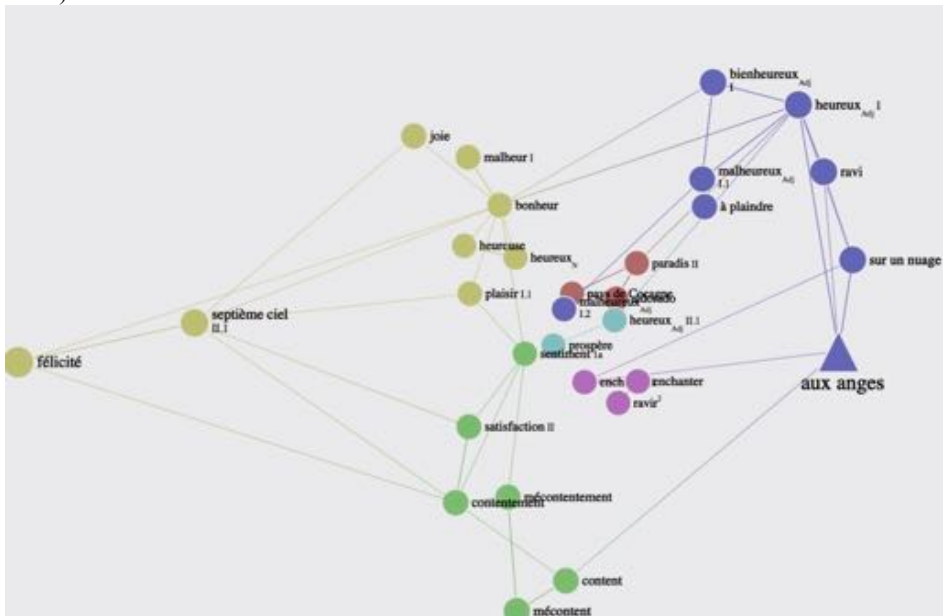


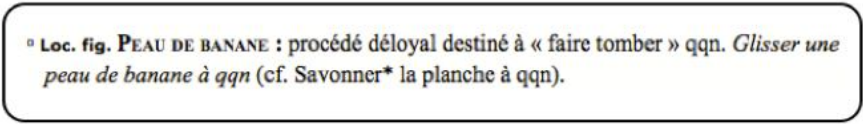
Fig. 3. Réseau de relations paradigmatiques de la locution prépositionnelle AUX ANGES (extrait de la base de données lexicales *RL-fr*).

En s'appuyant sur les paraphrases définitoires, on pourra qualifier le sens dénotatif des locutions de non-compositionnel, puisque leur sens global n'est pas le résultat d'un calcul

effectué sur les sens des lexèmes qui forment sa SLS⁷. Nous verrons plus loin (section 3.1) en quoi la non-compositionnalité d'un signifié locutionnel constitue un défi dans l'apprentissage d'une L2.

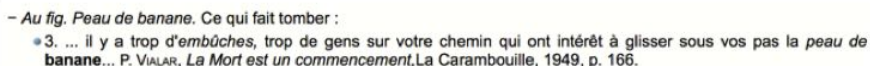
Certaines locutions – HÔTESSE DE L'AIR, ROUGE À LÈVRES, SALLE DE BAINS, SIÈGE SOCIAL – ont un sens dénotatif concret que l'on peut rendre précisément dans une paraphrase. Cependant, dans de nombreux cas, il n'en est pas ainsi. Comme le souligne Fellbaum⁸, les locutions sont là entre autres pour nous permettre de décrire des situations complexes d'une manière courte et condensée. Elles permettent de nommer, voire de qualifier, une grande diversité de situations. Pour ce faire, leur sens contient nécessairement une bonne dose de vague, qui se traduit par des paraphrases définitoires très générales. Elles s'avèrent par conséquent difficilement accessibles, surtout pour un apprenant.

Ci-dessous, dans les figures 4 et 5, nous reproduisons les définitions de PEAU DE BANANE proposées dans le *Petit Robert* électronique et dans le *Trésor de la Langue française informatisé* (désormais *TLFi*)⁹. Elles recourent respectivement aux lexèmes PROCÉDÉ et CE, qui autorisent tous les deux une multitude de référents de nature très diverse. De plus, la paraphrase du *Petit Robert* contient une expression placée entre guillemets, signalant sans plus de précision qu'elle est employée de manière inhabituelle.



° Loc. fig. PEAU DE BANANE : procédé déloyal destiné à « faire tomber » qqn. *Glisser une peau de banane à qqn* (cf. Savonner* la planche à qqn).

Fig. 4. Description lexicographique de la locution PEAU DE BANANE dans le *Petit Robert* électronique.



- Au fig. *Peau de banane*. Ce qui fait tomber :
« 3. ... il y a trop d'embûches, trop de gens sur votre chemin qui ont intérêt à glisser sous vos pas la *peau de banane*... P. VIALAR, *La Mort est un commencement*, La Carambouille, 1949, p. 166.

Fig. 5. Description lexicographique de la locution PEAU DE BANANE dans le *TLFi*.

Les didacticiens savent qu'une lexie au sens abstrait est plus difficile à apprendre que celle qui dénote une entité concrète (Laufer 1990). Qu'en est-il d'une unité au sémantisme vague ? Le cadre de cet article ne permet pas de discuter en détail le rapport entre l'abstrait et le vague ; ainsi, nous limiterons-nous à souligner simplement que, malgré certains recouvrements, il ne s'agit aucunement d'équivalence. Selon Mel'čuk *et al.* (1995), un sens est vague, lorsqu'il correspond à plus d'un référent extralinguistique ou plutôt, dirons-nous, à plus d'une catégorie de référents extralinguistiques¹⁰. En effet, PEAU DE BANANE peut renvoyer à des faits aussi variés qu'une question mal intentionnée, une visite inopportune ou toute action destinée à mettre quelqu'un dans une situation délicate. La locution VAGUE À L'ÂME, au contraire, est bien une lexie abstraite, puisque sa dénotation est construite en intensification à l'aide d'un ensemble de traits sémantiques, mais elle n'est pas vague sur le plan référentiel. Elle sert à nommer une émotion bien particulière.

Le sens dénotatif, paraphrasable tant bien que mal dans une définition lexicographique, joue un double rôle. D'une part sur le plan référentiel, il permet de circonscrire un ensemble de situations auxquelles la locution s'applique. De l'autre, il sert à établir des connexions paradigmatiques avec d'autres unités lexicales (cf. la figure 3 ci-dessus). Les difficultés que pose l'acquisition des lexies abstraites et vagues tiennent certainement à

plusieurs raisons ; notre explication se veut purement lexicale et ne peut, de ce fait, prétendre à exhaustivité.

Les lexies abstraites sont difficiles à retenir, car l'ensemble de traits définitoires en intension qu'il faut maîtriser, par exemple pour VAGUE À L'ÂME, doit ensuite être mis en relation avec un autre ensemble de traits définitoires, par exemple celui de TRISTESSE. Une lexie vague comme PEAU DE BANANE est difficile à retenir, puisqu'elle renvoie à des ensembles de référents très divers (CE ? PROCÉDÉ ?) dont parfois seul le locuteur perçoit des caractéristiques correspondant au sens lexical. La granularité descriptive d'une lexie vague la situe bien au-dessus de niveau de base, celui qui offre l'image d'un prototype du référent (cf. Rosch 1975, Kleiber 1990, Taylor 1995).

Ainsi, le sens dénotatif de plusieurs locutions – abstrait et/ou vague – rend leur acquisition particulièrement complexe, même si par ailleurs les lexicographes parviennent à en formuler, tant bien que mal, une définition.

Cependant, même en présence d'une définition explicite du sens dénotatif, une autre dimension sémantique doit souvent être prise en compte. La locution verbale COUPER LES CHEVEUX EN QUATRE signifie 'examiner une question avec une attention excessive pour des détails sans importance', ce qui la positionne dans le réseau lexical comme quasi-synonyme de CHERCHER LA PETITE BÊTE, des lexèmes PINAILLER, CHIPOTER, ainsi que, vulgairement, de ENCULER LES MOUCHES EN PLEIN VOL. Comme pour de nombreuses autres lexies locutionnelles, pour saisir sa particularité, il est, semble-t-il, nécessaire d'aller au-delà du sens dénotatif.

2.2.2 Sens figuratif ou figuré ?

Lors d'un cours de FLE, nous avons été amenée à expliquer aux étudiants de niveau A2 le sens de la locution nominale PEAU DE BANANE, tel qu'il est illustré dans l'exemple 7 ci-dessous :

7. Une petite peau de banane par-ci, un croc-en-jambe par-là, une méchanceté bénigne et un café offert pour sceller tout ça, vas y on est pote, la vie est belle. (Internet)

Contexte aidant, les étudiants sont parvenus à inférer le sens dénotatif de PEAU DE BANANE : 'situation défavorable mise en place, de manière sournoise, par X contre Y pour faire échouer les projets de ce dernier', mais nos tentatives – certainement maladroitement – de compléter la caractérisation sémantique de cette locution par une description de la métaphore et de l'image qu'elle véhicule (celle d'un individu, qui tout en avançant marche sur une véritable peau de banane, glisse et tombe) ont semé la confusion. Or, sans cette charge imagée, la connaissance de cette lexie locutionnelle ne peut pas être considérée comme complète.

Plusieurs travaux (cf. Gibbs 1994, Dobrovolskij et Piirainen 2005, Piirainen, 2008a, Dobrovolskij 2016 à titre d'exemple) soulignent que les locutions, tout comme les proverbes et bien d'autres unités polylexicales, font partie de ce que l'on dénomme en anglais *figurative language*. C'est à dessein que nous nous arrêtons sur cette notion dans sa forme anglophone. L'adjectif FIGURATIVE souvent employé dans ces travaux pour décrire le sémantisme des locutions est polysémique, avec deux acceptions susceptibles d'interférer, voire de prêter à confusion. Pris dans un premier sens, FIGURATIVE I peut être paraphrasé par 'used with a non basic meaning' et traduit en français par l'adjectif FIGURÉ, opposé à LITTÉRAL. Dans une seconde acception, FIGURATIVE II est un cognat du qualificatif français FIGURATIF, que nous paraphraserons simplement ici comme 'relatif à la représentation ou à l'image d'un objet'¹¹. Pour éviter tout malentendu, il convient d'expliquer le lien entre le caractère figuratif et figuré du sens locutionnel.

Un sens figuratif ?

Une locution telle que COUPER LES CHEVEUX EN QUATRE a, nous l'avons dit plus haut, un sens dénotatif. Cependant, la paraphrase qui en a été proposée semble incomplète. En effet, elle ne rend pas compte de la représentation qui surgit lorsque le sens dénotatif est actualisé dans un énoncé : celle d'un individu qui s'applique à séparer en plusieurs parties des objets aussi fins et insignifiants que des cheveux, tâche dont l'utilité reste difficile à établir. De même, en définissant PEAU DE BANANE (cf. l'exemple 7 ci-dessus) comme 'situation défavorable mise en place, de manière sournoise, par X contre Y pour faire échouer les projets de ce dernier', on ne mentionne d'aucune façon l'image qui accompagne le sens véhiculé.

Ainsi, on pourra admettre que le sens d'une locution est figuratif en ce qu'il véhicule une composante imagée. Selon Dobrovolskij (2016), une composante imagée (cf. *ibidem*, *the image component*) est faite de traces linguistiquement pertinentes de l'image sous-jacente¹². Il en ressort – comme souligné par l'auteur – que cette composante est bien plus qu'une représentation mentale. Elle est linguistiquement pertinente dans la mesure où elle détermine, du moins en partie, la combinatoire lexicale et les restrictions d'emploi.

La composante imagée se laisse en effet détecter dans le sémantisme de PEAU DE BANANE sur plusieurs plans. On la repère par la présence d'un premier actant X ou encore par la composante *d'une manière sournoise* dans la paraphrase définitionnelle. Ces éléments sémantiques peuvent être mis en relation avec une préférence nette que montre cette locution pour le collocatif verbal GLISSER III.2 qui avec une structure argumentale à trois places (X ~ Y à Z, où la position actancielle Y est remplie par PEAU DE BANANE) exprime la causation, et non pas le sens de base 'perdre l'équilibre en se déplaçant sur une surface lisse' (X ~ sur Y).

La définition de *composante imagée* proposée par Dobrovolskij (2016) est plus restreinte que l'image que nous avons tenté de décrire pour COUPER LES CHEVEUX EN QUATRE et PEAU DE BANANE, car elle n'inclut que les éléments repérables dans le sens dénotatif. En nous interrogeant sur leurs origines, nous devons à présent nous tourner vers ce qui, en sémantique des locutions, relève du figuré tout autant que du figuratif.

Un sens figuré ?

Certains travaux (Gibbs 1994) laissent penser que l'adjectif FIGURATIVE devrait être traduit par FIGURÉ pour caractériser le sens dénotatif des locutions par opposition à un sens littéral. Cependant, analyser le sémantisme d'une locution en le considérant comme figuré semble quelque peu problématique pour deux raisons. Premièrement, parce que cela laisse penser qu'une locution a d'abord un sens littéral. Si tel était le cas, pourrait-on encore parler d'une locution ?

Deuxièmement, notre approche s'appuie sur un modèle lexical en synchronie, avec des prérequis théoriques et méthodologiques bien précis. Nous nous intéressons aux locutions en tant qu'unités lexicales, donc en tant qu'entités constitutives du lexique français telles qu'elles se présentent aux locuteurs contemporains. Or, qualifier leur sens dénotatif de *figuré* nous conduirait à postuler l'existence de phrasèmes au statut lexical difficile à déterminer : ce ne serait pas des locutions, puisqu'elles auraient un sens littéral, donc potentiellement compositionnel. Aussi bien une modélisation lexicographique que le traitement didactique d'un tel construit seraient à notre avis hautement problématiques, notamment parce que loin de caractériser le sens lexical, ils s'appliqueraient plutôt aux mécanismes qui produisent des contenus imagés, c'est-à-dire figuratifs.

Il est probable, comme le suggèrent de nombreux auteurs (cf. Gibbs 1994, Dobrovolskij et Piirainen 2005 à titre d'exemple), que le sens dénotatif d'une locution soit effectivement une métaphore ou une métonymie lexicalisée, résultat de mécanismes cognitifs omniprésents dans la pensée et dans la langue, responsables entre autres d'extensions polysémiques dans le cas d'unités monolexémiques. Plutôt que de le qualifier de *figuré* et, par conséquent, de le considérer comme dérivé d'un sens littéral, nous le

traitions comme donnée première en synchronie : une locution peut avoir un caractère figuratif, lié entre autres à la présence d'une composante imagée telle que définie par Dobrovol'skij (2016). Dans la sous-section suivante, nous proposons une approche de contenus figuratifs des locutions à travers leur SLS.

Analyser les contenus figuratifs de locutions

Le sens dénotatif d'une locution est figuratif en ce qu'il contient des traces linguistiquement pertinentes de l'image sous-jacente, qui forment sa composante imagée. Leur pertinence linguistique peut se manifester par exemple sur le plan combinatoire (cf. Dobrovol'skij 2016) ou bien dans le registre : ENCULER LES MOUCHES EN PLEIN VOL se laisse paraphraser par 'attacher une importance excessive aux détails insignifiants', définition qui ne justifie en rien la marque d'usage *vulgaire*. Celle-ci est due au lexème verbal, dont le sémantisme est pourtant absent de la paraphrase définitoire.

La notion de *structure lexico-syntaxique*, SLS (cf. section 2.1 ci-dessus), permet en effet d'établir des liens entre la composante imagée du sens d'une locution et les unités qui la constituent en tant que syntagme. Si la composante imagée est bien présente, c'est en raison d'une mise en relation avec le syntagme libre à la SLS identique.¹³ Autrement dit, le contenu figuratif des locutions telles que PEAU DE BANANE et COUPER LES CHEVEUX EN QUATRE est à lier avec les syntagmes PEAU II DE I BANANE I et COUPER I.1a LES CHEVEUX I EN IV.3 QUATRE 1b. Leurs SLS sont construites de manière identique, mais en numérotant les lexies, nous indiquons l'apport sémantique de chaque élément lexical. Cet apport peut être relativement complexe, différent d'une locution à l'autre. Ci-dessous, les tableaux 1 et 2 permettent de comparer le cas de PEAU DE BANANE et celui de COUPER LES CHEVEUX EN QUATRE.

Tableau 1 : Collocation et locution (PEAU DE BANANE)

| PEAU II DE I BANANE I (syntagme libre) | PEAU DE BANANE (locution) |
|---|---|
| PEAU II – 'enveloppe extérieure d'un fruit' BANANE I – 'fruit de forme allongée et légèrement courbée dont la peau a des propriétés glissantes' <ul style="list-style-type: none"> • <i>La peau de banane a des propriétés physiques dangereuses pour celui qui marche dessus par inadvertance.</i> • <i>La banane est un fruit considéré comme particulièrement apprécié des singes.</i> • <i>Les singes sont considérés comme des animaux agiles et ingénieux, connus pour leurs tours de malice.</i> | PEAU DE BANANE – 'situation défavorable mise en place, de manière sournoise, par X contre Y, pour faire échouer les projets de ce dernier' <ul style="list-style-type: none"> • <i>Situation dangereuse, susceptible de déstabiliser Y.</i> • <i>X agit d'une manière agile et ingénieuse, non dépourvue de malice.</i> |

Dans la colonne de gauche (syntagme libre), en italiques, apparaissent des composantes relevant à la fois du sens de BANANE et des connotations liées non seulement à cette lexie, mais également au lexème SINGE qui n'apparaît pas explicitement¹⁴. La pertinence linguistique de cette référence simiesque se manifeste dans le collocatif GLISSER : glisser quelque chose à quelqu'un suppose en effet une certaine agilité, qu'on reconnaît habituellement aux singes justement.

Le cas de COUPER LES CHEVEUX EN QUATRE semble plus complexe encore en ce que le syntagme libre offre potentiellement beaucoup plus d'éléments que n'en contient la composante imagée de la locution. De nouveau, nous les réunissons dans la colonne de

gauche du tableau 2, en sélectionnant à droite ceux qui forment la composante imagée de la locution.

Tableau 2 : Collocation et locution (COUPER LES CHEVEUX EN QUATRE)

| COUPER I.1a LES CHEVEUX I EN IV.3 QUATRE 1b (syntagme libre) | COUPER LES CHEVEUX EN QUATRE (locution) |
|---|---|
| <p>COUPER I.1a – ‘séparer Y en plusieurs petites parties à l’aide d’un instrument tranchant’ CHEVEU I – ‘poil qui recouvre la tête’</p> <ul style="list-style-type: none"> • <i>Un cheveu un élément du corps très fin, délicat et fragile. Pris à l’unité, il est à peine visible.</i> • <i>COUPER sélectionne un argument Instrument, dont l’emploi n’est pas indispensable pour un Y aussi délicat qu’un cheveu.</i> • <i>Couper un cheveu (en quatre) conduit à diviser quelque chose de fin, fragile et insignifiant, en créant des parties encore plus fines et encore plus insignifiantes.</i> • <i>C’est un travail inutile.</i> • <i>Couper un objet fin en plusieurs parties encore plus fines est un travail long et délicat.</i> • <i>Un tel travail repose sur l’application, la précision et la méticulosité de l’individu qui le réalise.</i> | <p>COUPER LES CHEVEUX EN QUATRE – ‘examiner une question avec une attention excessive aux détails dépourvus d’importance’</p> <ul style="list-style-type: none"> • <i>travail inutile et dépourvu d’intérêt</i> • <i>perte de temps</i> • <i>agent X présenté sous un jour défavorable de pinailleur</i> |

Les éléments sémantiques réunis dans la colonne de gauche – composantes de sens lexical, implications conventionnelles (*entailments*) et contextuelles – sont nettement plus nombreuses que les traces linguistiquement pertinentes repérables à droite. En effet, la scène décrite par le syntagme libre forme une image suffisamment riche pour nourrir potentiellement deux types de contenus sémantiques différents : l’un positif, car mettant en avant les qualités telles que soin et précision apportés au travail, et l’autre, réellement lexicalisé, tendant vers le pôle négatif avec l’idée d’un entêtement inutile.

2.3 Propriétés linguistiques des locutions et difficultés d’apprentissage : un bilan

Cette section a permis de passer en revue les propriétés qui caractérisent les locutions en tant qu’unités lexicales, c’est-à-dire en tant qu’unités constitutives du lexique. D’un point de vue formel, ce sont des syntagmes aux structures plus ou moins complexes. De même, leur sens ne se limite pas, dans une bonne partie des cas, à la dimension dénotative qui, par ailleurs, peut être abstraite ou vague. Bien que non compositionnel en synchronie, il maintient souvent un lien avec les lexèmes de sa SLS. Les locutions semblent cumuler les propriétés qui sont des sources nombreuses de difficultés connues dans l’apprentissage de vocabulaire. Pour les maîtriser, un apprenant de L2 doit s’approprier une masse

d'informations de nature très différente. Cependant, comme nous le verrons dans la section 3, ces connaissances – relatives toujours à une locution particulière plutôt qu'à tous les phrasèmes de ce type – doivent être complétées par des compétences appropriées.

3. Fonctionnement linguistique des locutions et compétences d'apprenants

Pour décrire les locutions, on recourt aux notions de *figement*, de (*non-*) *compositionnalité*, de (*non-*) *analysabilité*, de *transparence* et d'*opacité*, de *motivation*. Si les définitions de ces concepts sont loin de faire l'unanimité dans la communauté des phraséologues, c'est parce que, nous semble-t-il, elles décrivent des aspects différents, souvent complémentaires, du fonctionnement sémantique et discursif des locutions. C'est d'ailleurs la raison pour laquelle il est si intéressant, à notre avis, d'étudier ces notions dans un contexte didactique. Dans cette section, notre objectif est de montrer dans quelle mesure il convient de les prendre en compte dans l'acquisition de différentes compétences en L2. Celles-ci, quel que soit le niveau auquel le CECRL (2001, North et Goodier 2018) les assigne, se répartissent en compétences réceptives et compétences productives¹⁵.

3.1 Locutions et compétences réceptives de l'apprenant

Les compétences réceptives en langue étrangère reposent à la fois sur des connaissances déclaratives et procédurales que l'apprenant doit mobiliser pour comprendre un message oral ou écrit, et sur une démarche interprétative qui lui permettra de faire un usage approprié de ces connaissances pour mettre en place des inférences et des devinements face à des éléments implicites ou simplement inconnus. La première tâche qu'il doit mener à bien consiste à décomposer des segments du discours pour isoler les éléments significatifs et leurs interactions tant sur le plan sémantique que pragmatique.

On pourrait difficilement envisager l'apprentissage d'une langue autrement qu'à partir de la compréhension : l'étudiant est en effet amené à donner une interprétation à un segment de discours, à le décomposer pour identifier les unités significatives. Par exemple, face à la locution *PEAU DE BANANE* dans l'énoncé (7), une première étape consiste à l'identifier comme un signifiant polylexical. Soit il s'agit d'une lexie déjà connue, soit l'étudiant doit partir à la découverte d'un sens inconnu. Il dispose, pour ce faire, d'un certain nombre d'indices co(n)textuels et lexico-syntaxiques.

Dans sa démarche interprétative, il est possible que, dans un premier temps, notre apprenant procède à une décomposition en N_1 (*PEAU II* 'enveloppe externe d'un fruit'), $Prép$ (*DE I*), N_2 (*BANANE* 'fruit à chair compacte, de forme allongée et courbée, avec une enveloppe jaune'). Or, en raison de la non-compositionnalité du sens dénotatif, le résultat d'un calcul effectué à partir de ces lexèmes s'accorde mal avec les indications cotextuelles (*CROC-EN-JAMBE*, *MÉCHANCETÉ*, *CAFÉ*) conduisant vers un autre sens potentiellement actualisé dans (7). Autrement dit, une réanalyse s'impose pour formuler une autre hypothèse.

7. Une petite peau de banane par-ci, un croc-en-jambe par-là, une méchanceté bénigne et un café offert pour sceller tout ça, vas y on est pote, la vie est belle. (Internet)

Dans (7), *PEAU DE BANANE* fait partie d'une coordination avec les lexèmes *CROC-EN-JAMBE* et *MÉCHANCETÉ II*, permettant d'identifier le champ sémantique d'attitudes et de comportements dont relève la locution. Un raisonnement analogique pourra conduire même à une structure argumentale similaire : tout comme *CROC-EN-JAMBE* et *MÉCHANCETÉ II*,

PEAU DE BANANE engage deux participants (celui qui la glisse et celui qui en subit les conséquences).

Un deuxième type d'indices à explorer s'offrent à travers la SLS de la locution. En poursuivant la recherche de signification, il convient d'observer les lexèmes qui forment le syntagme libre PEAU II DE I BANANE I. Il s'agit d'une réanalyse susceptible d'amener l'apprenant à une certaine propriété des peaux de banane qui les rend glissantes et, de ce fait, dangereuses pour l'équilibre et l'intégrité physique de quiconque marche dessus. Le succès de cette opération dépend aussi bien du vocabulaire dont on dispose que des connaissances extra-linguistiques et des inférences que l'on est en mesure de construire.

En ajustant la définition proposée par Langlotz (2006), on pourra dire qu'une locution est analysable, lorsque, à travers les contenus sémantiques¹⁶ associés aux éléments lexicaux qui la forment, on est en mesure d'identifier son sens. L'analysabilité n'est donc pas une propriété de l'unité lexicale elle-même, mais plutôt une caractéristique interactionnelle : elle tient entièrement à celui ou celle qui procède à l'analyse. En conséquence, la locution peut devenir transparente ; comme l'explique Svensson (2008 : 84) : "when presented with an expression, a language user understands it without any problems, without any other previous knowledge than understanding the separate words that make up the expression, then it is transparent."

L'analyse réalisée par le locuteur apprenant permet non seulement de construire une interprétation et, chemin faisant de comprendre le sens dénotatif de la locution, mais également de la motiver, c'est-à-dire d'établir le ou les liens qui connectent son sens à la signification de la collocation ou du syntagme libre correspondants. On pourra ainsi admettre que si une locution est sémantiquement transparente, elle devient motivée. Mais une fois de plus, il s'agit alors non pas de propriété inhérente des lexies locutionnelles, mais d'une caractéristique interactionnelle. On verra ci-dessous que la transparence est parfois trompeuse et conséquemment, la motivation caduque.

Les indices dont disposent les locuteurs apprenants sont parfois extrêmement minces. Dans une série de tests, deux groupes d'étudiants de Master Didactique des Langues, locuteurs natifs du français d'une part, étudiants étrangers de niveau B2 de l'autre, ont dû examiner une liste de treize locutions parmi lesquelles figurait TEMPÊTE DANS UN VERRE D'EAU. Dans un premier temps, elles leur ont été présentées de manière entièrement décontextualisée pour vérifier s'ils les connaissaient et, en cas de réponse négative, s'ils parvenaient à analyser leur SLS pour procéder à des inférences, en les motivant. À ce stade, de nombreuses réponses de locuteurs apprenants consistaient à expliquer que TEMPÊTE DANS UN VERRE D'EAU dénote un problème rencontré. En effet, tout comme pour COUPER LES CHEVEUX EN QUATRE discuté en 2.2.2, l'image d'une perturbation qui se déchaîne dans un petite quantité d'eau destinée à être bue est suffisamment riche pour donner lieu à plus d'une métaphore, entre autres celle qui en fait un obstacle empêchant d'avaler le contenu du verre. Les contenus sémantiques de la SLS ont donc bel et bien été analysés, sans que l'on puisse qualifier la locution de véritablement transparente pour nos apprenants. L'hypothèse formulée sur le sens dénotatif de la locution fut erronée, et la motivation établie non pertinente. Comme le souligne Langlotz (2006), la motivation n'est pas construite par le locuteur au hasard, car elle repose toujours sur un nombre limité de métaphores (et – ajoutons – de métonymies) possibles, mais il se peut que l'on ne choisisse pas celle qui est réellement encodée en L2.

Pour la deuxième étape du test, un corpus d'exemples a été distribué aux étudiants : quinze extraits en provenance de la base textuelle Frantext illustraient les emplois des locutions précédemment soumises à l'analyse. Pour TEMPÊTE DANS UN VERRE D'EAU, nous avons sélectionné le passage suivant¹⁷ :

8. Mais j'avais tenu au courant votre état-major impérial. Nous avons une petite tempête dans un verre d'eau. Cela va s'apaiser très vite. (Frantext)

Malgré le verbe S'APAISSER dans le cotexte droit, cet exemple (comme la plupart de ceux que nous avons soumis aux deux groupes) a été jugé peu utile pour orienter les hypothèses sur le sens de la locution. Ce résultat peut s'expliquer par le caractère sémantiquement condensé mis en avant par Fellbaum, que nous avons déjà mentionné dans la section 2.2.1 ci-dessus : une locution permet de désigner, de manière courte et synthétique, des situations par ailleurs complexes. Une locution offre ainsi un condensé sémantique, qui permet des économies d'indices cotextuels. De plus, observons que le choix d'une locution telle que TEMPÊTE DANS UN VERRE D'EAU permet d'ajouter une attitude propositionnelle, en l'occurrence un jugement, qualifiant le désaccord d'inimportant.

3.2. Locutions et compétences productives de l'apprenant

Pour transmettre un contenu sémantique à l'aide d'une locution, l'apprenant doit récupérer dans son stock lexical sa forme complète, c'est-à-dire sa structure lexico-syntaxique (SLS). Il n'est pas possible, en effet, de l'assembler en fonction du sens à transmettre – c'est à ce stade donc qu'il faut composer avec le figement du signifiant locutionnel.

Lorsque l'apprenant souhaite dire par exemple qu'un individu X réalise des efforts en collaboration avec Y en vue d'objectifs communs, il lui est certes possible d'employer le verbe COLLABORER ou COOPÉRER, ou encore de recourir à un syntagme libre et compositionnel comme TRAVAILLER ENSEMBLE, etc. Mais le français met aussi à disposition des locuteurs la locution RAMER DANS LE MÊME SENS. Le figement du signifiant se manifestant notamment dans le blocage de substitution paradigmatique (cf. Gross 1996) représente alors un défi certain. L'exemple 9 ci-dessous, produit par un locuteur étranger, illustre le cas d'infraction à ce blocage :

9. Ici, personne ne rame dans la même direction que moi.

(Conversation privée, sans jeu de mots, mars 2015)

En produisant la phrase (9), le locuteur a remplacé SENS par DIRECTION ; les deux quasi-synonymes étant en effet substituables dans de nombreux cotextes. Même si le signifié de SENS ne figure pas dans la définition de la locution (entièrement non-compositionnelle), il n'en joue pas moins un rôle pour ce qui est de son analysabilité et, par conséquent, pour la motivation du sens dénotatif. Si, ayant auparavant rencontré et analysé cette locution *via* sa SLS, l'étudiant est parvenu à identifier la composante imagée et à en inférer le sens global, il n'est pas étonnant que les relations paradigmatiques d'un lexème tel que SENS restent actives¹⁸. D'autant plus que, dans bien des cas, le blocage paradigmatique n'est pas complet : certaines locutions – par exemple MANGER DES YEUX – autorisent le recours aux quasi-synonymes, le cas échéant au verbe DÉVORER.

Les recherches menées sur 70 langues d'Europe et d'autres continents dans le cadre du projet *Widespread Idioms in Europe and Beyond* (Pirainen, 2008b) montrent que des locutions similaires sont largement attestées dans plusieurs langues de familles différentes : c'est le cas de NAGER CONTRE LE COURANT, par exemple. Cependant, ces équivalences doivent être observées avec une certaine prudence, car si ces ressemblances peuvent faciliter la démarche interprétative, elles constituent des sources de confusions potentielles dans la production des énoncés. Tout d'abord, même lorsqu'une locution est construite en L2 sur un patron syntaxique similaire, sa SLS peut être légèrement différente sur le plan lexical. Si dans plusieurs langues, un non-conformiste *nage contre le courant*, il peut également avoir à affronter une marée (en écossais, en portugais) ou une chute d'eau (en celte d'Irlande).

Ensuite, l'existence d'un équivalent locutionnel en L2 ne permet pas de conclure à l'identité sémantique. Dobrovolskij (2016) analyse entre autres le cas de SCHWARTZ AUF

WEIß en allemand et celui de ЧЕРНЫМ ПО БЕЛОМУ en russe, également traduits en français par NOIR SUR BLANC. Or, bien que la composante imagée d'un texte imprimée de manière claire et nette soit comparable, le sens lexicalisé par la locution dans ces langues diffère sensiblement. Si le russe (comme le français) met en avant le caractère évident de l'information transmise, l'allemand insiste sur son côté officiel lié à l'impression d'un document. Une certaine méfiance s'impose donc face aux locutions en provenance de langues différentes, même là où leurs SLS présentent des similarités importantes, particularités flexionnelles mises à part.

Quand un locuteur étranger choisit d'exprimer un sens par une locution plutôt que par un lexème, il lui est nécessaire de s'assurer que la composante imagée est bien celle qu'il veut faire apparaître. Il n'est pas rare que l'apprenant souhaite transmettre à son interlocuteur celle qui se trouve lexicalisée dans sa L1, soit parce qu'il ne connaît pas de locution équivalente en L2, soit parce qu'il n'y trouve pas la même composante imagée. Ainsi, face à un excès de soins et d'attentions qui aboutissent à un résultat contraire de celui qui était escompté, un locuteur polonophone pourra employer l'expression ZAGLASKAĆ KOTA NA ŚMIERC (lit. 'caresser le chat jusqu'à ce qu'il en crève'). Désireux d'en rendre le sens avec la composante imagée, il enchâsse une traduction mot à mot de la SLS dans une mise en garde de type « Comme on dit chez nous... ».

Ainsi, de quel côté que l'on se tourne, les difficultés se multiplient. Les propriétés structurelles et fonctionnelles des locutions constituent une source de problèmes considérables que les apprenants de L2 doivent savoir traiter et résoudre. D'une part, lors des premières rencontres dans le discours, elles imposent une démarche interprétative reposant sur une analyse et une recherche de liens entre le sens dénotatif, même lorsque celui-ci est entièrement non-compositionnel, et les lexèmes de sa SLS. De l'autre, leur emploi en discours ne peut se faire sans un traitement complexe¹⁹, dont il convient de tenir compte dans une démarche phraséodidactique.

4. Démarches pédagogiques à construire dans la pratique de classe de langue

Les propriétés spécifiques des locutions que nous venons de passer en revue et qui constituent des sources de problèmes importants dans l'apprentissage du vocabulaire doivent, bien évidemment, être prises en compte dans l'élaboration des cursus et dans la conception des activités de classe. Or, les formations en langue sont également soumises à d'autres contraintes. Nous ne nous arrêterons pas ici sur le manque de temps, qui pèse lourdement sur les choix pédagogiques, pour réfléchir à celles qu'impose une approche fortement utilitaire des méthodes actionnelles recommandées notamment par le *Cadre européen de référence pour les langues* (2001, North et Goodier 2018). Sans mettre en cause les méthodes actionnelles elles-mêmes, c'est la vision exclusive de l'apprenant « acteur social » promue par ces documents qui nous paraît pour le moins discutable (cf. Sikora 2016a)²⁰. Nous convenons volontiers cependant de leur utilité pour l'organisation des cursus en ce qui concerne le rythme de développement des compétences et surtout, l'évaluation de leur acquisition. C'est d'ailleurs cet aspect pratique qui nous paraît le plus intéressant pour la phraséodidactique. Réinterprétés par les linguistes didacticiens, le CECRL ainsi que ses documents satellites²¹ fournissent un référentiel de difficultés à envisager et des indices pour introduire progressivement dans les apprentissages des locutions abordables à tel ou tel niveau de compétences.

4.1 Compétences d'apprenants selon le Cadre européen commun de référence pour les langues (CECRL)

Dans la pratique de classe, l'utilité du CECRL tient aux niveaux de compétences qu'il définit et qui permettent de construire des programmes assurant une progression. La relative cohérence des apprentissages quel que soit leur cadre institutionnel est d'autant plus importante que dans leur immense majorité, les cours sont aujourd'hui orientés vers la valorisation des acquis par des certifications. Ainsi, en partant d'un niveau Pré-A1 (débutant complet), l'apprenant doit développer successivement un certain nombre d'aptitudes, en acquérant des connaissances linguistiques et socioculturelles. Au fur et à mesure que ses apprentissages progressent, il sera en mesure de réaliser de plus en plus de « tâches ». Concernant l'étendue de vocabulaire d'un étudiant A2, les nouveaux descripteurs du CECRL stipulent que celui-ci :

Peut produire de brèves expressions courantes afin de répondre à des besoins simples de type concret : détails personnels, routines quotidiennes, désirs et besoins, demandes d'information.

Peut utiliser des modèles de phrases élémentaires et communiquer à l'aide de d'expressions mémorisées, de groupes de quelques mots et d'expressions toutes faites, sur soi, les gens, ce qu'ils font, les lieux, les biens, etc.

Possède un répertoire limité de courtes expressions mémorisées couvrant les premières nécessités vitales des situations prévisibles ; des pannes fréquentes et des malentendus surviennent dans les situations imprévues. (North et Goodier 2018 : 137)

Un étudiant A2 doit donc être en mesure de parler de ses occupations quotidiennes, de préparer un voyage, d'apporter des informations sur ses goûts, ses loisirs et ses activités professionnelles par exemple. Parmi les tâches qu'il doit pouvoir réaliser figurent par exemple l'achat d'un billet, la réservation d'un hôtel ou la prise d'un rendez-vous, à l'aide – conformément à ses compétences grammaticales – de structures simples. Avec ce bagage lexical, l'apprenant « [p]eut utiliser des structures simples correctement mais commet encore systématiquement des erreurs élémentaires comme, par exemple, la confusion des temps et l'oubli de l'accord ; le sens général reste cependant clair » (*ibidem*, p. 139). Toutes les compétences sont ensuite progressivement développées jusqu'à un niveau C2 de locuteur autonome.

Les expressions idiomatiques parmi lesquelles comptent les locutions sont mentionnées seulement parmi les compétences du niveau C1. L'étudiant peut alors « comprendre un vaste éventail d'expressions idiomatiques, variations de registres et ajuster son propre registre » (*ibidem*, p. 144). Cependant, conformément à ces mêmes descripteurs, une bonne maîtrise de ces UP, reposant notamment sur une claire conscience des différents niveaux connotatifs, n'est attendue qu'en C2.

Les propriétés structurelles, sémantiques et fonctionnelles des locutions que nous avons discutées dans les sections précédentes, avec le lot de difficultés qu'elles occasionnent d'une part, les descripteurs de compétences du CECRL de l'autre pourraient conduire à la conclusion selon laquelle l'intégration de ces phrasèmes dans les apprentissages n'est pas à envisager avant le niveau C1. Dans les faits, cela signifierait que ce volet phraséodidactique n'est plus vraiment didactique, puisqu'à ce stade, l'apprenant est quasi autonome et n'a plus besoin d'une démarche guidée.

Néanmoins, une lecture attentive des descripteurs de compétences prévues en A2 (*supra*) indique à notre avis le contraire. HÔTESSE DE L'AIR, SOLDAT DU FEU, SALLE À MANGER ou encore ROUGE À LÈVRES sont bien des expressions courtes aux structures syntaxiques non seulement simples, mais aussi récurrentes en français. Faciles à mémoriser, elles permettent de parler des routines et des situations quotidiennes. Il est donc

bien possible, et à nos yeux souhaitable, d'enseigner le vocabulaire locutionnel à ce niveau peu avancé, tout en respectant la progression dans le développement des compétences, à condition cependant d'adopter des pratiques réfléchies et ciblées.

4.2 Locutions et phraséodidactique dans un cursus de langue

De nombreuses études (Bogaards 1994, Meara et Wolter 2004, Milton 2009 à titre d'exemple) montrent que lorsqu'une démarche pédagogique cible un ensemble d'unités lexicales, celles-ci sont acquises de manière plus efficace et durable qu'en cas de rencontres aléatoires dans des textes ou dans des échanges quotidiens. D'un point de vue quantitatif, la différence est notable : selon les activités menées en classe, le nombre de *mots* appris par heure d'exposition peut aller du simple au triple (cf. Milton 2009). Compte tenu des propriétés structurelles et fonctionnelles qui caractérisent les locutions et de l'ensemble de difficultés qu'elles présentent, il apparaît d'autant plus important d'adopter une démarche pédagogique appropriée, construite à partir d'une conceptualisation de ces unités lexicales. Leur fréquence peu élevée (Colson 2008) en corpus est un argument supplémentaire en faveur d'une phraséodidactique raisonnée : ayant peu de chances de rencontrer régulièrement une même locution dans le discours dans plusieurs contextes différents, l'apprenant a d'autant plus besoin d'acquérir des connaissances et des stratégies nécessaires pour l'intégrer dans son vocabulaire.

4.2.1 Place et objectifs de la phraséodidactique

La phraséodidactique n'est pas une démarche à part : elle doit au contraire s'intégrer à l'ensemble des activités de classe, en en partageant les finalités et le rythme de progression. Elle poursuit néanmoins deux types d'objectifs qui lui sont propres.

Comme pour tout enseignement lexical, son objectif premier est de développer le vocabulaire de l'apprenant d'un point de vue quantitatif. L'étudiant a besoin de maîtriser des unités du lexique, et celles-ci ne sont pas toujours des lexèmes. C'est ainsi qu'il élargit progressivement le nombre de domaines (cf. CECRL) dont il peut parler et la granularité des informations qu'il est en mesure de comprendre et de transmettre. Par exemple, pour traiter, à l'oral comme à l'écrit, des questions gastronomiques, il lui sera possible d'aller au-delà de PAIN, BEURRE, VIANDE, TOMATE, etc. avec des locutions telles que ASSIETTE DU TERROIR, MISE EN BOUCHE, CIGARETTE RUSSE, PIÈCE MONTÉE, POMME FRITE et, bien évidemment, PLAISIRS DE LA TABLE²² : le vocabulaire du champ sémantique de 'nourriture' en devient plus riche, plus diversifié, plus chargé culturellement.

Le second objectif pourrait être qualifié de *stratégique* : en plus d'apprendre une à une un certain nombre de locutions, il s'agit de mettre progressivement en place un *modus operandi*. En effet, le lexique d'une langue telle que le français offre une multitude de phrasèmes non-compositionnels, tout en s'enrichissant continuellement de nouvelles unités phraséologiques, y compris locutionnelles. On pourrait difficilement imaginer un parcours de L2 entièrement orienté vers la phraséodidactique, et quand bien même cela serait possible, il est peu probable qu'il parvienne à en faire un tour complet. En tant que locuteur, l'apprenant doit donc acquérir des stratégies qui lui permettent de traiter et de maîtriser en autonomie les locutions qu'il rencontre. Cyr (1998 : 5) indique à juste titre que « l'apprentissage d'une L2 peut être vu comme tout autre processus de traitement de l'information ». Or, dans le cas des locutions, ces informations sont, comme nous l'avons vu dans les sections précédentes, très complexes. Pour les traiter, l'apprenant a besoin de stratégies appropriées, c'est-à-dire de procédures et d'outils à mettre en œuvre pour améliorer ses apprentissages (cf. définition de *stratégie* proposée par Oxford 1990).

L'une des stratégies les plus fréquentes consiste à consulter une ressource dictionnaire. Cette stratégie peut, du moins partiellement, apporter le résultat escompté,

si l'information recherchée s'y trouve. Mais sa mise en place repose sur une autre démarche, à savoir celle qui consiste à repérer dans le texte une unité de sens au signifiant polylexémique. Autrement dit, l'apprenant doit avoir conscience qu'un sens dénotatif à prendre en compte dans le calcul compositionnel peut être porté par un phrasème, c'est-à-dire par une forme syntagmatique. La question est donc de savoir comment, sans se lancer dans un cours de linguistique et de lexicologie, simplement à travers des activités de classe, apprendre à apprendre le vocabulaire locutionnel.

4.2.2 Locutions dans l'enseignement de vocabulaire

Notre objectif dans cet article n'est pas de proposer une séquence pédagogique consacrée à l'apprentissage de locutions. La pratique de classe montre qu'il est difficile de construire un cours – un ensemble d'activités structurées qui tendent vers des objectifs bien déterminés – sans penser aux apprenants que l'on aura en face de nous. Ainsi, nous réunissons ci-dessous des éléments dont – à notre avis – il convient de tenir compte, si l'on veut concevoir une didactique du lexique incluant le vocabulaire locutionnel. Même si la phraséodidactique ne limite pas son champ d'action aux locutions, nous nous en tiendrons ici strictement à ces phrasèmes, car il s'agit de réfléchir à une pédagogie qui prenne en compte les propriétés formelles et fonctionnelles des unités phraséologiques d'un type particulier. Des propositions pédagogiques portant sur les locutions peuvent s'avérer inapplicables aux collocations ou aux pragmatèmes du fait de leurs caractéristiques différentes.

Malgré la complexité de leurs structures syntaxiques et sémantiques, il nous semble essentiel d'introduire des unités locutionnelles dès le niveau A2, voire A1, à condition néanmoins de les choisir soigneusement. Les critères à appliquer sont multiples : il est possible notamment de les sélectionner en tenant compte de leur complexité formelle et sémantique d'une part, et des recommandations du CECRL de l'autre. Si nous les discutons ci-dessous sous forme d'une liste, il ne faut en aucun cas y chercher une hiérarchie à respecter. En réalité, il s'agit de critères que l'on doit croiser.

1. Les domaines définis dans le CECRL correspondent *grosso modo* à ce que la linguistique qualifie de *champs sémantiques*. En A1, il s'agit de privilégier le vocabulaire relatif à la vie de famille, aux professions, à la nourriture, ainsi qu'à l'environnement matériel immédiat (objets quotidiens, meubles, accessoires, etc.). L'objectif à ce stade des apprentissages est de nommer des objets, plutôt que de les qualifier ou de s'intéresser aux relations entre eux. Il est dès lors possible de proposer un travail autour des locutions telles que HÔTESSE DE L'AIR, SOLDAT DU FEU, FER À REPASSER, MACHINE À LAVER, SALLE DE BAINS, PETIT COPAIN, etc. Deux bénéfices potentiels en découlent. En ce qui concerne le développement quantitatif du vocabulaire, les apprenants acquièrent des moyens de communiquer dans les champs sémantiques/domaines correspondant à leur niveau de compétences. Parallèlement, ils sont amenés vers des stratégies appropriées notamment vers une prise de conscience qu'un signifié n'est pas nécessairement associé à un lexème isolé : son signifiant est souvent polylexical.
2. Le travail avec un public débutant est orienté vers le concret, comme le montrent les domaines/champs sémantiques détaillés dans le CECRL. Il est donc préférable de se concentrer sur des locutions comme celles que nous venons d'évoquer, dont la description sémantique repose essentiellement sur le sens dénotatif, sans (trop de) recours à une composante imagée (cf. sous-section 2.2 *supra*). Au fur et à mesure que l'étudiant progresse vers des niveaux plus avancés, il est possible de prendre en charge les contenus figuratifs, en concevant des activités autour des images sous-jacentes des sens dénotatifs.

3. Si l'on s'en remet au descriptif des compétences proposé dans le CECRL, il convient également de tenir compte des propriétés structurelles du signifiant. Les étudiants A1 et A2 sont en mesure de s'exprimer avec des constructions simples. Par ailleurs, le vocabulaire nominal semble plus accessible pour les apprenants à cette étape (cf. la sous-section 2.1.3 *supra*). Tous ces indices conduisent à opter, dans un premier temps, pour des locutions nominales, avant d'introduire des phrasèmes locutionnels relevant des autres parties du discours.
4. Pour permettre aux étudiants A1 d'intégrer ces locutions dans leur vocabulaire productif, notamment à travers des exercices et activités de réemploi, il est nécessaire de s'assurer que la partie du discours de surface correspond aux fonctions que la locution remplit dans la phrase, c'est-à-dire à sa partie du discours profonde (cf. sous-section 2.1.3 *supra*). Conformément aux recommandations du CECRL, en A1, le locuteur apprenant recourt à des phrases simples de forme Sujet Verbe Complément. Le maniement de syntagmes nominaux qui s'emploient non pas comme des sujets ou des compléments, mais comme des modificateurs adverbiaux et/ou adjectivaux, pourra difficilement être envisagé avant un niveau de compétences B2.
5. À des niveaux avancés, il est important de proposer aux étudiants un travail de classe visant à la fois le sens dénotatif et son élaboration à partir de la composante imagée qu'il contient. Autrement dit, il s'agit d'accompagner les apprenants dans leur recherche de la motivation d'une locution pour les aider à identifier ces contenus figuratifs. Plusieurs raisons justifient un tel choix pédagogique. D'un point de vue linguistique, on ne peut pas employer une locution dans le discours sans maîtriser l'image qu'elle véhicule. Pour l'acquisition, conformément à la théorie de double codage (*dual coding theory*) développée par Clark et Paivio (1991), la mise en relation des informations lexicales et des contenus figuratifs faciliterait la mémorisation et le ré-emploi en contexte.

Les critères énumérés ci-dessus permettent de sélectionner un ensemble de locutions en accord avec le niveau de compétences des apprenants, avec un degré de difficulté qui n'est pas insurmontable pour le public concerné. Elles peuvent être intégrées dans les compétences réceptives et productives.

Les méthodes communicatives et actionnelles privilégient néanmoins l'usage de documents authentiques, originellement destinés aux locuteurs natifs. Si l'enseignant est parfois amené à les adapter à son public (par exemple en les raccourcissant), il n'est pas toujours possible, ni souhaitable, de les expurger complètement de tout élément idiosyncrasique. On peut alors y trouver une locution qui présente un niveau de difficulté trop élevé pour le niveau de compétences des apprenants. C'est à l'enseignant qu'il revient alors de faire un diagnostic et d'adopter une démarche appropriée, mais il lui faudra pour cela une connaissance des différentes facettes des unités du lexique et des problèmes qu'elles occasionnent. S'il juge une locution trop complexe sur le plan formel, sémantique et combinatoire pour s'attendre à ce qu'elle soit réemployée par l'apprenant, peut-être pourra-t-il limiter ses explications au sens dénotatif, en laissant (temporairement) de côté les contenus figuratifs ? Avec des apprenants plus avancés, au contraire, il sera possible d'organiser des activités ciblant justement la composante imagée, pourquoi pas dans une perspective comparée plurilingue dans des classes hétéroglottes ? L'enseignement n'est-il pas l'art de choisir les informations pertinentes pour soutenir les apprentissages *hic et nunc*, tout en préparant le terrain pour ceux à venir ?

5. Conclusion

L'objectif que nous nous sommes posé en entamant cette réflexion était de montrer en quoi une approche lexicologique des locutions devrait permettre d'envisager une démarche phraséodidactique dans l'enseignement de vocabulaire en L2. Certes, la phraséodidactique ne se limite aucunement aux locutions fortes, mais c'est à dessein que cette étude s'est focalisée sur ce type de phrasèmes. Il s'agissait de montrer comment leur conceptualisation dans un modèle lexical conduit à en penser une approche pédagogique. Pour ce faire, dans la section 2, nous avons étudié leurs propriétés formelles, sémantiques bien décrites par la lexicologie, afin de vérifier en quoi ces caractéristiques représentent des sources de difficulté potentielles pour les apprenants. Nous avons discuté ensuite (section 3), l'articulation entre ce que nous avons qualifié de propriétés fonctionnelles et les deux types de compétences – réceptives et productives. En réunissant les éléments ainsi dégagés, nous venons d'esquisser, dans la section 4, les grandes lignes pour un traitement pédagogique des locutions.

Faute de place, le présent article s'arrête à une étape où il serait intéressant de rebrousser chemin et d'envisager les rapports entre la lexicologie et la lexicographie d'une part, et la (phraséo)didactique de l'autre selon une perspective inverse. Nous pensons en effet que tout comme l'étude du lexique et la réflexion sur les entités qui le constituent aident les enseignants à concevoir des parcours d'apprentissage, une conscience claire des besoins des utilisateurs enseignants et apprenants permet aux lexicologues et aux lexicographes d'améliorer la macro- et la microstructure des ressources lexicales, en y incluant des informations nécessaires et en les structurant de manière à guider les uns et les autres. Présenter une unité phraséologique en tant qu'entrée lexicographique à part entière, avec un article qui la décrit devient une pratique courante (cf. *The Cambridge English Dictionary* en ligne, les bases *WordNet* de différentes langues ou les réseaux lexicaux comme le *RL-fr*), sans pour autant être devenu la règle. Certaines ressources fournissent, dans les articles dictionnaires, des informations très utiles pour la pratique pédagogique. Par exemple, le système d'étiquettes sémantiques (cf. Polguère 2011) implémenté dans les réseaux lexicaux tels que le *RL-fr* permettra à terme l'extraction d'unités lexicales relevant d'un même champ sémantique.

De plus, les dictionnaires fournissent parfois des indications précises (et précieuses !) quant aux niveaux de compétences d'apprenants susceptibles d'en acquérir la maîtrise. *The Cambridge English Dictionary* par exemple estampille la locution *DOWN IN THE DUMPS* de la marque C2, en précisant ainsi à quel niveau de compétences elle correspond. Autrement dit, les lexicologues indiquent explicitement que ce phrasème est à étudier par des apprenants C2, qui sont déjà des locuteurs autonomes.

Les échanges entre les linguistes – lexicologues et lexicographes tout particulièrement pour ce qui concerne les enseignements lexicaux – et les acteurs du terrain (enseignants, didacticiens, apprenants eux-mêmes) offrent de nouvelles perspectives à tous ceux qui y participent. Nous espérons avoir montré ce que la lexicologie peut apporter à l'enseignement : une modélisation des données lexicales et une conceptualisation des entités qui le forment. En adossant ses activités à un modèle robuste, l'enseignant se dote d'outils permettant de proposer des parcours d'apprentissage structurés, en anticipant les difficultés. Nous n'avons fait que signaler rapidement, en attendant de développer ces idées ailleurs, que les bénéfices sont réciproques, car la pratique de classe et la réflexion didactique indiquent en retour comment enrichir les informations lexicales proposées dans les ressources dictionnaires.

Bibliographie

- Aitchison, J. (2003). *Words in the Mind. An Introduction to the Mental Lexicon*. Oxford: Basil Blackwell.
- Allen, E. D. et Valette, R. (1972). *Modern Language Classroom Techniques: A Handbook*. New York: Harcourt, Brace, Jovanovich.
- Berthemet, E. (2017). *Manuels et apprentissage du lexique idiomatique*. 85^e Congrès de l'Association Francophone pour le Savoir (ACFAS), Université McGill, Montréal, Canada, 8-12 mai 2017.
- Bogaards, P. (1994). *Le vocabulaire dans l'apprentissage des langues*. Paris: Hatier/Didier.
- Cadre européen commun de référence pour les langues. Apprendre, évaluer, enseigner*. (2001). Conseil de l'Europe. <en ligne> rm.coe.int/16802fc3a8, consulté le 20 mars 2018.
- Clark, J. M. et Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3, 233-262.
- Colson, J.-P. (2008). Cross-linguistic phraseological studies. Dans S. Granger et F. Meunier (éds). *Phraseology. An Interdisciplinary Perspective*. Amsterdam, Philadelphia (Pa): John Benjamins Publishing Company, 191-206.
- Cuq, J.-P. (2004). Le lexique en situation d'apprentissage guidé : pour une méthodologie d'enseignement interventionniste dans l'enseignement du français langue étrangère. Dans É. Calaque et J. David (éds), *Didactique du lexique. Contextes, démarches, supports*. Bruxelles: De Boeck, 61-71.
- Cyr, P. (1998). *Les stratégies d'apprentissage*. Paris : Clé-international.
- Dobrovol'skij, D. (2016). The notion of "inner form" and idiom semantics. *Études et travaux d'Eur'ORBEM*, 1, 21-36.
- Dobrovol'skij, D. et Piirainen, E. (2005). *Figurative Language. Cross-Cultural and Cross-Linguistic Perspectives*. Amsterdam, Boston: Elsevier.
- Gibbs, R. W. J. (1994). *The Poetics of Mind. Figurative Thought, Language, and Understanding*. Cambridge, New York: Cambridge University Press.
- Gläser, R. (1988). The grading of idiomaticity as a presupposition for a taxonomy of idioms. Dans W. Hüllen et R. Schulze (eds), *Understanding the Lexicon*. Tübingen: Max Niemeyer, 264-279.
- Gréciano, G. (1994). Vorsicht, Phraseoaktivität! Dans B. Sandig (ed.), *Tendenzen der Phraseologieforschung*. Universitätsverlag Brockmeyer.
- Gross, G. (1996). *Les expressions figées en français. Noms composés et autres locutions*. Paris : Ophrys.
- Kleiber, G. (1990). *La sémantique du prototype*. Paris : Presses universitaires de France.
- Kleiber, G. (1999). *Problèmes de sémantique. La polysémie en questions*. Villeneuve d'Asq : Éditions du Septentrion.
- Klein, W. (1995). *Second Language Acquisition. Cambridge*. Cambridge: Cambridge University Press.
- Langlotz, A. (2006). *Idiomatic Creativity. A Cognitive-Linguistics Model of Idiom Representation and Idiom-Variation in English*. Amsterdam, New York: John Benjamins Publishing Company.
- Laufer, B. (1990). Why are some words more difficult? – Some intralexical factor that affect the learning of words. *International Review of Applied Linguistics in Language Teaching*, 28(4), 293-308.
- Lewis, M. (1993). *The Lexical Approach. The State of ELT and a Way Forward*. London: Language Teaching Publications.
- Lewis, M. (2008). *Implementing the Lexical Approach. Putting Theory into Practice*. Cheriton: Heinle.
- Meara, P. (2009). *Connected words: Word associations and second language vocabulary acquisition*. Amsterdam, Philadelphia (Pa): John Benjamins.
- Meara, P. et Wolter, B. (2004). V_Links: Beyond Vocabulary Depth. *Angles on the English Speaking World*, 4, 85-97.
- Mel'čuk, I. (2013). Tout ce que nous voulions savoir sur les phrasèmes, mais... *Cahiers de Lexicologie*, 102 (1), 129-149.
- Mel'čuk, I., Clas, A. et Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Bruxelles : De Boeck Supérieur.
- Milton, J. (2009). *Measuring Second Language Acquisition*. Bristol: Multilingual Matters.
- Nation, I. P. A. (1982). Beginning to learn foreign language vocabulary. *RELC Journal*, 13, 61-68.
- Nation, I. P. A. (1990). *Teaching and learning vocabulary*. New York: Newbury House.

- North, B., Goodier, T. (2018). *Cadre européen de référence pour les langues: apprendre, enseigner, évaluer. Volume complémentaire avec les nouveaux descripteurs*. Strasbourg: Conseil de l'Europe. <en ligne> rm.coe.int/cecr-volume-complementaire-avec-de-nouveaux-descripteurs/16807875d5, consulté le 20 mars 2018.
- Oxford, R. (1990). *Language Learning Strategies: What Every Teacher Should Know*. New York: Newbury House Publishers.
- Pausé, M.-S. (2017). *Structure lexico-syntaxique des locutions du français et incidence sur leur combinatoire*. Université de Lorraine et ATILF CNRS (UMR 7118), Nancy.
- Phillips, T. A. (1981). *Difficulties in Foreign Language Vocabulary Learning and a Study of Some Factors Thought to be Influential*. M.A. Project, Birkbeck College, University of London.
- Piirainen, E. (2008a). Figurative phraseology and culture. Dans S. Granger et F. Meunier (eds), *Phraseology. An Interdisciplinary Perspective*. Amsterdam, Philadelphia (Pa): John Benjamins Publishing Company, 207-228.
- Piirainen, E. (2008b). Phraseology in a European framework. A cross-linguistic and cross-cultural project on widespread idioms. Dans S. Granger et F. Meunier (eds), *Phraseology. An interdisciplinary perspective*. Amsterdam, Philadelphia (Pa): John Benjamins Publishing Company, 243-258.
- Polguère, A. (2011). Classification sémantique des lexies fondée sur le paraphrasage. *Cahiers de Lexicologie*, 98 (1), 197-211.
- Polguère, A. (2014). From Writing Dictionaries to Weaving Lexical Network. *International Journal of Lexicography*, 27 (4), 396-418.
- Polguère, A. (2016). La question de la géométrie du lexique. Dans F. Neveu, G. Bergounioux, M.-H. Côté, J.-M. Fournier, L. Hriba, et S. Prévost (éds.), *Actes du « 5e Congrès Mondial de Linguistique Française (CMLF 2016) »*. Université de Tours, <en ligne> [dx.doi.org/10.1051/shsconf/20162701002](https://doi.org/10.1051/shsconf/20162701002)
- Rodgers, T. S. (1969). On measuring vocabulary difficulty: an analysis of item variables in learning Russian-English vocabulary pairs. *International Review of Applied Linguistics*, 7, 327-343.
- Rosch, E. (1975). Cognitive Representation of Semantic Categories. *Journal of Experimental Psychology*, 104, 192-233.
- Schur, E. (2007). Insights into the structure of L1 and L2 vocabulary networks: intimations of small worlds. Dans H. Daller, J. Milton, et J. Treffers-Daller (eds.), *Modelling and Assessing Vocabulary Knowledge*. Cambridge, New York: Cambridge University Press, 182-203.
- Sikora, D. (2016a). Activité(s) de définition dans l'apprentissage d'une langue seconde. *Cahiers de Lexicologie*, 2 (109), 117-143.
- Sikora, D. (2016b). Identification d'unités phraséologiques et équivalence sémantique dans la traduction. In *Computerised and Corpus-based Approaches to Phraseology: Monolingual and Multilingual Perspectives/Fraseología computacional y basada en corpus: perspectivas monolingües y multilingües*. Genève: Tradulex. <en ligne> www.tradulex.com/varia/Europhras2015.pdf,
- Svensson, M. H. (2008). A very complex criterion of fixedness: Non-compositionality. Dans S. Granger et F. Meunier (eds), *Phraseology. An Interdisciplinary Perspective*. Amsterdam, Philadelphia (Pa): John Benjamins Publishing Company, 81-93.
- Taylor, J. R. (1995). *Linguistic Categorization*. Oxford: Oxford University Press.

¹ Je remercie chaleureusement Gaétane Dostie et les relecteurs anonymes : leur regard attentif et leurs remarques, questions et propositions d'amélioration m'ont beaucoup aidée dans la préparation de la version finale de cet article.

² Nous parlerons indifféremment de *langue étrangère*, *français langue étrangère* (FLE) et de *langue seconde L2*, en admettant avec Klein (1995) que toute langue apprise après une première est seconde. Pour cet exposé, il n'est pas indispensable de procéder à des distinctions plus fines entre une L2 et une L3.

³ Tous mes remerciements au relecteur anonyme pour avoir attiré mon attention sur ce cas de figure, qui est effectivement à distinguer de celui où l'apprenant doit lui-même décider de la forme fléchie à adopter.

⁴ Les similarités de structures en réseau ne signifient pas l'identité. Pour une présentation et discussion des différences dans l'organisation du vocabulaire chez un locuteur de L1 et chez un apprenant, voir Schur (2007).

⁵ La locution AUX ANGES a été discutée dans Sikora (2016b) en rapport avec un problème plus large d'identification des structures et, par conséquent, de la nature des locutions. Les arguments réunis montrent qu'il s'agit en effet non pas d'une locution verbale ÊTRE AUX ANGES, mais d'une locution prépositionnelle à emplois adjectivaux. Elle peut en revanche constituer une base pour des collocations formées avec les verbes supports ÊTRE, SEMBLER et METTRE.

⁶ Rappelons que le *RL-fr* est une implémentation d'un modèle relationnel de type *système lexical*, qui prend en compte tous les liens (paradigmatiques, syntagmatiques, ainsi que ceux d'inclusion formelle), sans les hiérarchiser (cf. Polguère 2014, 2016). Un tel réseau de connexions paradigmatiques représente une description relativement riche du sens dénotatif, même en l'absence de paraphrase définitoire. Une exploration d'un sous-réseau à la recherche d'informations sémantiques représente pour l'apprenant une tâche significative, c'est-à-dire l'impliquant personnellement. Or, comme l'observe Cuq (2004), ce caractère significatif des tâches réalisées en cours d'apprentissage est décisif pour son efficacité.

⁷ Nous reprenons ainsi la définition de non-compositionnalité proposée par Mel'čuk (2013) sans pour autant en reproduire la formalisation, qui ne semble pas indispensable pour les problèmes qui nous occupent ici.

⁸ Communication personnelle.

⁹ Accessible via la plateforme du *Centre National de Ressources Textuelles et Lexicales* (www.cnrtl.fr).

¹⁰ Cf. Mel'čuk *et al.* (1995 : 60)

¹¹ Notons que le TLFi, consulté via la plateforme CNRTL mentionne pour le français un emploi de *figuratif* au sens de *figuré* (cf. www.cnrtl.fr/definition/figuratif), acception qui semble d'ailleurs première du point de vue diachronique.

¹² Cf. Dobrovolskij (2016 : 23) : "the *image component* (...) of a figurative lexical unit is defined as the linguistically relevant traces of the mental image underlying its semantics."

¹³ Pausé (2017) parle à ce propos de *projection structurale*.

¹⁴ Nous n'utilisons pas ici le terme de *cadre* (cf. *frame*, Dobrovolskij 2016) découpant un domaine source pour la métaphore que serait alors la locution. Comme nous l'avons expliqué plus haut, il est problématique – dans un modèle synchronique – d'analyser une locution (donc une unité lexicale) appartenant au lexique contemporain comme métaphore d'un syntagme libre qui n'a pas sa stabilité au sein du module lexical. Nous convenons que la présente explication est un peu expéditive et qu'elle mériterait une justification plus étoffée, difficile à fournir dans les limites de cet article.

¹⁵ Chacun de ces deux types se subdivise en compétences relatives à l'écrit et à l'oral. Notons au passage que les compétences réceptives sont en général nettement plus étendues que les compétences productives.

¹⁶ Il est préférable de parler de *contenus sémantiques*, plutôt que de *sens*, puisque parfois, la réanalyse effectuée s'appuie également des connotations, comme dans le cas de PEAU DE BANANE.

¹⁷ Dans un sous-corpus de travail allant de 1950 à 2017 défini dans *Frantext*, seules six occurrences de cette locutions apparaissent. Nous avons sélectionné la concordance qui contenait une indication – S'APAISER – susceptible d'orienter les inférences.

¹⁸ Rappelons que plusieurs travaux consacrés au lexique mental (cf. Aitchison 2003 pour une discussion) mettent en avant le rôle structurant des relations paradigmatiques dans son organisation.

¹⁹ Nous ne parlerons pas ici de *flexibilité formelle* (Pausé 2017) qui autorise par exemple la passivation ou l'introduction de modificateurs, pas plus que nous n'évoquerons les défigements. Ces faits de langue sont affaire de locuteurs natifs et autonomes d'une langue (niveaux C1 et C2 du CECRL). Compte tenu de leurs compétences, ces derniers ne sont plus à considérer comme des apprenants à proprement parler.

²⁰ Par rapport à la version de 2001 du CECRL, la lecture du volume complémentaire du CECRL, paru en février 2018 (North et Goodier 2018) permet de constater une certaine ouverture, notamment vers une réflexion sur la langue, et donc un retour réflexif et critique sur les connaissances la concernant. C'est dire que l'apprenant y apparaît non plus seulement comme destiné à réaliser des tâches essentiellement sociales et des projets au sein d'un groupe. On peut donc espérer qu'à terme il soit envisagé selon une dimension plus personnalisée et personnalisante en tant qu'acteur de ses propres apprentissages (cf. Sikora 2016a).

²¹ Parmi les documents qui complètent le CECRL, citons à titre d'exemple l'outil d'autoévaluation intitulé *Portfolio européen des langues* (PEL).

²² Ces locutions ont été extraites de la base *RL-fr* via une requête portant sur l'étiquette sémantique 'aliment préparé'. Celle-ci a permis d'afficher toutes les lexies porteuses de cette étiquette. C'est dire qu'une ressource lexicale structurée de manière adéquate est un outil précieux dans le travail d'un enseignant. Nous y revenons dans la section 5, qui conclura nos réflexions.

- Gülich, E. & Mondada, L. (2001). Analyse conversationnelle. In G. Holtus, M. Metzeltin & C. Schmitt (dir.), *Lexikon der Romanischen Linguistik* (vol. 1, 2, p. 196-250). Tübingen : Niemeyer.
- Nonnon, É. (2012). La didactique du français et l'enseignement du vocabulaire, dans vingt ans de revues de didactique du français langue première. *Repères*, 46, 33-72.
- Pitsch, K. (2005). Interaction, auto-organisation et pratiques d'inscription. *Acquisition et interaction en langue étrangère*, 22, 73-99.
- Plane, S. (2005). Les activités définitionnelles au service des apprentissages lexicaux. *Pratiques*, 125/126, 115-138.
- Plane, S. & Lafourcade, B. (2004). Chapitre 3 : Pour une approche discursive de l'apprentissage du lexique : les activités définitionnelles. In É. Calaque & J. Davis (dir.), *Didactique du lexique* (p.47-59). Bruxelles : De Boeck.

ENTRE SENS ET COMPÉTENCES MÉTALINGUISTIQUES : LOCUTIONS EN PHRASÉODIDACTIQUE DU FLE

Dorota Sikora, Université du Littoral Côte d'Opale & UR HLLI (EA 4030)

Présentation et objectifs de l'article

Cet article propose une réflexion sur certains aspects de la compétence métalinguistique³⁴ dans l'enseignement du vocabulaire, et plus précisément en phraséodidactique. Il s'appuie sur les résultats d'un test réalisé auprès d'un public étudiant dans le cadre du programme de recherche *Images culturelles dans l'apprentissage des langues : enjeux didactiques et linguistiques* (ICal) développé à l'Université du Littoral Côte d'Opale. Ce n'est pourtant pas les caractéristiques figuratives des locutions testées qui nous occuperont ici, mais bien des questions de connaissances lexicales à propos de ces unités phraséologiques (désormais UP) et leur restitution par des locuteurs.

Le terme de *phraséodidactique* proposé par Gréciano (1984) implique une démarche particulière à adopter là où l'enseignement du vocabulaire prend pour objet des UP³⁵. En didactique des langues, celles-ci sont mieux connues sous le nom d'*expressions idiomatiques*, classe aux contours flous, à l'intérieur de laquelle se côtoient des locutions (*tempête dans un verre d'eau*), des collocations (*café noir*), des clichés (*Le compte y est*), des pragmatèmes (*peinture fraîche*), etc. Or, puisqu'il existe de nombreux types de phrasèmes, l'objectif de la phraséodidactique est avant tout de proposer des activités pédagogiques qui tiennent compte de cette diversité : une collocation n'a ni le même statut au sein du lexique, ni le même fonctionnement qu'une locution, ce qui conduit à en postuler un traitement différencié dans la pratique de classe. Pour l'assurer, l'enseignant a cependant besoin de conceptualiser les objets qu'il veut faire apprendre, d'en avoir une certaine représentation et d'être en mesure de la transmettre par des méthodes adéquates.

C'est en postulant une telle prise en compte des particularités des UP d'une part, et de l'autre, l'anticipation des difficultés qu'elles présentent pour les apprenants (cf. Sikora, 2018) que nous avons choisi de focaliser notre étude sur les locutions fortes³⁶, *ipse eo* phrasèmes non-compositionnels : il semble

³⁴ Nous employons le terme de compétence métalinguistique comme dénomination collective d'un ensemble de savoir-faire (i.e. de sous-compétences permettant d'organiser et de constamment mettre à jour ses connaissances), ce qui explique l'alternance du singulier et du pluriel par lequel nous y référons tout au long de ce texte.

³⁵ L'intérêt pour la phraséodidactique va croissant, avec notamment les travaux de Ettinger (1998), Boers (2001), Abel (2003), Boers *et al.* (2004), González Rey (2005, 2010), Boers *et al.* (2007), Granger & Paquot (2008), Maribel (2014).

³⁶ Nous reprenons ici le classement de phrasèmes proposé par Mel'čuk (2013). L'auteur distingue des locutions fortes, des semi-locutions et des locutions faibles. Les conventions typographiques adoptées sont également celles proposées par

prudent de ne pas aborder tous les types d'unités lexicales et phraséologiques à la fois. Il n'en reste pas moins que la question de compétences métalinguistiques que notre test a mise en avant, notamment dans leur dimension sémantique, mérite d'être posée pour l'ensemble des enseignements lexicaux.

Dans la section qui suit, nous présentons le cadre général et l'architecture du test réalisé pour en discuter les résultats en 3. La quatrième et dernière partie réunira nos conclusions.

Cadre général de la recherche et de l'élaboration du test

Images dans la langue : programme ICal

Le questionnaire fut réalisé dans le cadre du projet de recherche *Images culturelles dans l'apprentissage des langues : enjeux didactiques et linguistiques* (ICal), actuellement mené à l'Université Littoral Côte d'Opale sous la responsabilité d'Émilie Perrichon³⁷ et Dorota Sikora³⁸, en collaboration avec des chercheurs, linguistes et didacticiens, de l'équipe *Grammatica* de l'Université d'Artois (Arras, France). Il s'agit d'une réflexion interdisciplinaire sur l'image culturelle dans l'enseignement des langues étrangères, avec à la clé des propositions pédagogiques à appliquer en classe. Dans son volet linguistique, l'objectif poursuivi est de proposer une description linguistique cohérente et exhaustive des locutions pour des applications pédagogiques.

La notion d'*image culturelle*, retenue dans un premier temps sur des bases intuitives, offre un champ de référence suffisamment vaste pour réunir des études relevant de plusieurs disciplines. Cependant, appliquée aux locutions, elle doit être circonscrite avec plus de précision, condition *sine qua non* des traitements pédagogiques qu'on en envisagera par la suite. Nous nous sommes ainsi appuyées sur les réflexions de Dobrovol'skij (2016) et le concept de *composante imagée* (*image component*) présente dans le sémantisme de certaines locutions : il s'agit de l'ensemble de traces linguistiquement pertinentes de l'image sous-jacente. La pertinence linguistique de la composante imagée se manifeste entre autres dans la combinatoire lexicale de la locution, notamment dans ses collocatifs³⁹. C'est elle également qui est à l'origine de contraintes pesant quelquefois sur l'emploi de la lexie locutionnelle.

En concevant le test à soumettre aux futurs enseignants, nous avons donc adopté l'hypothèse selon laquelle le sens locutionnel réunit, de manière structurée, deux types de contenus :

- dénotatif, qui permet d'en identifier l'ensemble de référents et qui se laisse expliciter dans une paraphrase définitoire. Il conduit à tisser un réseau de relations paradigmatiques (synonymie, antonymie, etc.) à l'intérieur d'un champ sémantique ;
- imagé ou figuratif qui repose sur la composante imagée identifiée par Dobrovol'skij.

Présentation du test

Un test conçu dans le cadre d'une recherche peut servir au moins deux types d'objectifs. Il est parfois nécessaire de confirmer, d'infirmer, voire de nuancer une hypothèse adoptée par le chercheur. Il est dès lors possible de poser des questions en rapport direct avec celle-ci. Mais une enquête peut avoir comme but de ramener des informations qui permettront par la suite de construire des hypothèses. Les questions et tâches à réaliser prennent alors inévitablement une forme ouverte, de manière à ne pas écarter d'office les éléments de réponse non anticipés. Le prix à payer en est l'impossibilité de présenter des analyses quantifiées d'un point de vue statistique.

Objectifs du questionnaire et résultats à discuter

L'objectif général de notre questionnaire a été de réunir des informations sur la nature de la composante imagée et de l'ensemble des contenus figuratifs du sens locutionnel. Sont-ils clairement identifiés par les locuteurs et si oui, se limitent-ils aux traces linguistiquement pertinentes telles que la combina-

Mel'čuk. Elles sont destinées à traduire le statut lexical différent des phrasèmes. Faute de place, nous ne pouvons que renvoyer le lecteur au texte source.

³⁷ Enseignante-chercheuse à l'ULCO, spécialiste de didactique de langues au sein du Laboratoire HLLI (EA 4030).

³⁸ Enseignante-chercheuse à l'ULCO, spécialiste de sémantique lexicale au sein du Laboratoire HLLI (EA 4030).

³⁹ Voir, à titre d'exemple, le cas de PEAU DE BANANE discuté dans Sikora (2018).

toire ou les restrictions d'emploi ? Jouent-ils un rôle dans le travail d'analyse qu'un locuteur est amené à effectuer, lorsqu'il tombe sur un phrasème non-compositionnel inconnu ? Il est important de souligner que le questionnaire n'avait pas pour objectif de vérifier l'aisance dans la présentation des informations lexicales que des étudiants se destinant aux métiers d'enseignement seront susceptibles de présenter aux apprenants. La question des connaissances dont ils disposent et la façon dont celles-ci s'expriment (par une paraphrase ou par association avec d'autres lexies) a émergé face aux réponses obtenues, en écartant momentanément celle des contenus figuratifs. Autrement dit, les éléments fournis par le test et discutés ici constituent un apport inattendu, une valeur ajoutée en quelque sorte, qui mérite, à notre avis, d'être prise en compte dans le contexte plus large de la formation de futurs professeurs.

Choix de locutions

Le questionnaire a été soumis à deux groupes d'étudiants, locuteurs natifs du français d'une part, et de l'autre, apprenants FLE de niveau B2 du *Cadre européen commun de référence pour les langues* (voir également North & Goodier, 2018), inscrits dans des formations qui préparent à l'enseignement de cette langue soit à un public scolaire en France, soit à des apprenants étrangers. Au total, 62 participants ont répondu au test. Compte tenu de nos objectifs et du degré de maîtrise particulier à chacun des répondants, le choix des locutions pour le questionnaire était soumis à de nombreuses contraintes.

1. Il a fallu tenir compte du temps dont nous disposions : les réponses devaient être fournies en une séance de deux heures.
2. Étant donné le niveau de compétences B2 des participants apprenants, il fallait éviter des locutions dont l'emploi est trop complexe à cette étape de l'apprentissage. Ainsi, nous avons opté pour des locutions nominales et verbales, employées dans des fonctions syntaxiques traditionnellement réservées à ces parties du discours. Ainsi, avons-nous retenu *PRINCE CHARMANT* puisqu'on trouve ce phrasème en position de sujet ou de complément, en renonçant à *FLEUR BLEUE*, car toute nominale qu'elle est, elle s'emploie comme modificateur adjectival (pour une discussion du décalage entre la partie du discours de surface et le type d'emplois de certaines locutions, voir Sikora, 2018).
3. Compte tenu de notre objectif initial, qui était de vérifier le rôle de la composante imagée dans l'analyse et les stratégies de devinement⁴⁰ d'une part (cf. étape 1 du test), et de l'autre, les éventuelles différences entre les contenus figuratifs – par exemple entre *NOYER LE POISSON* et *BOTTER EN TOUCHE* (voir l'étape 3 du test ci-dessous) –, il a été nécessaire de trouver un équilibre entre des locutions potentiellement connues et inconnues. Le critère de fréquence que nous avons mis à contribution à ce stade n'en réduit pas néanmoins le caractère aléatoire. De manière générale, les locutions se caractérisent par une faible fréquence en corpus (Colson 2008). Dans la liste retenue, certaines ont une fréquence relativement élevée, notamment dans Frantext et dans le corpus français du web (FRWac) : c'est le cas de *PRINCE CHARMANT*⁴¹ et de *NOYER LE POISSON*. D'autres – *MARIAGE DE LA CARPE ET DU LAPIN*, *GARDER UN CHIEN DE SA CHIENNE* – plus rarement attestés, avaient été choisis pour vérifier les connaissances et mobiliser les compétences personnelles en vue de la seconde étape du test. Elles étaient destinées à déclencher des procédés d'analyse à la recherche du sens, quand celui-ci était inconnu.
4. Nous avons visé des locutions avec un réseau de connexions paradigmatiques, car celles-ci devaient servir deux objectifs. D'une part, compte tenu des difficultés que présente la définition du sens dénotatif, nous voulions offrir aux enquêtés la possibilité de le saisir à l'aide d'un quasi-synonyme. D'autre part, une comparaison avec des lexies quasi-synonymes pourrait faire apparaître plus nettement leurs propriétés sémantiques, tout particulièrement figuratives (cf. étape 3 décrite ci-dessous).

⁴⁰ Le terme de *devinement* désigne un type d'inférence réalisée par l'apprenant face à un élément inconnu (voir par exemple Defays et Deltour 2003).

⁴¹ Une fréquence élevée pour une locution reste néanmoins faible. Dans FrWac, avec 1212 occurrences, *PRINCE CHARMANT* a une fréquence relative de $7,5 \times 10^{-7}$.

Finalement, les locutions suivantes ont été sélectionnées : MARIAGE DE LA CARPE ET DU LAPIN, TEMPÊTE DANS UN VERRE D'EAU, VOIE DE GARAGE, CHEVALIER SERVANT, PRINCE CHARMANT, AVOIR LES DEUX PIEDS DANS LE MÊME SABOT, NOYER LE POISSON, TOURNER AUTOUR DU POT, COURIR DEUX LIÈVRES À LA FOIS, GARDER UN CHIEN DE SA CHIENNE, SE NOYER DANS UN VERRE D'EAU, DANSER SUR UN VOLCAN, JOUER AVEC LE FEU.

Architecture et déroulement du test

Le test comprenait trois étapes, chacune se focalisant sur l'un des trois types d'informations.

Étape 1. Les locutions ont été présentées aux participants de manière décontextualisée pour vérifier si le sens en était connu et si oui, comment il était appréhendé. Il était demandé de l'explicitier soit sous forme d'une définition (analytique ou phrastique), soit par des quasi-synonymes. Les participants précisaient explicitement s'ils connaissaient la locution (« Je sais ») ou s'ils en devinaient le sens (« Je devine »).

Étape 2. Un corpus de quinze exemples illustrant l'emploi de chacune des treize locutions était destiné à vérifier l'apport d'indices cotextuels à l'analyse de cas inconnus. Si celui-ci s'avérait utile, un retour à l'étape 1 était possible.

Étape 3. Dans huit extraits de corpus, les participants ont eu à remplacer la locution par une lexie quasi-synonyme (TEMPÊTE DANS UN VERRE D'EAU par DÉSACCORD, DANSER SUR UN VOLCAN par JOUER AVEC LE FEU) ou par un syntagme libre (COURIR DEUX LIÈVRES À LA FOIS par *s'occuper de plusieurs choses à la fois*). L'encadré présente la consigne et un exemple de corpus à traiter.

Dans les textes, essayez de remplacer la locution par un simple lexème quasi-synonyme. Est-ce que cette modification change le texte ? Comment pourriez-vous le décrire et/ou l'expliquer ?

TEMPÊTE DANS UN VERRE D'EAU → DÉSACCORD

Mais j'avais tenu au courant votre état-major impérial. Nous avons une petite tempête dans un verre d'eau. Cela va s'apaiser très vite.

Ch. de Gaulle, *Mémoires de guerre : t. 1 : L'Appel (1940-1942)*, 1954, p. 353

Ce jeu de substitution, dont les résultats étaient à décrire dans un très court commentaire, devait permettre d'identifier les différences de sens et de contenus figuratifs, véhiculés notamment par les locutions.

Résultats : focus sur les compétences métalinguistiques en langue maternelle

Nous limitons la discussion aux résultats retournés par les 35 participants de langue maternelle française. Plusieurs raisons sont à l'origine de ce choix. Tout d'abord, en étudiant les réponses, nous avons constaté que pour les participants-locuteurs non-natifs, des informations supplémentaires devraient être prises en compte dans une version affinée du test. Il est apparu par exemple qu'ils mettaient souvent en correspondance les locutions du test et celles de leurs langues maternelles respectives. Or, il s'agit là d'une différence importante entre la démarche des locuteurs natifs de français et celle des apprenants étrangers. Tel qu'il a été réalisé, notre questionnaire n'a pas permis d'apporter assez d'informations sur les stratégies de devinement appliquées dans les deux cas.

Ensuite, s'il est intéressant d'analyser les stratégies de devinement mises en œuvre, il semble problématique de se prononcer sur les compétences métalinguistiques, surtout quand il s'agit d'apprenants, même d'un niveau avancé : des recherches (Bialystok 1991, Pajooesh 2014) montrent que celles-ci sont à distinguer de la maîtrise de la langue générale. De plus, le fait de les posséder dans et pour sa langue maternelle ne signifie pas que l'on soit à même de les pratiquer dans une langue en cours d'apprentissage. Par ailleurs, les compétences productives en B2 n'incluent pas encore la formulation de définition et d'explication précises. Ainsi, il nous est possible – à partir des réponses fournies – de saisir les connaissances relatives à telle ou telle locution ou bien à la stratégie appliquée pour l'analyser, mais

une certaine prudence s'impose dans les généralisations concernant les compétences métalinguistiques des répondants.

Discussion des résultats

Locutions connues et inconnues

Les locutions étudiées sont très inégalement connues des participants. 20 % des réponses signalaient comme familier le sens de dix phrasèmes, la moyenne étant située à cinq des treize. Néanmoins, le sens défini pour certains cas ne correspondait pas nécessairement à celui qui est réellement lexicalisé. Par exemple, l'explication de VOIE DE GARAGE, au sens de situation compliquée, insoluble, laisse penser que le sémantisme de la locution reste peu cerné. On note également un recours très différencié aux stratégies de devinement. 14 participants (soit 40 %) les ont systématiquement appliquées à tous les cas inconnus, alors que d'autres (11 %) ne s'y sont pas lancés du tout ou bien très peu (pour une ou deux locution(s)). Les raisons de ces postures différentes seront intéressantes à explorer dans une recherche future.

PRINCE CHARMANT, NOYER LE POISSON, TOURNER AUTOUR DU POT et JOUER AVEC LE FEU se situant parmi les locutions les plus attestées, apparaissent également comme mieux connues, avec plus de paraphrases à la fois claires et adéquates. Tel n'est pas le cas de TEMPÊTE DANS UN VERRE D'EAU ou AVOIR LES DEUX PIEDS DANS LE MÊME SABOT, dont respectivement trois et deux explications de sens correspondent à celui qui est réellement lexicalisé. Dans une majorité des cas, soit ni définition, ni quasi-synonyme ne sont proposés, soit la paraphrase est issue de procédés d'analyse (*ipse eo* relève des stratégies de devinement) aboutissant à un résultat inexact – elle est alors marquée de l'étiquette « Je devine ». Une analyse de TEMPÊTE DANS UN VERRE D'EAU a conduit l'un des participants au sens « mauvaise ambiance qui règne dans un endroit donné », alors que pour AVOIR LES DEUX PIEDS DANS LE MÊME SABOT, plusieurs hypothèses sémantiques furent formulées : « être bête », « ne pas être capable d'effectuer correctement un mouvement », « être coincé dans un domaine » ou encore « ne pas pouvoir faire de choix », « être bloqué ».

La diversité des sens inférés est inhérente à toute procédure d'analyse qu'entreprend un usager de la langue face à une locution inconnue, puisque – comme l'observe Dobrovolskij (2016), l'image véhiculée peut conduire à plus d'un sens à envisager⁴². Le risque d'orienter l'analyse vers une interprétation qui ne correspond pas au sens dénotatif de la locution est donc inévitable, mais la procédure elle-même présente un intérêt incontestable. Elle constitue ce que Cuq (2004) qualifie de *tâche significative* : celle qui par l'implication du locuteur ou de l'apprenant dans la réflexion sur ses propres connaissances, laisse une trace mémorielle plus importante qu'une simple consultation de dictionnaire.

Cependant, il est parfois difficile d'évaluer tant les connaissances lexicales des répondants que les résultats de l'analyse de la locution en raison des formulations qui en ont été adoptées, c'est-à-dire en raison de la forme de restitution.

Forme de la restitution des résultats

Rappelons qu'afin de ne pas limiter les réponses, deux modes d'explication du sens pouvaient être adoptés : une définition (analytique ou phrastique) et des quasi-synonymes lexémiques ou locutionnels. Tous les participants ont alterné les deux moyens, aussi bien pour les locutions considérées comme connues que pour celles soumises à une analyse.

En ce qui concerne les définitions proposées, il est intéressant de noter que plusieurs d'entre elles, sans être inexacts à proprement parler, conduisent à poser la question du degré de maîtrise de la lexie locutionnelle. Nous venons de citer le cas de VOIE DE GARAGE paraphrasé par *être dans une situation qui semble compliquée, insoluble*. Ailleurs, celui de TOURNER AUTOUR DU POT est défini par « ne pas dire la vérité ». Le test étant anonyme, dans le souci notamment de ne pas le faire apparaître comme une forme de contrôle de connaissances, nous n'avons pas pu interroger les participants plus à même. Face

⁴² Dans Sikora (2018), nous analysons le cas de COUPER LES CHEVEUX EN QUATRE dont l'analyse conduit à deux sens : l'un pourrait être paraphrasé par « effectuer un travail de manière soignée et attentive » et l'autre – réellement lexicalisé – par « étudier un problème avec une attention inutile aux détails sans importance ».

à ces réponses, deux explications ont été envisagées. Notons qu'elles ne s'excluent guère et peuvent être vraies en même temps.

Premièrement, le sens de ces locutions (et de celles qui avaient reçu des descriptions similaires) est saisi de manière très générale : *TOURNER AUTOUR DU POT* est simplement inscrit dans le champ sémantique de la communication, associé à son synonyme moins riche *DIRE*. Pour *VOIE DE GARAGE*, les liens avec d'autres lexies semblent encore plus ténus, au point qu'il est difficile de préciser dans quel champ sémantique la réponse le situe. Le caractère aléatoire de ces connexions se manifeste également, lorsque le sens est caractérisé à l'aide de quasi-synonymes, souvent appartenant à des parties du discours différentes de celle de la locution. Pour *CHEVALIER SERVANT*, on propose « aider, rendre service », pour *MARIAGE DE LA CARPE ET DU LAPIN* – « différent ». Si l'on admet avec Meara et Wolter (2004) et Meara (2009) que la maîtrise d'un *mot* se mesure au nombre et à la qualité des relations qui l'attachent aux unités lexicales déjà acquises, on peut considérer que ces locutions sont simplement peu intégrées dans le vocabulaire des locuteurs.

Deuxièmement, de nombreuses recherches (Vygotsky 1986, Pajoesh 2014) montrent que, même lorsqu'une unité lexicale est connue et couramment employée par le locuteur, celui-ci n'a pas nécessairement la compétence métalinguistique nécessaire pour en formuler une définition et pour identifier avec précision les liens paradigmatiques déterminants pour sa place dans le lexique d'une langue.

Les compétences métalinguistiques, que ce soit dans une langue maternelle ou étrangère, doivent être clairement distinguées de celles qui permettent simplement à un locuteur de se servir d'une langue (Bialystok 1991). Elles reposent sur une certaine représentation des savoirs concernant le système et son fonctionnement, ainsi que sur des opérations cognitives permettant constamment d'améliorer cette modélisation, en la restructurant, en la rendant plus explicite et plus solidement interconnectée. L'analyse de ses propres connaissances linguistiques, en l'occurrence lexicales, et les activités consistant à les expliciter, par exemple sous forme de définitions, constituent une étape importante dans l'apprentissage d'une L2 (cf. Sikora, 2016), certes, mais également dans son enseignement.

Conclusions

Notre test a mis en avant la nécessité d'un travail sur les connaissances lexicales que de futurs enseignants seront certainement amenés à transmettre. Certes, ces résultats pourront être consolidés et nuancés auprès d'un panel de participants plus large. De nombreux indices entraperçus seulement, par exemple, la possible mise à contribution des connaissances en L1 chez les apprenants, méritent qu'on s'y penche avec attention.

Lors de notre expérimentation, de futurs enseignants découvraient pas à pas les différentes facettes d'un échantillon de locutions dans leur langue maternelle. Il s'avère, du moins pour l'ensemble retenu, que leur connaissance reste incertaine : ces lexies semblent en plus faiblement ancrées dans le vocabulaire. Mais les définitions proposées dans les réponses, notamment lorsqu'elles manquent de précision (par exemple, lorsque l'on explique *CHEVALIER SERVANT* par « aider, rendre service »), indiquent que le fait de connaître une UP n'est qu'un élément à prendre en compte. Il y va également des compétences métalinguistiques des locuteurs. Celles-ci s'appuient, certes, sur la maîtrise des lexies locutionnelles, mais également sur une organisation consciente des informations (sémantiques, grammaticales, pragmatiques) les concernant. D'ailleurs, ce besoin de conscientisation et de (re)restructuration de connaissances lexicales, puis celui d'entraînement à leur explicitation se sont clairement manifestés dans les réponses et dans les discussions qui les ont suivies.

Références bibliographiques

- Abel, B. (2003). English idioms in the first language and second language lexicon: a dual representation approach. *Second Language Research*, 19(4), 329-358.
- Bialystok, E. (1991). Metalinguistic Dimensions of Bilingual Language Proficiency. In E. Bialystok (ed.) *Language Processing in Bilingual Children* (p. 113-40). Cambridge : Cambridge University Press.

- Boers, F. (2001). Remembering figurative idioms by hypothesising about their origins. *Prospect*, 16(3), 3543.
- Boers, F., Demecheleer, M. & J. Eyckmans. (2004). Etymological elaboration as a strategy for learning idioms. In P. Bogaards & B. Laufer (eds) *Vocabulary in a Second Language Selection, Acquisition and Testing* (p. 5378). Amsterdam : John Benjamins.
- Boers, F., Eyckmans, J. & H. Strengers. (2007). Presenting figurative idioms with a touch of etymology - More than mere mnemonics? *Language Teaching Research*, 11(1), 43-62.
- Colson, J.-P. (2008). Cross-Linguistic Phraseological Studies—An Overview. In S. Granger & F. Meunier (eds) *Phraseology* (p. 191-206). Amsterdam, Philadelphia : John Benjamins Publishing Company.
- Cuq, J.-P. (2004). Le lexique en situation d'apprentissage guidé. Pour une méthodologie d'enseignement interventionniste dans l'enseignement du français langue étrangère. In É. Calaque & J. David (éds) *Didactique du lexique* (p. 6171). Bruxelles : De Boeck.
- Defays, J.-M. & Deltour, S. (2003). *Le français langue étrangère et seconde. Enseignement et apprentissage*. Wavre : Mardaga.
- Dobrovolskij, D. (2016). The notion of “inner form” and idiom semantics. *Études et Travaux*, 1(1), 2136.
- Ettinger, S. (1998). Einige Überlegungen zur Phraseodidaktik. In W. Eismann (ed.) *Studien zur Phraseologie und parömiologie* (p. 201217). Bochum : Universitätsverlag Dr. N. Brockmeyer.
- González Rey, M. I. (2005). L'espace réservé à la phraséologie dans la didactique du F.L.E. In Á. S. Ramos (éd.) *Espacio y Texto en la Cultura Francesa. Espace et Texte dans la Culture Française* (p. 14211439). Alicante : Universidad de Alicante,.
- González Rey, M. I. (2010). *La Phraséodidactique en action : Les expressions figées comme objet d'enseignement*. Lyon : ENS LSH/DGESCO.
- Granger, S. & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger et F. Meunier (eds) *Phraseology* (p.27-50). John Benjamins Publishing Company,.
- Gréciano, G. (1984). Pour un apprentissage des unités phraséologiques. *Nouveaux Cahiers d'Allemand*, 2, 95113.
- González, M. (2014). *Outils et méthodes d'apprentissage en phraséodidactique*. Éditions modulaires européennes. Fernelmont : EME Éditions.
- Meara, P. (2009). *Connected Words*. Amsterdam and Philadelphia: John Benjamins.
- Meara, P. & B. Wolter. (2004). V-Inks : Beyond Vocabulary Depth. *Angles on the English Speaking World*, 4, 8597.
- Mel'čuk, I. A. (2013). Tout ce que nous voulions savoir sur les phrasèmes, mais... *Cahiers de Lexicologie*, 102 (1), 12949.
- North, B. & T. Goodier. (2018). Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer. Volume complémentaire avec de nouveaux descripteurs. Starbourg : Conseil de l'Europe. rm.coe.int/cecr-volume-complementaire-avec-de-nouveaux-descripteurs/16807875d5 (consulté le 18 juin 2018).
- Pajoosh, P. (2014). Knowledge of Word Concepts and Referents. In J. Milton et T. Fitzpatrick (eds) *Dimensions of Vocabulary Knowledge* (p. 7391). Palgrave : Macmillan.
- Sikora, D. (2016). Activité(s) de définition dans l'apprentissage d'une langue seconde. *Cahiers de Lexicologie*, 2 (109), 11743.
- Sikora, D. (2018). Locutions en apprentissage de langue seconde. In SHS Web of Conferences (Vol. 46, p. 05002). EDP Sciences.
- Vygotsky, L. S. (1986). *Thought and Language*. Cambridge, Mass : MIT Press.

DE QUELQUES ASPECTS DE COMPÉTENCE PHRASÉOLOGIQUE EN L2: LOCUTIONS EN TEST

Dorota Sikora
Université du Littoral Côte d'Opale

Alexandra Tsedryk
Mount Saint Vincent University

RÉSUMÉ

Les apprenants de français langue seconde maîtrisent-ils des locutions, c'est-à-dire des unités phraséologiques au sens non-compositionnel? Quelle(s) démarche(s) adoptent-ils face à une locution inconnue? Cherchent-ils à en découvrir le sens et si, oui quelles stratégies adoptent-ils pour y parvenir? Faut-il des compétences spécifiques pour apprendre des unités phraséologiques, en particulier des locutions? L'article explore les réponses à un questionnaire construit autour de treize locutions françaises, soumis à des apprenants de niveaux de compétences B1 et B2, que le *Cadre européen commun de référence pour les langues* qualifie d'*utilisateurs indépendants*. L'analyse des résultats montre que les connaissances de vocabulaire locutionnel sont aléatoires et que, face à des unités inconnues, les apprenants mettent en place peu de stratégies susceptibles de les conduire au sens. Le rôle de la phraséodidactique dans l'enseignement de vocabulaire et dans le développement de compétences indispensables dans les apprentissages apparaît ainsi comme essentiel dans le cheminement vers l'autonomie.

Mots clés: locutions, phraséologie, phraséodidactique, compétences phraséologiques, stratégies d'apprentissage

1. INTRODUCTION

Depuis une quarantaine d'années, les réflexions sur l'apprentissage et l'enseignement des unités phraséologiques (désormais UP) dans un cursus de langue seconde (L2) convergent vers un consensus : de par leurs spécificités syntaxiques, sémantiques et pragmatiques, les formules polylexicales exigent une démarche didactique appropriée. Celle-ci ne peut se construire sans prendre en compte les caractéristiques des processus d'apprentissage de l'étudiant d'une part, et les propriétés des UP elles-mêmes de l'autre.

Notre article présente les résultats d'un test réalisé auprès d'un groupe d'apprenants de français langue seconde de niveau intermédiaire-avancé à l'Université Mount Saint Vincent à Halifax (Nouvelle Écosse, Canada). Treize locutions leur ont été soumises, avec deux tâches à effectuer, afin de dégager des éléments de réponses aux questions suivantes : (1) Quel est le degré de connaissance du vocabulaire locutionnel relativement fréquent (comme *PRINCE CHARMANT*) ou non (*GARDER UN CHIEN DE SA CHIENNE*)? (2) Lorsqu'une locution est inconnue, quelle est la démarche des apprenants ? Que font-ils pour en découvrir le sens? (3) Comment transposent-ils les informations sémantiques ou les hypothèses qu'ils forment dans une paraphrase définitoire, dans une connexion paradigmatique ou dans un exemple d'emploi ?

Nous commencerons, dans la section 2, par circonscrire le domaine de la phraséodidactique, préciser son objet et réfléchir sur les compétences qui permettent à l'apprenant de traiter les UP. La partie suivante sera consacrée à la présentation du test, à son architecture et aux conditions de sa réalisation. Les résultats quantitatifs et qualitatifs seront exposés dans la section 4, puis discutés en 5. Dans les Conclusions que nous en tirons (section 6), nous plaiderons pour une phraséodidactique explicite et structurée, capable de pallier les difficultés des apprenants face notamment au vocabulaire locutionnel.

2. PHRASÉODIDACTIQUE ET COMPÉTENCES PHRASÉOLOGIQUES

Bien que évoquée par Bally (1909), la phraséodidactique est un courant de recherches relativement récent. C'est seulement depuis les années 80 du XX^e siècle, avec les travaux de linguistes allemands, et notamment ceux de Gréciano (1984), que l'idée d'un module didactique spécifique, visant l'apprentissage des unités phraséologiques a inspiré quantité d'études, essentiellement de nature expérimentale. Elles ont en commun l'hypothèse selon laquelle, en L2, l'apprentissage des UP nécessite un enseignement approprié. Autrement dit, on admet qu'il existe au sein du lexique d'une langue, des unités polylexicales dont le traitement cognitif et, par conséquent, la démarche pédagogique qui l'accompagne, ne sont pas les mêmes que ceux qu'on doit employer pour le vocabulaire lexémique. Il s'agit dès lors, de la part des enseignants, de proposer un ensemble d'activités qui, en se focalisant sur les UP, seront susceptibles d'en faciliter l'acquisition.

S'il est généralement admis que les UP reçoivent un traitement holistique de la part d'un locuteur natif, la démarche d'un apprenant semble souvent s'organiser différemment (cf. à titre d'exemple Siyanova-Chanturia et al. 2011). Compte tenu de la complexité de ces phénomènes, les études expérimentales sont inévitablement conduites à privilégier un angle d'approche. Ainsi, l'acquisition des UP a été étudiée entre autres selon les axes tels que :

- traitement des UP par des locuteurs natifs et par les apprenants (Abel 2003 ; Bortfeld 2012 ; Caillies 2009 ; Edmonds 2013) ;
- intégration des UP et le développement de connaissances lexicales (Sulowska 2013, 2018) ;

- contribution des UP, repérées notamment sur des bases fréquentielles en tant qu'unités pré-fabriquées, au développement des compétences réceptives et des compétences productives (Stengers et al. 2011), dont la fluence (Théophanous & Perez-Bettan 2018) ;
- démarches pédagogiques et activités à proposer pour enseigner, de manière efficace, les UP (Boers 2001 ; Boers et al. 2004 ; Boers et al. 2007 ; González Rey 2007, 2010, 2018).

Contrairement au courant sociolinguistique actuellement dominant en didactique, qui se manifeste notamment dans l'approche actionnelle du *Cadre européen commun de référence pour les langues* (CECRL), la phraséodidactique se rapproche naturellement des recherches linguistiques réalisées en phraséologie, s'appuyant sur l'exploration de corpus. Elle en contracte par ailleurs les difficultés, par exemple quand il s'agit de classer les UP, dont la structure polylexicale et le fonctionnement multidimensionnel (formel, sémantique, syntaxique, discursif, etc.) rend toute démarche typologique particulièrement délicate. En effet, les multiples caractéristiques des UP conduisent à des classements, et chemin faisant à des terminologies, très variés. Ainsi, on parlera des *séquences formulaires* (cf. *formulaic sequences* chez Wray 2002), de séquences préfabriquées pour désigner des segments aux degrés de figement différents quelle que soit leur nature. En lexicographie (Mel'čuk 2013), on prend en compte au contraire leur nature et les types de contraintes sémantiques, lexicales et syntaxiques qu'elles subissent. Une perspective discursive conduit à adopter des termes tels que *routines conversationnelles* (Coulmas 1981 ; Groupe de Fribourg 2012) ou *actes stéréotypés de langage* (Kauffer 2011).

L'objectif d'une démarche phraséodidactique est double. Premièrement, il consiste à développer des connaissances lexicales de l'apprenant du point de vue quantitatif, en y intégrant de nouvelles UP. Autrement dit, il s'agit de l'aider à en maîtriser de plus en plus, en les intégrant dans le vocabulaire à la fois réceptif et productif (Sikora 2018a). Or, d'une part, ni le volume horaire destiné à un parcours d'apprentissage, ni le contenu des programmes ne permettent en général d'accorder aux UP l'attention didactique qu'elles méritent (Théophanous 2013 ; Tsedryk 2018). D'autre part, leur multitude et leur vivacité – pour des raisons diverses, elles sont nombreuses à apparaître et à disparaître dans les échanges quotidiens – rend difficile la prise en compte ne serait-ce que d'une partie significative en classe de langue. D'où la nécessité de définir un second objectif phraséodidactique : aider l'apprenant à construire des compétences sur lesquelles il pourra s'appuyer pour traiter, intégrer dans son vocabulaire et employer en discours des UP jusque là inconnues, rencontrées souvent de manière aléatoire, au gré de ses contacts de la langue cible. Ce second objectif est donc qualitatif.

Qu'entend-on alors par compétences phraséologiques? De manière générale, une compétence est définie comme un savoir-faire qui conduit l'apprenant à faire un usage approprié des connaissances et de l'ensemble des ressources qu'il possède pour comprendre le sens, repérer les caractéristiques sémantiques et formelles, identifier le fonctionnement discursif d'une UP pour l'intégrer dans son vocabulaire productif.

Si les connaissances phraséologiques se développent tant du point de vue quantitatif que qualitatif pendant un cursus d'apprentissage (cf. Sułkowska 2018), la multitude et la diversité des UP obligera toujours l'apprenant à en traiter de nouvelles, inconnues jusque là. Il sera ainsi amené à s'appuyer non seulement sur les connaissances déjà acquises, mais également sur des straté-

gies appropriées. Les stratégies d'apprentissage des UP s'inscrivent dans la dimension qualitative des compétences phraséologiques.

Ainsi, le développement de stratégies nécessaires pour maîtriser les UP, c'est-à-dire des procédures et des outils mis à contribution de manière consciente pour améliorer les apprentissages, constitue l'un des principaux objectifs de la phraséodidactique. Parmi les stratégies qui font partie intégrante des compétences phraséologiques (Sikora 2018a), figurent notamment des :

- stratégies cognitives, qui consistent à analyser l'expression, comparer avec les langues connues, cf. (Oxford 1990) ;
- stratégies cognitives de premier contact (*initial handling strategies*, Yongqi 2010) telles que recours au dictionnaire (simple ou étendu), devinement à partir d'indices linguistiques ;
- stratégies métacognitives (Oxford 1990), qui consistent à examiner la matière à apprendre, découvrir le fonctionnement ;
- stratégies mnémoniques (Oxford 1990) qui reposent sur le classement, les associations, c'est-à-dire sur la formation des liens et le positionnement par rapport aux connaissances déjà acquises, utilisation des images, des sons, etc.

3. PRÉSENTATION DU TEST ET DE SON ENVIRONNEMENT

L'étude dont cet article présente les résultats partiels s'inscrit dans le cadre plus large du programme de recherche *Images culturelles dans l'apprentissage des langues : enjeux didactiques et linguistiques* (ICal) développé à l'Université du Littoral Côte d'Opale (France) en collaboration avec l'Université Mount Saint-Vincent à Halifax (Canada). Ce projet, qui réunit les spécialistes en linguistique et en didactique des langues propose d'explorer les images culturelles et leur rôle dans l'enseignement et l'apprentissage des langues, notamment lorsque la démarche pédagogique se donne pour objectif de former un locuteur plurilingue.

Dans son volet linguistique, *ICal* se focalise sur les locutions telles qu'elles sont définies par (Mel'čuk 2013) : phrasèmes non-compositionnels qui appartiennent au module lexical d'une langue, en l'occurrence du français. Les locutions telles que *CASSER LA CROÛTE* sont ainsi identifiées en tant qu'unités lexicales (lexies) au même titre que les lexèmes, par exemple *MANGER I.1*. Pour la phraséodidactique, elles sont particulièrement intéressantes en raison des contenus figuratifs (Dobrovol'skij & Piirainen 2005 ; Piirainen 2008 ; Dobrovol'skij 2016) qu'elles sont nombreuses à véhiculer. Or, si cette particularité de leur sémantisme représente une difficulté dans l'apprentissage (Sikora 2018a), elle peut également être exploitée en phraséodidactique (voir, à titre d'exemple, les propositions de Boers (2001), Abel (2003), Boers et al. (2007)).

Le questionnaire dont nous présentons et discutons ici des résultats partiels a été conçu tant à destination de locuteurs natifs du français que d'apprenants de L2. Son objectif général était de réunir le plus d'éléments de réponse à la question de savoir comment sont traités les contenus figuratifs par ces deux publics. Le test possède une formule ouverte dans la mesure où il n'est pas destiné à confirmer ou à infirmer une hypothèse de travail, mais d'apporter des données pour en construire. Treize locutions ont été retenues (voir les critères du choix ci-dessous), nombre relativement modeste, mais permettant aux participants de traiter l'ensemble des tâches définies.

3.1 Architecture globale du test

L'ensemble du questionnaire organisait le travail demandé en trois étapes :

1. Dans un premier temps, les 13 locutions ont été présentées de manière entièrement décontextualisée afin de vérifier si elles faisaient partie du vocabulaire des participants. Si ce n'était pas le cas, les participants ont été encouragés à les soumettre à l'analyse pour en deviner le sens à partir de la structure lexico-syntaxique du phrasème (Pausé 2017), en se servant notamment des contenus figuratifs. Cet intérêt pour les images véhiculées (ou non) par les éléments lexicaux formant la locution et leur (éventuel) pouvoir explicatif qui a été décisif pour la décontextualisation. Nous avons voulu savoir ce qui pouvait être inféré de la structure lexico-syntaxique elle-même¹.
2. Une deuxième étape comprenait un travail sur des extraits de corpus contenant les locutions du test, de manière à vérifier s'ils fournissaient des indices pour comprendre leur sens là où il n'était pas connu. Cette tâche réalisée en France, auprès de 62 étudiants de licence et de master, discutée dans (Sikora 2018b), ne sera pas abordée ici.
3. Une dernière réflexion, à présenter sous forme d'une courte explication rédigée sans contrainte formelle, s'appuyait sur un exercice de substitution. Dans des extraits de corpus, nous avons remplacé la locution qui y figurait (par exemple JOUER AVEC LE FEU) par son quasi-synonyme (DANSER SUR UN VOLCAN). Les participants étaient invités à expliquer la modification sémantique qui en découle, toujours dans le but de saisir les contenus figuratifs liés aux locutions en question. Jusqu'à présent, seuls les étudiants de français L1 ont été concernés par cette tâche : les résultats en ont été discutés dans (Sikora 2018b) et nous n'y revenons plus dans la présente étude.

L'architecture globale du test fut déterminante pour le choix des locutions, à commencer par le nombre. Il nous a fallu le limiter de manière à ce que l'ensemble soit potentiellement traitable en une heure, notamment par des participants de L1 français. Il nous a fallu équilibrer les locutions connues – permettant ainsi une réflexion de l'étape 3, et inconnues, pour inciter l'analyse, le devinement et la recherche d'indices contextuels des deux premiers points.

Dans la mesure où les participants de français L2 ont, pour la plupart, un niveau de compétences B1 et B2, nous avons ciblé des phrasèmes nominaux et verbaux, qui du point de vue de leurs fonctions syntaxiques s'emploient conformément à la nature de leur gouverneur, c'est-à-dire en tant que sujets ou compléments pour les premiers, et dans les fonctions verbales pour les seconds. C'est dire que nous n'avons pas retenu par exemple MAIN DANS LA MAIN, qui tout en étant un syntagme nominal, connaît uniquement des usages adjectivaux et adverbiaux.

C'est en fonction de cet ensemble des contraintes (compétences du public, temps de réalisation) et d'objectifs (tâches à réaliser, informations à ramener) que les 13 locutions suivantes ont été intégrées : MARIAGE DE LA CARPE ET DU LAPIN, TEMPÊTE DANS UN VERRE D'EAU, VOIE

¹ D'autant plus que les étudiants de français L2 à qui il a été proposé de réaliser l'étape 2 avaient trouvé peu, voire pas du tout d'indice contextuel dans les extraits de corpus susceptible d'illustrer l'emploi de ces locutions (cf. Sikora 2018b).

DE GARAGE, CHEVALIER SERVANT, PRINCE CHARMANT, AVOIR LES DEUX PIEDS DANS LE MÊME SABOT, NOYER LE POISSON, TOURNER AUTOUR DU POT, COURIR DEUX LIÈVRES À LA FOIS, GARDER UN CHIEN DE SA CHIENNE, SE NOYER DANS UN VERRE D'EAU, DANSER SUR UN VOLCAN, JOUER AVEC LE FEU.

3.2 CONDITIONS DE RÉALISATION DU TEST ET MÉTHODES D'ANALYSE

L'organisation modulaire de notre test permet d'en extraire telle ou telle partie pour un usage autonome, avec des questions de recherche spécifiques. C'est ainsi que nous avons procédé pour vérifier le degré de connaissance des locutions ci-dessus et les stratégies appliquées face à celles qui n'étaient pas connues des participants. Nous avons en effet extrait le document relatif à la première étape : désormais, nous désignerons ce document sous le nom de Test.

Ce Test a été proposé à 20 étudiants apprenants du français L2 à l'Université Mount Saint Vincent à Halifax au Canada, en fin d'une année universitaire pendant laquelle ils avaient suivi un cours de grammaire ou de conversation à raison de deux heures et demies par semaine. Dix-sept d'entre eux étaient issus de milieux anglophones ; trois évoluaient dans un environnement bilingue français-anglais. Tous attestaient d'un niveau intermédiaire-avancé (surtout B1 et quelques B2).

Notre objectif était de réunir les éléments de réponses aux questions suivantes : (1) Lorsque les locutions sélectionnées selon les critères ci-dessus étaient connues des participants, parvenaient-ils à en paraphraser le sens, à construire un exemple de leur emploi, à les connecter à des lexies (quasi-)synonymes? (2) Lorsque les locutions sélectionnées leur étaient inconnues, s'engageaient-ils systématiquement dans des démarches d'analyse, en appliquant notamment des stratégies de devinement pour accéder à leur sens? Nos objectifs s'articulaient en une série de questions plus précises : (a) Quel est le degré de maîtrise sémantique des locutions proposées? (b) Quels types de difficultés les apprenants ont-ils? (c) Y a-t-il un lien entre la réussite et la fréquence des locutions et/ou la langue maternelle d'apprenants ?

Nous nous sommes donc focalisées sur les stratégies cognitives (Oxford 1990), dont celles de premier contact (Yongqi 2010). Comme nous l'avons indiqué plus haut, les 13 locutions ont été présentées aux participants de manière décontextualisée. Pour chacune d'entre elles, il a été demandé de réaliser deux tâches suivantes : (a) Tâche 1 (T1) : Chaque participant devait préciser s'il en connaissait le sens (*Je sais*) ou s'il le devinait (*Je devine*). Dans les deux cas, il fallait ensuite l'expliciter sous forme de paraphrase, conformément à la consigne : *Connaissez-vous ces expressions ? Ce sont des locutions. Expliquez le sens de chacune d'entre elles* ; (b) Tâche 2 (T2) : Les participants étaient invités à construire un exemple illustrant le sens de chaque locution et en fournir un quasi-synonyme. La consigne en a été formulée de la façon suivante : *Proposez une phrase qui montre le sens de cette locution, selon vous. Ensuite, proposez son équivalent sémantique, si c'est possible*.

Afin d'éviter tout malentendu, deux exemples – TOMBER À L'EAU et MÈRE POULE – ont été fournis avec les réponses de type de celles qui étaient attendues. Les étudiants pouvaient disposer du temps qui leur était nécessaire dans les limites d'une séance de deux heures. L'usage de ressources lexicales n'a pas été permis.

Chacune des deux tâches pouvait être considérée comme *réussie*, *partiellement réussie* ou bien *non réussie*. Pour la T1, la paraphrase proposée par chacun des participants (lorsqu'elle l'était) fut comparée aux définitions extraites de trois ressources dictionnaires (cf. liste de références). De même, la réussite dans la T2 pouvait être totale (l'exemple illustre correctement le sens de la locution), partielle (l'exemple était acceptable, mais il véhiculait une ambiguïté sémantique dans la mesure où il convenait à plusieurs significations différentes dont la correcte) ou se montrer entièrement inappropriée (cf. la section 4.2 ci-dessous).

Une analyse quantitative des réponses relatives à chaque locution a ensuite été effectuée. Elle a consisté à calculer : (1) la proportion des locutions connues et de celles dont on devinait le sens pour la T1 et (2) la proportion des réussites, des réussites partielles et d'échecs pour la T1 et pour la T2.

4. RÉSULTATS

Pour contrôler la fiabilité d'analyse par deux évaluatrices, un protocole de dépouillement de données a été élaboré. Sur les 520 réponses possibles (20 participants, avec 26 réponses possibles par participant), 327 ont été réellement obtenues. Nous n'avons pas inclus dans les calculs les cases laissées vides, ni les Je ne sais pas, Aucune idée, etc. Les analyses ont été conduites de façon indépendante et ensuite comparées pour chaque jugement. Le coefficient d'objectivité est élevé, car les degrés d'entente atteignent une concordance de résultats de 80% en moyenne pour l'ensemble des réponses. Les exemples analysés différemment ont ensuite été discutés, et l'attribution des points a été décidée d'un commun accord.

4.1. RÉSULTATS QUANTITATIFS

Le Tableau 1 présente le total des réponses fournies et les pourcentages d'apprenants qui « connaissent l'expression » vs ceux qui essaient de deviner son sens. Le total des réponses possibles par locution est 40 (20 participants, deux tâches). Nous observons une grande variabilité dans le nombre de réponses pour des locutions 20 différentes, allant de 10 (pour VOIE DE GARAGE) à 38 (pour PRINCE CHARMANT). Seulement trois locutions (PRINCE CHARMANT, JOUER AVEC LE FEU et AVOIR LES DEUX PIEDS DANS LE MÊME SABOT) ont reçu un taux élevé de réponses. En moyenne, chaque locution a obtenu 25 réponses seulement, ce qui montre rapidement l'hésitation des apprenants vis-à-vis les locutions fournies.

Selon le Tableau 1, seulement deux locutions sont « connues » par la majorité des apprenants (PRINCE CHARMANT et JOUER AVEC LE FEU). Reste à savoir dans quelle mesure cette connaissance est complète, ce qui sera considéré ultérieurement, car dire qu'on connaît n'équivaut pas toujours à connaître vraiment. Les apprenants qui ont essayé de deviner le sens des locutions sont plus nombreux, ce qui suggère que pour la plupart, elles leur ne sont pas familières et qu'ils ont dû déployer les stratégies de devinement pour arriver au sens.

TABLEAU 1

Réponses fournies et le pourcentage de « je connais » vs « je devine »

| Locution | Réponses fournies (T1 + T2) | “Je connais” | “Je devine” | “Je connais” vs “Je devine” |
|---|-----------------------------------|-----------------|----------------|--------------------------------|
| PRINCE CHAR- MANT | 38 | 14 | 6 | 70% |
| JOUER AVEC LE FEU | 35 | 17 | 3 | 85% |
| AVOIR LES DEUX PIEDS DANS LE MÊME SABOT | 33 | 2 | 18 | 10% |
| NOYER LE POIS- SON | 29 | 1 | 16 | 6% |
| MARIAGE DE LA CARPE ET DU LA- PIN | 27 | 0 | 17 | 0% |
| TOURNER AU- TOUR DU POT | 26 | 4 | 11 | 27% |
| DANSER SUR UN VOLCAN | 26 | 2 | 14 | 13% |
| CHEVALIER SER- VANT | 24 | 1 | 12 | 8% |
| COURIR DEUX LIÈVRES À LA FOIS | 23 | 2 | 13 | 13% |
| TEMPÊTE DANS UN VERRE D’EAU | 20 | 1 | 12 | 8% |
| SE NOYER DANS UN VERRE D’EAU | 19 | 1 | 11 | 8% |

| | | | | |
|-------------------------------------|----|---|---|----|
| GARDER UN CHIEN DE SA CHIENNE | 17 | 0 | 7 | 0% |
| VOIE DE GARAGE | 10 | 0 | 7 | 0% |

Le Tableau 2 contient les pourcentages des taux de réussite pour chaque tâche séparément et ensuite les moyennes pour les deux tâches. Les taux de réponses non adéquates sont élevés ; en gras sont marquées les locutions pour lesquelles le taux d'échec est plus de 75%. En italiques sont notées les locutions pour lesquelles le taux d'échec est moins de 25%. Par exemple, le sens de la locution JOUER AVEC LE FEU a été bien expliqué par la totalité des participants (taux d'échec 0% pour la T1).

TABLEAU 2

La réussite et l'échec pour la T 1 (définition) et la T 2 (production)

| Locution | % de T1 (réussie/ par- tiellement réussie) | % de T1 <i>non réussie</i> | % de T2 (réussie/ par- tielle ment réussie) | % de T2 non réussie | Moyenne, échec (T1 et T2) |
|--|---|-------------------------------|---|------------------------|---------------------------------|
| PRINCE CHARMANT | 65% | 35% | 55% | 44% | 40% |
| JOUER AVEC LE FEU | 100% | 0% | 77% | 23% | 12% |
| AVOIR LES DEUX PIEDS DANS LE MÊME SABOT | 15% | 85% | 7% | 93% | 89% |
| NOYER LE POISSON | 6% | 94% | 8% | 92% | 93% |
| MARIAGE DE LA CARPE ET DU LAPIN | 65% | 35% | 67% | 33% | 34% |
| TOURNER AU- TOUR DU POT | 40% | 60% | 45% | 55% | 58% |
| DANSER SUR UN VOLCAN | 63% | 37% | 80% | 20% | 29% |

| | | | | | |
|-------------------------------------|-----|------|-----|------|------|
| CHEVALIER SERVANT | 77% | 23% | 64% | 36% | 30% |
| COURIR DEUX LIÈVRES À LA FOIS | 67% | 33% | 89% | 11% | 22% |
| TEMPÊTE DANS UN VERRE D'EAU | 23% | 77% | 0% | 100% | 89% |
| SE NOYER DANS UN VERRE D'EAU | 34% | 66% | 14% | 86% | 76% |
| GARDER UN CHIEN DE SA CHIENNE | 0% | 100% | 0% | 100% | 100% |
| VOIE DE GA- RAGE | 0% | 100% | 0% | 100% | 100% |

La moyenne des taux d'échec pour les T1 et T2 permet de constater que le sémantisme de certaines locutions a été plus transparent que d'autres. Nous les divisons en quatre groupes selon la moyenne des taux d'échec : (1) locutions « plutôt accessibles », avec le taux d'échec inférieur à 25%, (2) locutions « plus ou moins accessibles », avec le taux d'échec entre 26% et 50%, (3) les locutions « plus ou moins inaccessibles », avec le taux d'échec entre 51% et 75% et (4) les locutions « plutôt inaccessibles », avec le taux d'échec entre 76% et 100%. Le Tableau 3 est une récapitulation des résultats.

Les résultats quantitatifs laissent identifier les locutions dont le sens a été opaque : elles sont donc considérées comme « plutôt inaccessibles ». Nous voulions savoir s'il existait un lien entre la fréquence des locutions en corpus et la réussite, même si la fréquence des locutions est en général très basse². Après avoir consulté les corpus Frantext (1800-2018 : 144 264 677 tokens) et FrWac (2010 : 1 613 206 614 tokens), nous avons identifié les locutions les plus fréquentes (prince charmant, jouer avec le feu et tourner autour du pot) et les moins fréquentes (garder un chien de sa chienne³, mariage de la carpe et du lapin, danser sur un volcan, courir deux lièvres à la fois), selon les deux corpus. Deux d'entre elles ont la fréquence la plus élevée

² Par exemple, la locution « la plus fréquente » de la liste des locutions est PRINCE CHARMANT dont la fréquence relative dans le corpus FrWac est 7.5×10^{-7} , alors que celle de DANSER SUR UN VOLCAN qui est une des plus rares dans notre liste est de 8.1×10^{-9} . Cependant, il ne peut s'agir que d'orientation générale, puisque – comme le montrent justement González-Fernández et Schmitt (2015) à propos de collocation – une fréquence en corpus (surtout littéraire) ne rend pas compte de la fréquence avec laquelle l'apprenant rencontre réellement une unité en discours. Un corpus du Web tel que FrWac semble plus fiable sur ce plan.

³ Il est possible de considérer CHIEN DE SA CHIENNE comme locution, car les verbes *réserver/garder* peuvent être alternés, en formant une collocation avec CHIEN DE SA CHIENNE.

(prince charmant et jouer avec le feu). Elles figurent dans la liste de locutions « plutôt accessibles » ce qui fait penser à une influence positive de la fréquence sur l'apprentissage. Cependant, en général, la fréquence ne semble pas être le critère le plus important pour la réussite, car parmi les locutions « plutôt accessibles » et « plus ou moins accessibles » figurent en même temps des locutions fréquentes et rares, alors que la liste de « locutions plutôt inaccessibles » contient seulement une locution rare et le reste appartient au milieu de la liste. Notre hypothèse que les locutions les plus fréquentes seront les plus familières aux apprenants, tandis que les locutions les plus rares ne seront pas du tout accessibles n'est pas confirmée. D'autres facteurs influenceraient la réussite. Nous les examinerons dans la section 5 « Discussion ».

TABLEAU 3

Classement de locutions selon la transparence du sens pour les apprenants
et les correspondances anglaises

| Classement de locutions | Taux d'échec (T1+T2)/2 | Locutions (taux d'échec) | Équivalent anglais le plus proche |
|----------------------------|------------------------|--|---|
| Plutôt accessible | 0%-25% | jouer avec le feu (12%)* COURIR DEUX LIÈVRES À LA FOIS (22%) | play with fire ??? TRY TO DO TWO THINGS AT ONCE |
| Plus ou moins accessible | 26%-50% | danser sur un volcan (29%) chevalier servant (30%) mariage de la carpe et du lapin (34%) PRINCE CHARMANT (40%)* | be skating on thin ice knight in shining armor be like oil and water PRINCE CHARMING |
| Plus ou moins inaccessible | 51%-75% | TOURNER AUTOUR DU POT (58%) | TO BEAT AROUND THE BUSH |

| | | | |
|------------------------|----------|---|---|
| Plutôt inaccessible | 76%-100% | se noyer dans un verre d'eau (76%) | ???BE AS HELPLESS AS BABE |
| | | tempête dans un verre d'eau (89%) | tempest in a teapot ???lack in initiative |
| | | AVOIR LES DEUX PIEDS DANS LE MÊME SABOT (89%) | to muddy the waters ???to have a grudge against someone |
| | | noyer le poisson (93%) | DEAD END |
| | | garder un chien de sa chienne (100%) | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | VOIE DE GARAGE (100%) | |

*Locutions que les apprenants ont considérées comme « familières »

Nous avons aussi examiné le lien entre la tâche de définition et la tâche de production incluant la reformulation. Nous avons comparé les résultats obtenus pour la T1 et pour la T2. En plus des locutions pour lesquelles le taux d'échec est 100% dans les deux tâches, les taux d'échec se sont avérés plus élevés dans la T2 que dans la T1 pour six locutions. Pour les deux locutions que les apprenants ont classées comme familières (PRINCE CHARMANT et JOUER AVEC LE FEU), le taux de réussite a été plus bas dans la T2 (100% vs 77% pour PRINCE CHARMANT et 65% vs. 55% pour JOUER AVEC LE FEU), ce qui confirme la difficulté d'utiliser des locutions en contexte de production chez les apprenants et suggère que la tâche de tissage des liens paradigmatiques est un problème épineux pour les apprenants.

4.2 RÉSULTATS QUALITATIFS: RÉUSSITES ET DIFFICULTÉS

Cette section examine la qualité des définitions proposées et illustre des difficultés d'apprenants par les exemples tirés du corpus. Rappelons que chacune des deux tâches proposées pouvait être entièrement ou partiellement réussie. Une paraphrase du sens (Tâche 1) n'était que partiellement aboutie lorsque sa formulation était trop générale, de manière à englober potentiellement plusieurs signifiés. C'est le cas du lexème « incompatibilité » indiqué pour MARIAGE DE LA CARPE ET DU LAPIN. De même, la Tâche 2, qui consistait à produire un exemple illustrant le sens, était en partie réussie, lorsque sans être incompatible avec le sémantisme de la locution, la phrase proposée ne permettait pas d'admettre que celui-ci était réellement maîtrisé. C'est le cas notamment de: *Jack est plein d'énergie. Linda a très peur et silencieuse. C'est un mariage de la carpe et du lapin*, où il est question d'un couple, ce qui, du point de vue sémantique apparaît comme trop littéral et restreint.

Notre but était d'observer des tendances dans les réponses et d'identifier les types de difficultés rencontrées par des apprenants lors d'application de stratégies de devinement. Nous avons

identifié trois types majeurs de problèmes dans l'élaboration des définitions : (1) décomposition du sens dénotatif ; (2) confusion avec des homonymes et (3) interférence avec la langue maternelle. Nous les illustrons dans ce qui suit.

La première difficulté provient de la décomposition et analyse du sens dénotatif des composantes de la locution. Par exemple, pour la locution NOYER LE POISSON ('embrouiller volontairement une situation, faire des digressions dans le but de tromper ou de lasser'[Antidote]), 11 réponses sur 17, soit 65%, proposent le sens 'faire une chose impossible', comme illustré dans (1a-c). Le syntagme *noyer le poisson* est alors analysé par l'apprenant comme cela se fait dans un syntagme libre, et non dans un phrasème non-compositionnel ; en analysant le sens dénotatif de *noyer* et de *poisson*, il arrive à cette interprétation car il est effectivement impossible de noyer le poisson au sens littéral.

- (1) a. Essayer d'accomplir tous les devoirs c'est comme essayer de noyer le poisson. Autrement dit, cette chose est difficile et ça semble impossible.
 b. Habiter sur Mars c'est comme noyer le poisson.
 c. Je ne peux pas croire que tu as eu 100 %*sur l'examen ! Tu as noyé le poisson !⁴

Les exemples en (2) illustrent la confusion chez l'apprenant due à l'homonymie. Ainsi, pour expliquer la locution VOIE DE GARAGE ('situation sans issue') 4 réponses sur 7 se sont basées sur le sens de VOIX, lexie homonyme de VOIE.

- (2) a. *une voie forte et haute
 b. Quand la prof essaie d'avoir l'attention de la classe, elle utilise une *voie de garage.

Comme l'on pouvait s'y attendre, la langue maternelle d'apprenants (l'anglais) a joué un rôle en même temps positif et négatif dans l'interprétation du sens. Par exemple, la locution AVOIR LES DEUX PIEDS DANS LE MÊME SABOT ('être incapable de faire quoi que ce soit, rester passif, n'avoir aucune initiative, ne manifester aucune débrouillardise' [Antidote]), serait confondue avec la locution anglaise PUT ALL ONE'S EGGS IN ONE BASKET (équivalent de METTRE TOUS SES ŒUFS DANS LE MÊME PANIER), car le sens 'investir tout dans quelque chose' a été proposé, tel qu'illustré dans (3a). Une interprétation semblable, 'concentrer tout l'effort et l'énergie à une place ou dans un domaine' est illustrée dans (3b).

- (3) a. Elle *applique [fait une demande] seulement à MSVU, elle a les deux pieds dans le même sabot. Autrement dit, elle espère vraiment qu'elle ira à MSVU. Elle n'a pas pigé d'autres options.
 b. Jérémie a quitté l'école pour devenir musicien ; il a mis les deux pieds dans le même sabot. Autrement dit, il a décidé de seulement se concentrer sur sa musique et rien d'autre.

⁴ Les phrases d'exemples qui contiennent des erreurs grammaticales ou lexicales sont marquées par un astérisque.

Enfin, nous avons constaté l'évitement d'engagement linguistique chez des apprenants. Considérons les exemples en (4) proposées par les participants pour la T2 pour la locution PRINCE CHARMANT. Alors que ces phrases sont parfaitement acceptables dans un discours, elles n'explicitent pas le sens de la locution par le contexte. Nous les appelons des *exemples douteux*⁵ car ils sont sécurisants pour un apprenant qui évite le risque de se tromper en employant la locution dans un contexte moins précis. Des degrés d'acceptabilité de réponses sont différents, car (4a) semble plus précis et plus acceptable que (4c) très général.

- (4) a. Un jour tu trouveras ton prince charmant.
 b. Ce garçon est mon prince charmant.
 c. C'est mon prince charmant.

Nous avons aussi observé l'influence de la langue maternelle sur la réussite dans les définitions. Les expressions accessibles pour les apprenants ont souvent un équivalent locutionnel en anglais: soit un équivalent direct lexicalement et syntaxiquement presque identique (JOUER AVEC LE FEU / PLAY WITH FIRE, PRINCE CHARMANT / PRINCE CHARMING), soit un équivalent sémantique (CHEVALIER SERVANT / KNIGHT IN THE SHINING ARMOUR).

En ce qui concerne les locutions « plutôt inaccessibles », souvent, il n'y a pas de locution en anglais exprimant le même sens locutionnel. Par exemple, SE NOYER DANS UN VERRE D'EAU ('ne pas savoir se débrouiller, ne pas pouvoir surmonter de petits obstacles') n'a pas d'équivalent direct en anglais (cf. Tableau 3).⁶ Une expression assez proche sémantiquement BE AS HELPLESS AS A BABE ('unable to manage independently' ['ne pas être en mesure de résoudre un problème indépendamment']) ne dénote pas exactement la même situation, car il s'agit d'une généralisation de qualités de l'individu X, alors que la composante 'ne pas savoir surpasser un petit obstacle' manque. Par contre, dans la locution MAKE A MOUNTAIN OUT OF A MOLEHILL ('to make a slight difficulty seem like a serious problem'[Cambridge Dictionary] / 'traiter un problème trivial en tant qu'un problème grave' proposé comme traduction par certains (expressio.fr) n'est pas tout à fait équivalent car le sens 'exagérer' est surtout mis en évidence en anglais.

5. DISCUSSION

Revenons à nos questions de recherche sur lesquelles nous nous sommes focalisées dans cette étude :

- (1) Lorsque les locutions sélectionnées selon les critères ci-dessus étaient connues des participants, parvenaient-ils à en paraphraser le sens, à construire un exemple de leur emploi, à les connecter à des lexies (quasi-)synonymes ?

⁵ Nous empruntons ce terme, en l'adaptant aux exemples, à Hasselgren (1994), qui qualifie de *doutés lexicaux* (*lexical Teddy Bears*) ce type de choix sécurisants en ce qui concerne le vocabulaire.

⁶ Dans le Tableau 3 nous avons proposé les variantes de traductions en anglais qui nous ont semblées les plus proches sémantiquement. Nous avons consulté les dictionnaires Cambridge et Collins et avons privilégié les variantes de l'anglais américain versus l'anglais britannique, car nos apprenants habitent en Amérique du Nord.

Des 13 locutions proposées seulement deux ont été familières aux apprenants. Quand les locutions étaient connues des participants, de manière générale, ils parvenaient à en paraphraser le sens, mais le taux de la réussite a été régulièrement plus bas pour la T2 que pour la T1. Le paraphrasage exige l'usage de (quasi-)synonymes ou de constructions lexico-syntaxiques quasi-équivalentes dans les substitutions, ce qui représente, comme on le sait, une faiblesse chez ce type d'apprenant (Tsedryk 2016). Ce résultat témoigne du manque de connaissances phraséologiques chez des apprenants et du besoin d'incorporer la phraséodidactique, en général, et l'enseignement de locutions, en particulier, dans le cursus scolaire, ce qui n'est pas fait de façon régulière dans les manuels de langue de niveau intermédiaire-avancé (Tsedryk 2018).

- (2) Lorsque les locutions sélectionnées inconnues aux étudiants, s'engageaient-ils systématiquement dans des démarches d'analyse, en appliquant notamment des stratégies de devinement pour accéder à leur sens? Plus précisément, (a) Quel est le degré de maîtrise sémantique des locutions proposées? (b) Quels types de difficultés les apprenants ont-ils? (c) Est-ce qu'il y a un lien entre la réussite et la fréquence des locutions et/ou la langue maternelle d'apprenants ?

Lorsque les locutions étaient inconnues aux apprenants, plusieurs peinaient à en faire une analyse et identifier le sens, en s'abstenant de donner des réponses. Si l'on s'attendait à obtenir 40 réponses par locution, en moyenne, on n'a obtenu que 25. Certains apprenants n'ont pas proposé de réponses du tout, alors que d'autres se sont limités à des stratégies d'évitement, ce qui conduit à des résultats parfois difficiles à interpréter. Cette tendance témoigne clairement de la complexité de l'apprentissage des locutions en L2 et de difficultés qu'un apprenant éprouve souvent à deviner le sens de celles-ci en raison de leur non-compositionnalité (Sikora 2018a). Cette difficulté représente également un défi pour les locuteurs natifs de français. Ainsi, pour les locuteurs natifs testés en France dans la tâche de définition, les locutions « les plus difficiles » ou « plutôt inaccessibles » ont été les mêmes que dans le groupe de L2 (Sikora 2018b). Généralement, le taux de réussite est bas chez nos participants : souvent les stratégies de devinement n'aboutissent pas à la réussite dans la tâche de définition ; celle-ci n'est pas triviale pour un apprenant à cause de l'effort cognitif nécessaire et aussi le manque de métaconnaissances (Sikora 2018b). Les difficultés d'apprenants que nous avons identifiées s'expliquent par le recours au sens dénotatif des composantes de locutions, par le manque de connaissances linguistiques (confusion entre les homonymes, cf. Laufer 1997) et quelquefois par un transfert négatif de la langue maternelle. Nous n'avons pas remarqué de lien évident entre la réussite dans les tâches et la fréquence de locutions. Ceci peut s'expliquer par une fréquence très basse de locutions dans les corpus et dans les discours, en général, et, comme conséquence, absence ou quasi-inexistence d'apprentissage incident de locutions.

Même si les difficultés d'apprenants étaient évidentes, le sens de certaines locutions s'est avéré plus accessible. Nous avons remarqué un lien positif entre la réussite et l'existence d'un équivalent locutionnel en L1. Notamment, les locutions en français ayant un équivalent locutionnel en anglais ont reçu un taux plus élevé de réponses et de réussites dans les deux tâches. En revanche, dans plusieurs cas, l'identité sémantique pour chaque locution a été impossible à trouver, et la composante imagée fut difficile à être ciblée pour nos apprenants, tel que démontré par la liste de locutions « plus ou moins / plutôt inaccessibles ». Cette problématique déjà éclairées

dans les travaux théoriques (cf. Dobrovol'skij 2016 ; Sikora 2018a,) a été attestée dans notre étude-pilote de façon expérimentale. Pour cette raison, il est important en didactique de langues de prêter une attention particulière au choix de locutions et à leur traduction en langue maternelle de l'apprenant. Des activités pédagogiques visant la comparaison de l'image véhiculée par la locution en langues et cultures différentes seraient souhaitables au niveau intermédiaire-avancé (B1 et B2 du CECRL).

Comme on pouvait s'y attendre, quelques locutions se sont avérées plus analysables, plus transparentes que d'autres pour les apprenants, même si l'équivalent locutionnel en L1 n'était pas toujours disponible (DANSER SUR UN VOLCAN, MARIAGE DE LA CARPE ET DU LAPIN, COURIR DEUX LIÈVRES À LA FOIS). On peut supposer que leurs composantes imagées ont été plus facilement identifiées par les apprenants, car ils ont pu utiliser ces locutions en contexte de production langagière avec succès. D'autre part, ces expressions facilement décomposables contiennent des mots-formes appartenant plutôt au domaine du concret, ce qui pourrait faciliter leur compréhension. Il est intéressant de noter que nos résultats font écho à l'analyse psycholinguistique de Caillies (2009) en français L1 pour les mêmes locutions. Selon cette étude, la locution *TOURNER AUTOUR DU POT* a été perçue comme moins décomposable que *JOUER AVEC LE FEU*. Des recherches empiriques sur des locutions différentes en français (en L1 et en L2) doivent être effectuées afin d'obtenir plus de données descriptives de locutions, ce qui faciliterait ensuite le choix de locutions à enseigner aux apprenants de langues.

Il faudrait s'arrêter sur des limites de cette étude et les questions pour les recherches ultérieures. Premièrement, les résultats du test sont en lien direct avec le choix de locutions et pourraient montrer d'autres tendances dans les stratégies cognitives de devinement si on choisissait d'autres locutions. Est-ce que le taux d'absence de réponses et les stratégies d'évitement seraient aussi nombreuses? Est-ce que les stratégies utilisées par des apprenants seraient les mêmes? Deuxièmement, le nombre de participants a été plutôt limité. Est-ce qu'on obtiendrait les mêmes taux d'échec ou de réussite pour les mêmes locutions, si l'on testait des centaines d'apprenants? Questionner plus d'apprenants en se basant sur le même protocole d'analyse de données serait avantageux pour confirmer les tendances dans nos résultats. Également, il serait intéressant de comparer les démarches d'apprenants de L2 avec ceux de locuteurs natifs pour exactement les mêmes tâches afin de vérifier la connaissance de ces locutions et les stratégies de devinement déployées par les natifs en cas de locutions qui leur sont méconnues. Cette démarche pourrait affiner nos analyses et compréhension des stratégies de devinement. Troisièmement, le test devrait être plus explicite sur l'usage de plusieurs phrases dans la T2 pour éviter les doudous lexicaux et obtenir des réponses plus détaillées.

6. CONCLUSIONS

Notre étude-pilote avait un double objectif: (1) vérification de la connaissance des 13 locutions françaises (choisies selon des critères particuliers) par des apprenants intermédiaires-avancés de français L2 lors des tâches de définition et de production ; (2) description quantitative et qualitative de difficultés d'apprenants vis-à-vis du vocabulaire locutionnel à partir des données expérimentales. Pour les locutions proposées, les connaissances sont peu développées, voire inexistantes, ce qui souligne le besoin d'une attention particulière au développement de la compétence

phraséologique à ce niveau d'études. Tel qu'attendu, en raison de la non-compositionnalité de locutions, deviner le sens de locutions à partir d'indices linguistiques et le contenu figuratif véhiculé n'a pas été facile pour les apprenants. Les difficultés étaient liées surtout à la décomposition erronée du sens dénotatif, ainsi qu'aux problèmes dus à l'homonymie et à l'interférence de la L1. La réussite d'accès au sens de quelques locutions semble être liée plutôt à l'existence d'un équivalent locutionnel en langue maternelle et non à la fréquence de locutions. Il convient de considérer ces facteurs en élaborant une liste de locutions à enseigner pour les niveaux B1 et B2. Ces conclusions provisoires ne font que souligner l'importance de l'enseignement explicite et structuré d'unités phraséologiques dans le cursus scolaire. C'est alors aux linguistes et aux didacticiens de se retrousser les manches.

REFERENCES

- Abel, Beate. 2003. English idioms in the first language and second language lexicon: a dual representation approach. *Second Language Research* 19(4). 329-358.
- Bally, Charles. 1909. *Traité de stylistique française*. Genève : Georg.
- Boers, Frank. 2001. Remembering figurative idioms by hypothesising about their origins. *Prospect* 16(3). 35-43.
- Boers, Frank & Demecheleer, Murielle & Eyckmans, June. 2004. Etymological elaboration as a strategy for learning idioms. In Bogaards, P. B. Laufer (éd), *Vocabulary in a Second Language Selection, Acquisition and Testing*. 53-78. Amsterdam: John Benjamins.
- Boers, Frank & Eyckmans, June & Strengers, Hélène. 2007. Presenting figurative idioms with a touch of etymology - More than mere mnemonics? *Language Teaching Research* 11(1). 43-62.
- Bortfeld, Heather. 2012. What native and non-native speakers' images for idioms tell us about figurative language? *Advances in Psychology* 134. 275-295.
- Caillies, Stephanie. 2009. Description de 300 expressions idiomatiques : Familiarité, Connaissance de leur signification, plausibilité littérale, « décomposabilité » et « prédictibilité ». *L'Année psychologique* 109(3). 463-508.
- Coulmas, Florian. 1981. *Conversational Routine*. La Haye: Mouton.
- Dobrovol'skij, Dmitrij. 2016. The notion of "inner form" and idiom semantics. *Études et Travaux* 1. 21-36.
- Dobrovol'skij, Dmitrij & Piirainen, Elisabeth. 2005. *Figurative Language: Cross-Cultural and Cross-Linguistic Perspectives*. Amsterdam: Elsevier.
- Edmonds, Amanda. 2013. Une approche psycholinguistique des phénomènes phraséologiques: le cas des expressions conventionnelles. *Langages* 189(1). 121-138.
- González Rey, I. 2007. *La didactique du français idiomatique*. Fernelmont: Éditions E.M.E.
- González Rey, I. 2010. *La Phraséodidactique en action: Les expressions figées comme objet d'enseignement*. Lyon: ENS LSH/DGESCO.
- González Rey, I. 2018. Phraséotext – le français idiomatique: Une méthode d'enseignement-apprentissage en phraséodidactique du FLE ». In Soutet, O. S. Mejri, I. Sfar (éds). *La phraséologie: Théories et applications*. 301-318. Paris: Honoré Champion.

- Gréciano, Gertrud. 1984. Pour un apprentissage des unités phraséologiques. *Nouveaux Cahiers d'Allemand* 2. 95-113.
- Groupe de Fribourg. 2012. *Grammaire de la période*. Bern: Peter Lang.
- Hasselgren, Angela. 1994. Lexical Teddy Bears and Advanced Learners: a Study into the Ways Norwegian Students Cope with English Vocabulary. *International Journal of Applied Linguistics* 4 (2). 237-260.
- Kauffer, Maurice. 2011. Actes de langage stéréotypés en allemand et en français. Pour une redéfinition du stéréotype grâce à la phraséologie. *Nouveaux Cahiers d'Allemand*. 35-53.
- Laufer, Batia. 1997. What's in a word that makes it hard or easy. Some intralexical factors that affect the learning of words. In Schmitt, N. M. McCarthy (ed). *Vocabulary: Description, Acquisition and Pedagogy*, 140-155. Cambridge: Cambridge University Press.
- Mel'čuk, Igor A. 2013. Tout ce que nous voulions savoir sur les phrasèmes, mais... *Cahiers de Lexicologie* 102(1): 129-149.
- Oxford, Rebecca L. 1990. *Language Learning Strategies - What Every Teacher Should Know*. New York: Newbury House Publishers.
- Pausé, Marie-Sophie. 2017. *Structure lexico-syntaxique des locutions du français et incidence sur leur combinatoire*. Nancy : Université de Lorraine and ATILF (Doctoral dissertation).
- Piirainen, Elisabeth. 2008. Figurative phraseology and culture . In Granger, S. F. Meunier (ed). *Phraseology : An Interdisciplinary Perspective*. 207-228. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Sikora, Dorota. 2018a. Locutions en apprentissage de langue seconde. SHS Web Conf 46, (https://www.shs-conferences.org/articles/shsconf/pdf/2018/07/shsconf_cmlf2018_05002.pdf) (Accessed 2019-03-26.)
- Sikora, Dorota. 2018b. Entre sens et compétences métalinguistiques : locutions en phraséodidactique du FLE. *Lettre de l'AIRDF* 64. 60-66.
- Siyanova-Chanturia, Anna & Conklin, Kathy & Norbert, Schmitt. 2011. Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research* 27(2). 251-272.
- Stengers, Hélène & Boers, Frank & Housen, Alex. 2011. Formulaic sequences and L2 oral proficiency: Does the type of target language influence the association? *International Review of Applied Linguistics* 49(4). 321-343.
- Sułkowska, Monika. 2013. *De la phraséologie à la phraséodidactique*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Sułkowska, Monika. 2018. Phraséodidactique et compétences phraséologiques : recherches, analyses et bilan des résultats. In Soutet, O. S. Mejri, I. Sfar (ed.) *La phraséologie: Théories et applications*. 319-338. Paris: Honoré Champion.
- Théophanous, Olga. 2013. Le traitement de la préconstruction lexicale dans un corpus de méthodes de français langue étrangère. Garcia-Debanc, C. C. Masseron, Ch. Ronveaux (ed). *Enseigner le lexique*. 173-192. Namur : Presses Universitaires de Namur.
- Théophanous, Olga & Perez-Bettan, Anne. 2018. Apport des séquences préfabriquées dans la fluence verbale en langue seconde/étrangère. In Soutet, O. S. Mejri, I. Sfar (ed.) *La phraséologie: Théories et applications*. 339-354. Paris : Honoré Champion.

- Tsedryk, Alexandra. 2016. *La compétence paraphrastique en français langue seconde*. Bern : Peter Lang.
- Tsedryk, Alexandra. 2018. Comment ne pas tomber dans les pommes en étudiant le français : l'apprentissage d'expressions idiomatiques en L2 aux niveaux intermédiaire et avancé. (Paper presented at the 32^e Congrès Mondial du Conseil International d'Études Francophones. Université de La Rochelle, France, 4-10 June 2018.)
- Wray, Alison. 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.
- Yongqi, Gu. 2010. Learning Strategies for Vocabulary Development. *Reflections on English Language Teaching* 9(2). 105-118.

Dictionnaires et corpus consultés :

- Antidote RX, version 9 [Logiciel], Montréal, Druide informatique, 2018.
- Dictionnaire Cambridge en ligne. (<https://dictionary.cambridge.org/>)
- Dictionnaire Collins en ligne. (<https://www.collinsdictionary.com/dictionary/english>)
- Frantext. Base de données en ligne. (<https://www.frantext.fr/>)
- FrWac. Base de données en ligne. (<https://www.sketchengine.eu/frwac-french-corpus/>)
- Le Petit Robert. Dictionnaire de la langue française. Version électronique 2018.
- Rey, Alain & Sophie Chartreau. 2002. *Dictionnaire des expressions et locutions*. Paris : Éditions Le Robert.

INTRODUCTION

L'ESPACE DANS LE ROMAN : ÉTAT DES LIEUX ET PERSPECTIVES

Catherine HAMAN
Dorota SIKORA

Qui pense espace et littérature ne peut que songer immédiatement à l'ouvrage fondamental de Maurice Blanchot, *L'Espace littéraire*, qui définit métaphoriquement celle-ci comme le lieu d'un absolu, séparé du réel. Notre recueil adoptera a contrario une démarche *géocentrée* ; il se propose d'étudier l'espace dans le roman comme image recomposée d'un référent extratextuel, dans une perspective qui n'emprunte ni à Blanchot ni aux pratiques structuralistes des années 60-70. Il vise en effet à inscrire l'œuvre dans un contexte géographique. La forme générique qu'il prend pour objet est le roman français postérieur à 1980.

Il a longtemps été d'usage en critique de mettre l'accent sur la temporalité dans le roman, de rapprocher roman et Histoire, rapprochement favorisé par le développement et la prépondérance de l'histoire littéraire, par l'existence du roman historique, et par la valeur documentaire accordée aux textes littéraires du passé. L'étude de l'espace a longtemps occupé une portion congrue, se limitant pour l'essentiel à l'analyse du décor et des pratiques descriptives. *Temps et Récit* de Paul Ricœur se compose de trois tomes ; Gérard Genette travaille longuement sur le temps en narratologie. Mais il ne consacre que deux brefs articles à l'espace dans *Figures*¹, employant le terme au sens figuré, et traitant pour l'essentiel de la spatialité du langage, la période structuraliste se caractérisant, on le sait, par la prééminence du fait linguistique.

La notion de géographie littéraire est cependant relativement ancienne ; apparue au début du XX^e siècle, lors de la mode du provincialisme littéraire, elle ne s'impose alors pas en France, contrairement à l'histoire littéraire. Deux ouvrages cependant méritent mention, celui d'Albert Thibaudet² et celui d'André Ferré³. Il faut attendre les années 90 et le *tournant spatial* pour que celle-ci revienne en force et occupe un large pan des études critiques. Le modèle historique dominant, mis à mal par la fin des *grands récits* fédérateurs et les théories de la *fin de l'histoire*, cède le pas à de nouvelles conceptions, à une *géo-histoire*, pour reprendre le néologisme de Fernand Braudel⁴ : son présupposé est qu'il est impossible de penser la dimension spatiale des sociétés sans sa dimension temporelle, et réciproquement⁵.

Le roman contemporain pour sa part, nous le verrons, passe aussi par l'espace, le lieu, le paysage, pour décrire l'histoire humaine, individuelle, sociale. Rappelons du reste que le critique littéraire américain Frédéric Jameson décrit le postmodernisme comme une spatialisation de la culture sous la pression du capitalisme.

APPROCHES CRITIQUES DE L'ESPACE EN LITTÉRATURE AU COURS DU XX^{ÈME} SIÈCLE

Quoique la prédominance du modèle historique ait été indéniable, il nous faut rappeler que quelques chercheurs ont malgré tout travaillé tout au long du vingtième siècle à la question de l'espace dans le roman. Dès le début du siècle, dans des ouvrages qui ne seront traduits que dans les années 70-80, Mikhaïl Bakhtine puis Youri Lotman affirment que la structuration de l'espace fictionnel est indispensable à l'élaboration du sens. Pour eux, la forme que l'espace emprunte dans le roman traduit toujours une vision du monde. M. Bakhtine⁶, dans la perspective d'une poétique historique, définit les repères et marqueurs spatio-temporels d'une œuvre comme caractéristiques d'une époque historique et les désigne du nom de chronotopes. Y. Lotman, adoptant une démarche de sémiotique culturelle, affirme que l'être humain structure l'espace selon des oppositions binaires : haut/bas, gauche/droite, devant/derrière... Ces oppositions se retrouvent dans le roman, associées à des polarités axiologiques du type « valable-non-valable, bon-mauvais, les siens-les étrangers, accessible-inaccessible, mortel-

1 G. Genette, « Espace et langage », dans *Figures I*, Seuil, Coll. « Tel Quel », 1966, « La littérature et l'espace » dans *Figures II*, Seuil, Coll. « Tel Quel », 1969.

2 A. Thibaudet, « Pour la géographie littéraire », NRF, Avril 1929, recueilli dans *Réflexions sur la littérature*, Gallimard, coll. « Quarto », 2007.

3 A. Ferré, *Géographie de Marcel Proust avec index des noms de lieux et des termes géographiques*, Le Sagittaire, 1939.

4 Voir par exemple F. Braudel, *La Méditerranée et le monde méditerranéen à l'époque de Philippe II*, A. Colin, 1949.

5 « La grande hantise qui a obsédé le XIX^e siècle a été, on le sait, l'histoire [...] L'époque actuelle serait peut-être celle de l'espace », M. Foucault, « des espaces autres », cité par M. Collot, *Pour une géographie littéraire*, Corti, coll. « Les Essais », 2014, p. 16.

6 M. Bakhtine, *Esthétique et théorie du roman*, Gallimard, 1978.

immortel, etc.⁷ ». Y. Lotman distingue la *topographie*, c'est-à-dire la représentation d'espaces concrets, de la *topologie*, qui illustre les structures signifiantes de base, à savoir les structures invariantes d'une culture – par exemple la frontière qui divise « tout l'espace du texte en deux sous-espaces, qui ne se recoupent pas mutuellement⁸ » –, dont l'existence motive fondamentalement l'intrigue romanesque. Outre-Atlantique, la « forme spatiale » de Joseph Frank⁹, illustrée par exemple par les œuvres de Proust et de Joyce, met l'accent sur des procédés narratifs qui bouleversent la chronologie du récit – sans pour autant la faire disparaître – en juxtaposant des scènes ou événements dans un effet de simultanéité. Ce type d'organisation narrative identifiée par J. Frank privilégie les relations synchroniques au détriment d'une composition diachronique, ce qui favorise le surgissement d'images spatiales semblables à une série d'instantanés.

En France, dans la mouvance de la phénoménologie, qui étudie l'homme dans son être-au-monde et intègre de facto le sujet dans une spatialité¹⁰, Gaston Bachelard travaille dès les années 30 à la notion d'espace et publie *La poétique de l'espace* en 1957¹¹. La représentation de celui-ci dans l'œuvre reflète pour lui une vision singulière de l'espace réel et la dimension affective qui lui est associée. Étudiant l'image de la maison, du nid, de la coquille, il élabore une « topoanalyse », ou « étude psychologique systématique des sites de notre vie intime¹² ». Tout en étudiant essentiellement la poésie, G. Bachelard évoque cependant quelques romanciers admirés, comme Henri Bosco, Rilke, et le Hugo de *Notre-Dame de Paris*. Dans son sillage, des critiques de la conscience et de l'imaginaire comme Georges Poulet puis Jean-Pierre Richard vont travailler à partir des années 50 à la perception subjective de l'espace de certains écrivains, Proust notamment. Dans les travaux de J.-P. Richard, le mot *paysage* ne désigne pas le ou les lieux où un écrivain a vécu ou voyagé, mais une image singulière du monde, rattachée à un style et une sensibilité. La peinture du paysage littéraire doit se lire comme une mise en forme réciproque du monde et de l'œuvre.

La période structuraliste se consacre peu, on l'a dit, à l'étude de l'espace. Cependant, à l'occasion de la publication par Georges Matoré, lexicologue et professeur à la Sorbonne, de *L'Espace humain. L'expression de l'espace dans la vie, la pensée et l'art contemporains*¹³, qui propose un lexique spatial représentatif de l'esprit de l'époque, Gérard Genette commente l'ouvrage dans « Espace et langage¹⁴ ». Il reviendra à la question de l'espace dans « La littérature et l'espace¹⁵ ». Il y propose quatre types de spatialisations : la spatialisation primaire (le langage), la spatialité de l'écriture (la graphie), la spatialité stylistique (figures rhétoriques, sens figuré) et la spatialité de la littérature dans son ensemble (lectures, croisements entre différents textes). L'usage qu'il fait de l'espace est, on le voit, centré autour du langage et de ses caractéristiques, selon les modes de pensée propres au structuralisme. La question d'une possible référentialité est esquivée, le texte est perçu comme une entité autonome, close.

Roland Bourneuf, dans un article fondateur en ce qu'il cherche à poser les fondements d'une étude globale de l'organisation de l'espace dans le roman¹⁶, aborde de son côté la question sous trois angles différents mais complémentaires : « l'espace dans sa relation avec l'auteur, avec le lecteur, avec les autres éléments constitutifs du roman¹⁷ ». Le premier aspect s'apparenterait à une poétique de l'espace, similaire à celle de G. Bachelard. Le second aspect renvoie à l'interférence entre l'espace imaginaire et l'univers réel du lecteur, comme Michel Butor l'expose dans « L'espace du roman¹⁸ ». Pour Bourneuf, l'espace est enfin un tout, les différents lieux étant liés les uns aux autres et devant s'interpréter dans leur globalité. Il serait donc nécessaire de travailler à la « topographie¹⁹ » du roman, afin d'en repérer la configuration originale : espace clos, ouvert, unipolaire ou multipolaire, etc. À la fin des années 70, étudiant le roman du XVIII^e siècle²⁰, Jean Weisgerber prolonge les travaux de Y. Lotman et définit l'espace en termes relationnels : le récit se construit à partir de structures spatiales binaires telles que gauche/droite, haut/bas ou encore avant/arrière. À celles-ci se rattachent un jugement de valeur ou des

7 Y. Lotman, *La Structure du texte artistique*, Gallimard, 1973, p. 311.

8 *Ibid.*, p. 321.

9 J. Frank, « Spatial Form in Modern Literature. An Essay in Two Parts », *The Sewanee Review*, vol. 53, n°2, 1945.

10 Voir M. Heidegger, *Bâtir, habiter, penser*, [1958], Gallimard, coll. « Tel », 1980.

11 G. Bachelard, *La Poétique de l'espace*, Presses universitaires de France, 1957.

12 *Ibid.*, p. 27.

13 G. Matoré, *L'Espace humain. L'expression de l'espace dans la vie, la pensée et l'art contemporains*, Nizet, 1976.

14 G. Genette, « Espace et langage », *op. cit.*

15 G. Genette, « La littérature et l'espace », *op. cit.*

16 R. Bourneuf, « L'Organisation de l'espace dans le roman », Québec : *Études littéraires*, vol. 3, n°1, 1970.

17 *Ibid.*, p. 80.

18 M. Butor, « L'Espace du roman », *Essais sur le roman*, Gallimard, 1964.

19 R. Bourneuf, *art. cit.*, p. 82.

20 J. Weisgerber, *L'Espace romanesque*, Lausanne, L'Âge d'homme, 1978.

significations particulières. Le critique refuse de cantonner son étude aux sphères du décor et de la description. L'espace n'est pas selon lui limité, contrairement à ce que pense la théorie littéraire traditionnelle, à une fonction ornementale ou d'encadrement, qui l'exclut de toute progression chronologique, de l'avancée de l'intrigue, de l'évolution des personnages et de la production de sens.

De même, dans *Le Discours du roman*²¹, Henri Mitterrand, prolongeant les pratiques de la narratologie, voit l'espace comme un « champ de déploiement des actants et de leurs actes²² ». L'espace fait émerger le récit, détermine les relations entre les personnages et influence leurs actions. L'espace prend forme grâce à la description mais aussi par la combinaison de plusieurs éléments (narration, personnages, temps, actions). Dans la lignée de Y. Lotman lui aussi, H. Mitterrand affirme qu'une étude spatiale ne se cantonne pas au pur topographique mais exige également une topologie qui dégage ses valeurs symboliques et idéologiques. Il appelle du reste de ses vœux un « répertoire morphologique et fonctionnel des lieux romanesques²³ ».

La sémiologie de l'espace, de son côté, considère que l'espace urbain est jalonné de signifiants interprétables comme un langage. On peut lire l'espace urbain comme on lirait un texte, chargé qu'il est de signes porteurs de sens. Ainsi, dans les années quatre-vingt, Roland Barthes s'essaie à une sémantique de la ville²⁴, M. Butor élabore une « grammaire urbaine²⁵ », Michel Foucault élabore le concept d'hétérotopie, qui désigne la spécificité de certains espaces, souvent clos ou enclavés, caractérisés par une discontinuité avec ce qui les entoure²⁶. Dans les trois cas, il s'agit d'apprendre à lire l'espace, et de proposer un parallèle entre la littérature et les autres productions humaines : tableau, ville, etc. Cette sémiotique touche indirectement aux formes romanesques et à leur analyse, dans la mesure où espace et espace textuel se voient associés. À l'occasion du tournant spatial, les anthropologues et les géographes, tels Michel de Certeau ou Edward Soja, recourront également à un parallèle entre espace et texte, tout comme le fera aussi la géopoétique. M. de Certeau considère en effet l'espace urbain comme une « texturologie²⁷ », où chaque marcheur réécrit la ville qu'il sillonne. Sans doute est-ce aussi ce que chaque romancier pratique.

LE TOURNANT SPATIAL

Depuis une trentaine d'années, notre rapport à l'espace s'est profondément modifié, dans le cadre d'une mutation épistémologique générale, provoquée par un certain nombre de problématiques nouvelles, sur fond de déclin de la prééminence occidentale : la chute du Mur de Berlin en 1989 et les conflits ethniques sanglants qui en ont résulté, la crise écologique, les diasporas, les luttes de frontières culturelles. D'autres évolutions interviennent aussi dans ce changement de regard, comme la démocratisation des vols long-courriers, l'accélération des moyens de transport, le développement exponentiel d'un espace virtuel facilitant les échanges et brouillant la dichotomie proche/lointain... Autant d'invites à repenser notre situation dans le monde.

L'idée d'un *tournant spatial* comme phénomène interdisciplinaire dans les sciences sociales a été énoncée pour la première fois en 1989 par le géographe californien Edward Soja dans son essai *Postmodern Geographies*²⁸. Il met ce phénomène en relation avec les problématiques territoriales et les *cultural studies*. La fiction romanesque va elle aussi à cette période relayer les interrogations relatives à l'espace et à l'identité, ainsi qu'aux migrations et à l'exil. Le textualisme et le tout linguistique des années 60-70 sont remplacés par une ouverture sur le monde. L'espace devient, tout autant que le temps et le langage, un outil d'interprétation de l'existence humaine, individuelle et collective, qui affirme désormais son implication dans toute forme de savoir. De même, les nouvelles lectures critiques ne l'interprètent plus comme un simple arrière-plan. Il est perçu comme ressort et enjeu de l'intrigue, axe structurant et signifiant et architecte d'univers fictifs.

En France, de très nombreux travaux ont été consacrés ces dernières décennies à l'inscription de la littérature dans l'espace et à la représentation des lieux dans les textes littéraires, textes particulièrement susceptibles de

21 H. Mitterrand, *Le Discours du roman*, Presses Universitaires de France, 1980.

22 *Ibid.*, p.190.

23 *Ibid.*, p. 193.

24 R. Barthes, « Sémiologie et urbanisme », *L'Aventure sémiologique*, Seuil, 1985.

25 M. Butor, « La Ville comme texte », *Répertoire V*, Minuit, 1982, p. 33-42

26 Voir M. Foucault, « Des espaces autres. », Conférence au Cercle d'études architecturales, 14 mars 1967, repris dans *Architecture, Mouvement, Continuité*, n° 5, 1984, p. 46-49.

27 M. de Certeau, *L'Invention du quotidien*, Gallimard, 1990, p.141.

28 E.W. Soja, *Postmodern Geographies. The Reassertion of Space in Critical Social Theory*, London, New York : Verso, 2011.

retranscrire la relation concrète, affective et symbolique qui unit l'homme aux lieux. Beaucoup de colloques ont été organisés sur ces questions, de nombreux ouvrages et numéros de revue ont traité ce sujet. Il serait impossible de tout citer. Nous renvoyons à la présentation à la fois synthétique et très éclairante de Michel Collot, disponible en ligne²⁹. Il est également notable que la narratologie contemporaine consacre l'essentiel de ses analyses aux questions spatiales. Citons parmi beaucoup d'autres les recherches de Marie-Laure Ryan, Ansgar Nünning et Katrin Dennerlein.

La géographie de la littérature

Quatre tendances essentielles sont repérables dans ce renouveau des études spatiales. Pour plus de détails, je renvoie aux remarquables analyses de M. Collot³⁰. La géographie de la littérature se propose tout d'abord d'étudier le contexte spatial dans lequel les œuvres sont écrites et diffusées. Le principe est essentiellement référentiel. La perspective est géographique, mais aussi historique, sociale et culturelle. Sacrifiant à la mode actuelle du *mapping*, en vogue dans les sciences humaines et sociales, la géographie de la littérature recourt aussi fréquemment à la cartographie pour mieux saisir la façon dont la géographie humaine de l'Europe a influencé le roman. En s'appuyant sur un large corpus de textes d'époques et d'espaces différents, Franco Moretti a publié un *Atlas du roman européen. 1800-1900*³¹. Une centaine de cartes spatiales de la littérature cherchent à démontrer que la géographie réussit à engendrer le roman moderne, et qu'il y a solidarité entre l'expansion du genre romanesque et la formation des États-nations. F. Moretti souhaite en effet étudier « l'espace dans la littérature » mais aussi « la littérature dans l'espace³² ». La géographie de la littérature implique aussi une réflexion sur les différentes ères littéraires et culturelles et leurs implications géopolitiques (études post-coloniales, etc.).

La géocritique

La géocritique naît dans les années 2000 ; elle étudie les représentations de l'espace dans les textes eux-mêmes et se situe plutôt sur le plan de l'imaginaire et de la thématique littéraires. Elle s'emploie à étudier les images et les significations produites par les lieux. Bertrand Westphal³³ est l'inventeur du terme. L'expansion du thème géographique dans la critique contemporaine accompagne le *retour au réel* après la période formaliste, elle est à rapprocher de l'importance que revêt l'espace en philosophie ces cinquante dernières années, notamment chez Gilles Deleuze et Félix Guattari, Michel Foucault, David Harvey ou Henri Lefebvre. B. Westphal s'inspire de la dialectique de G. Deleuze et F. Guattari entre territorialisation et déterritorialisation pour « repenser le lien entre espaces humains et littérature³⁴ ». Il analyse les interactions réciproques entre espace réel et représentations de l'espace : le référent spatial d'un texte est déjà lui-même en partie chargé de références littéraires, une réalité géographique est pour lui assimilable à un texte, puisque le réel est toujours déjà façonné par la littérature. Les objets d'étude de la géocritique sont des lieux chargés d'histoire et de culture, pour lesquels il met en parallèle les différentes images qu'en ont proposées divers écrivains. La théorie de B. Westphal est au croisement des études littéraires, de la géographie, des études urbaines, de la sociologie, de la philosophie et de l'architecture.

Sa démarche est *géocentrée*, l'accent étant placé « davantage sur l'espace observé que sur l'observateur saisi dans sa spécificité³⁵ ». Renonçant à une approche *égocentrée*, la géocritique ne concentre pas son attention sur un auteur ou une époque, mais sur un espace spécifique – région, ville, pays – selon un principe à la fois synchronique, diachronique, thématologique et imagologique. Tout comme Marc Augé, B. Westphal perçoit le réel postmoderne comme instable, incertain et problématique. Les notions de vrai et de fictif y perdent leurs significations opposées et tranchées. Il considère que « c'est la littérature postmoderne qui s'adapte le mieux à cette nouvelle version du réel, le "réel déréalisé" ; c'est peut-être elle qui offre les meilleures options de lecture du monde, en

29 Voir : www.canal-u.tv/video/universite_toulouse_ii_le_mirail/tendances_actuelles_de_la_geographie_litteraire_michel_collot.14338. Il est également possible de consulter la bibliographie détaillée de Marie-Hélène Voyer dans *Terrains vagues. Poétique de l'espace incertain dans le roman français et québécois contemporain*, Québec : Nota Bene, 2018.

30 M. Collot, *Pour une géographie littéraire*, Paris, Corti, 2014.

31 F. Moretti., *Atlas du roman européen (1800-1900)*, Seuil, 2000.

32 *Ibid.*, p. 9.

33 B. Westphal, *La Géocritique : réel, fiction, espace*, Minit, 2007.

34 *Ibid.*, p. 17.

35 B. Westphal, « Pour une approche géocritique des textes. Esquisse », B. Westphal (dir.), *La Géocritique mode d'emploi*, Limoges, Presses universitaires de Limoges, 2000, p. 32.

vertu de sa fictionnalité même³⁶ ». Robert T. Tally et Eric Prieto poursuivent le même type de recherche outre-Atlantique. Toutefois, pour R. T. Tally, la géocritique appliquée à la littérature ne consiste pas en une analyse de la représentation littéraire d'un espace spécifique. Il projette de développer de nouveaux modèles théoriques et critiques afin de mieux saisir la manière dont les auteurs et les lecteurs cartographient le monde.

La géopoétique

Une troisième orientation, nommée géopoétique et née aux alentours de 1980, étudie les rapports entre l'espace et les formes et/ou genres littéraires. Elle interroge les conditions territoriales de la création et analyse en quoi la géographie conditionne l'écriture. C'est une théorie de la création littéraire, son approche est à la fois stylistique et poétique et s'attache aux formes littéraires qui modèlent l'image des lieux. Deux poètes, Michel Deguy d'abord, puis Kenneth White, promeuvent cette réflexion sur les relations qu'entretiennent la littérature et l'espace, le second en expose même les principes théoriques en définissant son *Plateau de l'Albatros* comme une *Introduction à la Géopoétique*³⁷. La littérature se doit de renouer les liens perdus avec le réel, de ne plus s'enfermer dans la clôture du texte : elle n'est pas seulement pour K. White un art du langage, elle implique une vision du monde. Les géopoètes pratiquent le nomadisme, tant intellectuel que physique, privilégient le mouvement et la transgression des frontières, qu'elles soient géographiques, linguistiques ou disciplinaires. Ils se projettent dans l'espace de la page aussi bien que dans celui du paysage. Langage, espace et écriture finissent par relever d'une expérience similaire de projection hors de soi et d'émergence d'un sujet en lien étroit avec ses expériences et graphiques et géographiques.

Cette démarche peut être rapprochée d'une tendance, parallèle pour ce qui est de l'attention portée à la nature, perçue comme primordiale, celle de l'écocritique. Le terme *ecocriticism* est utilisé pour la première fois en 1978 par William Rueckert, un essayiste américain : l'objet d'analyse de ce courant est la thématique de la nature d'un côté et la relation homme/nature et homme/environnement de l'autre. Il s'agit de renverser la perspective anthropocentriste et familière à la tradition occidentale. L'attention de l'*ecocriticism* se concentre ensuite vers les enjeux sociaux de la problématique environnementale. La recherche écocritique européenne, elle, tend à s'orienter vers une éco-poétique, plus soucieuse d'une dimension proprement littéraire des œuvres³⁸.

La géographie littéraire

La pensée-paysage de M. Collot³⁹ emploie les outils de la poétique pour travailler à l'espace en littérature, et étudie les structures spatiales et textuelles de la représentation, ainsi que leur relation à l'évolution des formes littéraires. Elle a des affinités avec la géopoétique, puisqu'elle s'interroge sur la manière dont la relation entre l'homme et son environnement se traduit dans les arts – elle dialogue aussi régulièrement avec la philosophie, la peinture... Collot porte l'accent sur la notion de paysage, qui nous invite à repenser notre appréhension schématique et formatée de l'espace. Contrairement à Descartes, à la vision duquel l'Occident reste figé, M. Collot n'oppose pas la chose pensante et la chose étendue, l'esprit et le corps. Or, le paysage les conjugue par sa dimension phénoménologique, lui qui résulte d'une interaction entre l'homme et l'espace. Dans le sillage de Merleau-Ponty, M. Collot considère la conscience comme corporelle et donc spatiale, et pour lui la vision d'un paysage renvoie toujours indirectement une image de soi. Pour le phénoménologue, un sujet ne peut regarder un objet sans que son propre corps soit enfermé dans la perspective. Comme la géopoétique, la pensée-paysage cherche dans l'ouverture au monde une confrontation et avec soi-même et avec autrui. Le paysage est de surcroît un espace *en mouvement*, en évolution, qui résiste aux classifications réductrices, ne serait-ce que grâce aux changements de perspective du spectateur. La pensée-paysage est donc une ego-géographie, qui s'élabore à partir de la confrontation d'un sujet et d'un espace. Cette lecture critique s'apparente à la critique de la conscience et de l'imaginaire, aux travaux de J.-P. Richard, G. Poulet ou encore Jean Starobinski. Contrairement à la géocritique, elle est favorable aux monographies et à l'étude des espaces imaginaires.

36 B. Westphal, *La Géocritique : réel, fiction, espace, op. cit.*, p. 150.

37 K. White, *Le Plateau de l'Albatros : Introduction à la géopoétique*, Grasset, 1994.

38 Voir à ce sujet S. Posthumus, *French Écocritique : Reading Contemporary French Theory and Fiction Ecologically*, Toronto : University of Toronto Press, 2017.

39 M. Collot, *La Pensée-paysage : philosophie, arts, littérature*, Arles : Actes Sud ; Versailles : ENSP, 2011.

Contrairement aux recherches linguistiques développées dans le cadre structuraliste, celles qui sont entreprises dans le vaste courant des sciences cognitives ont marqué, depuis les années 80 du siècle dernier, un véritable *tournant spatial*. Il est difficile, voire impossible, de parler d'une seule *linguistique cognitive*, comme on le fait pour l'école structuraliste, tant les approches, théories et méthodes sont multiples. Quelles que soient les positions théoriques adoptées, les linguistiques cognitives explorent, avec des outils nouveaux, les questions pourtant anciennes de relations entre les systèmes de connaissance, la pensée, la langue et la perception.

En tant qu'individus, nous existons dans l'espace, *eo ipso* nous y sommes localisés, nous nous y déplaçons, nous modifions l'emplacement des objets qui nous entourent et, de ce fait... l'espace environnant. Tout en le faisant, nous en avons une perception sensorielle, dont les résultats s'organisent en un système conceptuel. Autrement dit, les êtres humains conceptualisent l'espace, s'y repèrent, et plus encore, ils disposent de langues naturelles pour représenter les relations spatiales. C'est ce lien entre le système symbolique⁴⁰ qu'est une langue et les catégories conceptuelles qui s'est montré particulièrement stimulant pour les linguistes cognitivistes. Des études comparées réalisées sur des données provenant de langues différentes ont permis de montrer que celles-ci diffèrent sensiblement quant à l'encodage, tant lexical que grammatical, des paramètres spatiaux. Certaines, comme l'allemand, véhiculent dans le contenu sémantique des lexèmes verbaux non seulement l'information relative à la localisation d'un objet par rapport à un autre, mais également à leurs positions. Dans d'autres – en français par exemple – le moule lexical retient seulement le premier de ces paramètres, c'est-à-dire la localisation. L'observation des paramètres spatiaux qu'une langue lexicalise ou, au contraire, qu'elle exprime dans ses structures grammaticales a permis de repérer des tendances récurrentes et d'établir des typologies. Ainsi, selon celle proposée par Leonard Talmy⁴¹, le français est une langue à *trajectoire verbalisée* dans la mesure où le sémantisme du verbe contient des composantes spécifiant le parcours effectué, alors que l'expression de la manière dont on se déplace reste optionnelle. *Les enfants vont au cinéma* est une façon par défaut d'évoquer leur déplacement, qui, en cas de besoins communicatifs particuliers seulement, sera complétée par des précisions telles que *à pied, en courant, en bus* (ou autre moyen de transport). En affinant le classement de L. Talmy, Dan I. Slobin⁴² préfère placer le français parmi les langues à *basse saillance de manière*.

Même si, face à des données linguistiques extensives extraites de corpus, on s'aperçoit finalement que les contours de ces classements ne sont pas toujours clairs et nets⁴³, la question que se sont posée les linguistes – cognitivistes ou non – était celle des rapports entre une langue et la conceptualisation de l'espace : le pensons-nous, ou plus encore, le percevons-nous conformément aux moyens linguistiques à notre disposition ? Les travaux des psycholinguistes Soonija Choi et Melissa Bowerman⁴⁴ ont montré que, lors du processus d'acquisition du langage, des enfants âgés de 17 à 20 mois, loin de construire les concepts spatiaux à partir d'expériences directes, sont sensibles aux moules lexicaux qu'offrent leurs langues respectives. Pour dire les choses simplement : en décrivant une même configuration spatiale, les enfants coréens et anglophones mettent en avant des paramètres différents. Les premiers séparent nettement les déplacements spontanés de ceux qui sont causés par un facteur externe, ce qui n'est pas le cas des seconds.

C'est parce que, explique D. I. Slobin, en tant que locuteurs, nous devenons plus attentifs aux catégories que notre langue maternelle nous conduit à exprimer. En observant une configuration spatiale ou un déplacement, nous cherchons à en repérer certaines caractéristiques, en en délaissant d'autres, moins ou pas du tout pertinentes du point de vue linguistique. L'hypothèse de D. I. Slobin consiste à admettre que nous pensons pour parler

40 R. W. Langacker, *Foundations of Cognitive Grammar. Theoretical Prerequisites*, Stanford : Stanford University Press, 1987.

41 L. Talmy, « Lexicalization patterns : semantic structure in lexical forms » dans *Language typology and syntactic description. Grammatical categories and the lexicon*, (dir. T. Shopen), Cambridge : Cambridge University Press, 1985, vol. III, p. 57–149, voir aussi. L. Talmy, *Toward a Cognitive Semantics*, Harvard, MIT Press, 2000.

42 D. I. Slobin, « The Many Ways to Search for a Frog : Linguistic Typology and the Expression of Motion Events » dans *Relation Events in Narrative : Typological and Contextual Perspectives* (dir. S. Strömquist et L. Verhoeven), Mahwah (NJ) : Lawrence Erlbaum Associates, 2004, p. 219-257. Il est important de souligner que la thèse de D. I. Slobin n'a rien du déterminisme whorfien. B. Whorf (« Language, Mind, and Reality », 1942, web.stanford.edu/dept/SUL/library/extra4/sloan/mousesite/Secondary/Whorfframe3.html, consulté le 31 mai 2017) soutenait en effet que la langue forme une grille de perception. Elle empêcherait le locuteur de percevoir et de conceptualiser ce pour quoi elle n'a pas de moyens d'expression explicites. D. I. Slobin considère que les ressources linguistiques dont dispose un locuteur le conduisent à observer et à conceptualiser les situations en anticipant les formulations qu'il adoptera.

43 Voir par exemple A. Kopecka, « L'expression du déplacement en français : l'interaction des facteurs sémantiques, aspectuels et pragmatiques dans la construction du sens spatial », *Langages*, 173, 2009, p. 54 -73.

44 S. Choi et M. Bowerman, « Learning to express motion events in English and Korean », *Cognition*, 41, 1991, p. 83-121.

(*thinking for speaking*), ce qui résulte des styles narratifs différents que développent les locuteurs selon la langue qu'ils parlent⁴⁵.

L'espace intéresse les linguistes également par ce que l'on pourrait appeler son potentiel métaphorique. Des notions spatiales telles que la localisation et le déplacement constituent une source de métaphores conceptuelles avec des domaines cibles multiples. Plusieurs se sont lexicalisées dans le cours de l'histoire d'une langue (voir l'article de Georges Kleiber pour ce qui est de l'*espace* lui-même), pour désigner des réalités qui n'ont pas un caractère spatial : *champ*, *zone*, *parcours*. Certaines, *espace littéraire* par exemple, sont devenues des concepts scientifiques. D'autres encore restent vivantes, comme celles des relations amoureuses appréhendées comme un voyage, pour reprendre l'un des exemples de George Lakoff⁴⁶.

LE ROMAN CONTEMPORAIN ET L'ESPACE

Romanciers de l'extrême-contemporain et espace

On ne peut que mettre en regard le développement de l'étude critique de l'espace et la mutation progressive des formes littéraires, comme Joseph Frank l'avait souligné dès 1945⁴⁷. Plus récemment Marc Brosseau, dans son ouvrage intitulé *Des romans-géographes*⁴⁸, consacré au roman contemporain, fait valoir l'existence d'une pensée spatiale du roman depuis les années 70. On ne peut manquer d'observer, maints chercheurs l'ont souligné, une rupture progressive dans les textes narratifs, tout au long de XX^e siècle, avec l'axe linéaire du récit et le devenir des personnages, au profit d'une attention plus grande au contexte et à la périphérie. Julien Gracq, écrivain-géographe, en est un des exemples les plus marquants, lui dont les paysages finissent par envahir tout le champ. À l'ère du nouveau roman, M. Butor abandonne une forme pourtant novatrice au profit d'une écriture qui se veut pure exploration de l'espace, selon un principe à la fois typographique et thématique⁴⁹. De même, beaucoup de récits contemporains pourraient se définir comme des *récits d'espace*, pour reprendre la formule de M. de Certeau⁵⁰. Georges Perec apparaît de ce point de vue, avec *Espèces d'espaces* notamment, comme un précurseur et un modèle. Dans une perspective plus sociologique et politique, François Bon, dans *Paysage fer*⁵¹, suit le tracé du train qui le mène en Lorraine pour structurer le fil de sa narration, démarche souvent reprise par la suite⁵². On note par ailleurs une véritable recrudescence de la littérature de voyage – et parallèlement d'études sur ce type de littérature. La parution de *L'Usage du monde* de Nicolas Bouvier en 1985, du *Petit Guide pédestre de la littérature française au XVII^e siècle* de Michel Chaillou en 1990, et, plus près du nous, du *Dépaysement* de Jean-Christophe Bailly, est à cet égard éclairante. Aline Bergé-Joonekindt parle à ce sujet d'« un véritable tournant paysager⁵³ » dans le roman d'après 1980, qui accompagne le *retour au réel*.

Quelques grandes tendances sont repérables : certains, tels Pierre Bergougnoux, Pierre Michon, Jean-Loup Trassard ou Richard Millet, disent et redisent leur terre natale, font revivre l'idée d'un terroir qui modèle l'identité et s'oppose à une mondialisation par trop uniformisante. D'autres, comme Patrick Modiano, sont d'inlassables arpenteurs, sillonnant Paris, les lieux du quotidien, d'une démarche inquiète et questionneuse ; cela est vrai aussi de François Bon et d'Annie Ernaux, mais chez eux le paysage se colore d'une dimension sociologique et politique, et privilégie banlieues et espaces périphériques ; Jean-Marie-Gustave Le Clézio, Jean Rolin, Jean-Philippe Toussaint, Christian Garcin, explorent le monde selon leurs nécessités singulières. Antoine Volodine et Michel Houellebecq, chacun à leur manière, voyagent dans des territoires dystopiques. Pierre Senges, Michel Chaillou, Pascal Quignard, Jean Echenoz et bien d'autres sillonnent des espaces chargés d'histoire et de

45 D. I. Slobin, *op. cit.*

46 Nous pensons bien évidemment à l'influent ouvrage de G. Lakoff et M. Johnson *Metaphors we live by* (Chicago : Chicago University Press, 1980). Dans un registre plus décontracté, G. Lakoff cite cet exemple dans une conférence que l'on retrouvera sur www.youtube.com/watch?v=Eu-9rpJITY8 (consultée le 30 octobre 2019).

47 Voir la note 7.

48 M. Brosseau, *Des Romans-géographes*, L'Harmattan, 1996.

49 On pense bien sûr à des ouvrages expérimentaux comme *Mobiles* (1962).

50 M. de Certeau, *L'Invention du quotidien* [1980], Gallimard, coll. « Folio essai », 1990, p.170 et sq.

51 F. Bon, *Paysage fer*, Lagrasse, Verdier, 1999.

52 Voir « Les non-lieux littéraires comme lieux rhétoriques », *Ville infectée, ville déshumanisée*, S. Freymuth, J. F. P. Bonnot, T. Obergöker (dir.), Peter Lang, coll. « Comparatisme et société », vol. 29, 2014, p. 261 (note 44).

53 Voir A. Bergé-Joonekindt, « le tournant paysager de la littérature contemporaine. Une traversée des modernités », dans A ; Bergé et M. Collot (dir.), *Paysage et modernité(s)*, Bruxelles, Éditions Ousia, 2007, p. 89-101.

littérature. Tant d'autres mériteraient d'être évoqués, et le seront fort heureusement pour quelques-uns dans les différents chapitres.

Perspectives critiques du recueil

Cet ouvrage se propose d'apporter un prolongement aux travaux touchant aux liens qui unissent espace et roman contemporain⁵⁴, et de suggérer quelques pistes nouvelles. Dans la mouvance tour à tour de la géocritique, de la géographie littéraire, des approches narratologiques ou sociologiques, des chercheurs français et étrangers s'emploient dans ce recueil à penser ce que la production romanesque d'après 1980 dit de la perception contemporaine de l'espace, étudiant les lieux et les *non-lieux*, que M. Augé⁵⁵ a définis avec tant de pertinence. Si le principe de référentialité n'a pas été exclu, les différentes interventions ont considéré que les romans ne renvoyaient pas directement au réel mais à une expérience singulière de celui-ci, une récréation, une transposition. La postmodernité remet en question dans bien des domaines l'opposition classique du vrai et du faux, du factuel et du fictionnel, et perçoit pour cette raison la littérature comme un moyen fondamental pour penser l'expérience humaine et l'espace où elle se déploie⁵⁶. L'expérience spatiale étant toujours celle d'un sujet, elle est aussi toujours inscrite dans une culture. E. Soja évoque dans cette optique le *thirdspace*, qui contrairement aux deux premiers modes de perception que sont le *firstspace* et le *secondspace*, est un espace de *représentation*, de construction subjective, où les identités et les héritages culturels individuels se confrontent, un espace nimbé d'imaginaire et de symboliques antérieures pour celui qui le considère. C'est cette construction que l'ouvrage se propose d'étudier dans le roman contemporain, la façon dont il s'approprie l'espace culturel et en restitue l'image. L'univers de chaque romancier sera donc conçu, plutôt que comme un objet mimétique – on sait la vanité de telles conceptions – comme un *dispositif* qui porte une pensée sur le monde.

Une première partie, consacrée à une *approche linguistique*, tiendra lieu de réflexion préalable sur la façon dont la langue configure l'espace. Dans le discours, la matière linguistique sert à représenter des lieux et des relations spatiales, notamment le déplacement, qui peut se définir comme modification de localisations successives par rapport à un ou plusieurs repères. C'est également ce que fait souvent le texte littéraire. Cependant, bien plus que les autres formes de discours, il crée plutôt qu'il ne recrée un espace, tout aussi fictionnel que les personnages et les objets qui en font partie. Comme eux, il participe d'un univers romanesque. Ainsi, pour explorer la – ou plus exactement – les dimensions spatiales du roman contemporain français, ce volume s'ouvre par une enquête sur le mot *espace* lui-même. Georges Kleiber analyse ce nom qu'il qualifie de *sommital* : il appartient à une catégorie d'hyperonymes très généraux, qui recouvrent des domaines d'expérience entiers. On le retrouve dans le sens de *lieu*, d'endroit, de *place*, et de bien d'autres, non sans éviter une certaine circularité, comme le pointe G. Kleiber. Mais qu'est-ce que l'*espace* d'un point de vue sémantique ? « Un *Janus bifrons* » – constate le linguiste, une véritable énigme sémiotique, avec une face à la fois familière, puisque nous savons tous de quoi il s'agit, et mystérieuse, dès que l'on s'arrête pour en fixer les traits. Pour G. Kleiber, ce constat de difficulté n'est pourtant qu'un point de départ pour une enquête minutieuse, qui s'appuie sur des indices disséminés dans la langue elle-même.

Le texte de Dorota Sikora est un exemple d'application d'outils linguistiques à l'analyse d'un roman. Dès la première lecture, *Les Âmes grises*, roman de Philippe Claudel apparaît comme une œuvre éminemment spatiale. Sa narration joue avec les catégories de proximité et d'éloignement, qui coïncident avec ce qui est connu et inconnu pour les personnages de l'univers romanesque. En partant de l'organisation globale de l'espace dans le roman de P. Claudel, D. Sikora analyse comment, strate par strate, les choix lexicaux, les procédés syntaxiques de détermination définie, la structure anaphorique des références spatiales contribuent à découper dans le territoire familier du narrateur des segments spatiaux de plus en plus resserrés et refermés sur eux-mêmes. Cependant, ce

54 Voir B. Blanckeman et B. Havercroft (dir.), *Narrations d'un nouveau siècle (romans et récits français, 2001-2010)*, Presses de la Sorbonne Nouvelle, 2013 ; J. Baetens et D. Viart (dir.), *États du roman contemporain*, T. 2., Minard, 1999 ; M. Collot, *Pour une géographie littéraire*, Corti, 2014 ; S. Coyault-Dublanchet, *La province en héritage. Pierre Michon, Pierre Bergounioux, Richard Millet*, Genève : Droz, 2002 ; F. de Chalonge (dir.), *Espace et récit de fiction*, Villeneuve d'Ascq : Presses universitaires du Septentrion, 2005 ; M. Dambre et W. Asholt (dir.), *Un retour des normes romanesques dans la littérature française contemporaine*, Presses de la Sorbonne Nouvelle, 2010 ; B. Westphal, *Le Monde plausible. Espace, lieu, carte*, Minuit, 2011. Des indications bibliographiques renverront dans ces livres à des ouvrages ou articles que nous ne pouvons pas tous évoquer.

55 M. Augé, *Non-Lieux, Introduction à une anthropologie de la surmodernité*, Seuil, 1992.

56 « [...] c'est la littérature postmoderne qui s'adapte le mieux à cette nouvelle version du réel, le 'réel déréalisé' ; c'est peut-être elle qui offre les meilleures options de lecture du monde, en vertu de sa fictionnalité même », B. Westphal, *La géocritique, op. cit.*, p. 150.

texte romanesque ne fait pas que créer un espace fictionnel : par un jeu polyphonique, le lecteur, lui aussi, y trouve sa place.

Une seconde partie, *Les lieux du contemporain*, cherche à caractériser la façon dont le roman s'approprie quelques espaces perçus comme emblématiques de notre époque. Bruno Blanckeman, travaillant à un vaste corpus d'œuvres récentes, retrace les configurations majeures de l'espace dans le roman contemporain, présenté tour à tour comme une coquille culturelle, un territoire marginal, dystopique, ou encore un lieu d'identité et de mémoire. Florence de Chalonge s'intéresse à l'usage singulier du topos du *regard à la fenêtre* chez Jean-Philippe Toussaint ; la fenêtre, lieu frontière favorisant les échappées et les intrusions, devient dans *Nue* une balise spatiale de la narration, mais aussi une façon d'articuler le visible et l'invisible, le présent et le passé. Laura Tudoras, étudiant *Les Heures souterraines* de Delphine de Vigan et *Disparaître* d'Étienne Ruhaud, cartographie les réalités urbaines surmodernes pour en désigner les aberrations et montrer comment elles menacent l'individu dans son identité par des mécanismes d'inclusion-exclusion. Isabelle Dangy, en s'appuyant sur quelques romans emblématiques, rappelle combien, dans le prolongement du postjaponisme, le Japon occupe toujours une place importante dans l'imaginaire français contemporain. Elle se concentre sur la représentation de l'espace domestique japonais, défini comme un espace allégorique.

La troisième partie, *À l'ère de la mobilité*, se veut une réflexion sur le lien que les romanciers actuels entretiennent avec le mouvement et les modes de traversée de l'espace. Anne Cousseau, réfléchissant à la mondialisation, qui fait de notre terre un « village global », selon la formule de Herbert Marshall McLuhan, analyse le roman comme caisse de résonance de ces mutations et les étudie sous l'angle du territoire, de l'ailleurs, du flux et de la mobilité, en travaillant plus particulièrement à *Naissance d'un pont* de Maylis de Kerangal et *L'invention des corps* de Pierre Ducrozet. Matthieu Freyheit étudie la métaphore voyageuse qui caractérise la période contemporaine, favorisée par le développement du digital ; celle-ci dérive d'une mythologie inspirée de l'imaginaire américain. Par l'analyse de *L'invention des corps* de Pierre Ducrozet, il cartographie le territoire états-unien tel qu'il se dit dans le roman, espace de nomadisme et d'errance, entre fantasme et aporie. *Terminal Frigo* de Jean Rolin, commenté par Pierre Hyppolite, retrace un cheminement de port en port, selon un protocole déambulatoire n'obéissant pas à la logique géographique mais à une démarche subjective. Jean Rolin nous offre une herméneutique de ces espaces frontières que sont les villes portuaires, par une mise en récit de perceptions spatiales et de territoires palimpsestes, dont la configuration contemporaine ne peut s'expliquer que par leur histoire singulière. Prenant pour objet les logiques spatiales contemporaines, Émilie Ieven et Manon Delcour questionnent la figure de l'écrivain flâneur et arpenteur dans *Caprice de la reine* de Jean Echenoz et *Une vie en l'air* de Philippe Vasset. Valeria Gramigna travaille à l'écriture nomade de Maylis de Kerangal, interrogeant chez elle la question du mouvement, qui participe de la construction d'une identité, d'une création-documentaire sur le monde, d'un nouveau mode de relations humaines ou encore de la perception de l'espace.

Une dernière partie enfin, *Le Monde est une bibliothèque*, étudie les univers reposant sur des strates mémorielles, construits en référence à des espaces antérieurs, littéraires ou artistiques, façon de réintégrer de la profondeur temporelle dans l'espace. Perspective qui rejoint ce que Bruno Blanckeman nomme les « fictions spéculaires⁵⁷ » : « Plusieurs œuvres s'affichent ainsi comme des écritures-lectures qui questionnent l'univers de références directes dans lequel elles s'inscrivent (le monde d'aujourd'hui) autant que celui des références littéraires dont elles sont issues (la bibliothèque) [...] ⁵⁸ ». Catherine Haman s'attache aux paysages qui hantent l'œuvre de Tanguy Viel, à leur charge symbolique, construite sur des oppositions et des frontières marquées, et restituées à travers le filtre d'un espace mental. Elle analyse aussi la forte présence de strates culturelles : par le biais d'un paradigme indiciaire, Tanguy Viel revisite certains *lieux communs*, affirmant par là son appartenance à un espace littéraire et intermédial. Quand Alain Corneau demande un scénario original pour un film de fiction consacré aux musiciens Sainte-Colombe et Marin Marais, P. Quignard refuse, mais écrit un roman qu'ils adapteront ensemble ; Carl Veters travaille sur le passage du signe linguistique au signe iconique, et expose comment l'espace du roman, entièrement construit dans l'optique d'une sobriété janséniste, est ensuite rendu à l'image grâce à des choix esthétiques et techniques concertés. Chiara Rolla analyse comment l'œuvre érudite et encyclopédique de P. Senegès se fraye un chemin à travers les textes et les savoirs. Véritable odyssee littéraire, *Environs et mesures*

57 B. Blanckeman, « La littérature française au début du XXI^{ème} siècle : tendances en cours », dans Michèle Touret (dir.), *Histoire de la littérature française du XX^{ème} siècle*, Presses Universitaires de Rennes, coll. « Histoire de la littérature française », 2008, p. 451.

58 *Ibid.*

vagabonde dans un passé et des espaces fictionnels devenant le lieu d'un dialogue entre hier et aujourd'hui, ici et ailleurs. Seneg s'emploie particulièrement à dénoncer, en alternant le grave et le facétieux, quelques-unes des conséquences les plus désastreuses de la soif de circonscrire le monde.

PARTITION DE L'ESPACE DANS *LES ÂMES GRISSES* DE PHILIPPE CLAUDEL

Dorota SIKORA

INTRODUCTION : OBJECTIFS, DIFFICULTÉS ET QUESTIONS DE MÉTHODE

Questions et objectifs de recherche

Dès la première lecture, le roman de Philippe Claudel apparaît comme éminemment spatial. Certes, certaines séquences descriptives contribuent à construire l'espace dans lequel évoluent « les âmes grises ». Les séquences descriptives offrent par exemples des évocations du cours d'eau, témoin des événements : « Sur la droite du parc musarde une rivière étroite et sinueuse, la Guerlante, dont le nom rend bien compte de son cours ralenti tout en tourbillons et nénuphars. L'eau imprègne tout¹ ».

Le titre choisi par le traducteur américain² confirme l'intuition de lecteur non averti : *By the Slow River* réfère non pas aux habitants, mais aux lieux qui constituent le théâtre des événements racontés. D'un point de vue de linguiste (qui est le mien), *by a slow river*, tout comme *sur la droite du parc*, est un syntagme prépositionnel Préposition + Syntagme nominal (*Prép* + *SN*) à valeur localisatrice. Il sert à situer des individus, des objets et des faits dans un espace de référence. Or, selon Leonard Talmy, la procédure inhérente à toute opération de localisation « involves processes of dividing a space into subregions or segmenting it along its contours, so as to “narrow in” on an object's immediate environment³ ».

Cependant, ces expressions localisatrices disséminées aussi bien dans des séquences descriptives que narratives ne sont pas les seules responsables des intuitions spatiales qui accompagnent la lecture du roman. Elles reposent en fait sur des indices linguistiques, qui relèvent de plusieurs niveaux d'analyse : lexical, syntaxique, discursif, non sans impliquer des phénomènes de polyphonie. La présente étude retrace une enquête dont l'objectif était de repérer et d'identifier les moyens linguistiques mis à contribution tout au long du roman non seulement pour fixer les repères spatiaux de l'univers romanesque, mais pour y attirer et y intégrer le lecteur.

L'espace fictionnel et son analyse linguistique : problèmes épistémologiques et méthodologiques

En soumettant un texte littéraire à une analyse linguistique, plusieurs difficultés sont à envisager. La première est d'ordre épistémologique et elle est liée au traitement référentiel des entités fictionnelles. En effet, dans la section suivante, il sera question de l'organisation globale de l'espace dans *Les Âmes grises*. On s'y repère par rapport aux lieux familiers pour le narrateur, à la fois énonciateur et acteur des événements : il s'agit donc de la *déixis* personnelle et spatiale. Il nous faut connaître des paramètres de la situation d'énonciation – la personne et la position de l'énonciateur pour identifier les lieux qu'il évoque. Mais qu'en est-il lorsque l'espace n'est situé nulle part, puisqu'il fait partie d'un univers fictif créé dans le roman ? Peut-on dès lors parler de centre déictique qui se constitue du point de vue de narrateur, alors que par ailleurs, la fonction référentielle d'un roman est assurée par des moyens anaphoriques⁴ ?

La réponse adoptée dans cette étude est celle que propose Nelson Goodman⁵ : une expression linguistique telle que « le Château » réfère non pas à une entité réelle, mais à la description du Château dans le roman. Une telle solution permet entre autres de parler de déixis mémorielle interne, puisqu'elle opère sur une représentation partagée entre le narrateur et le lecteur construite par le texte.

1 Ph. Claudel, *Les Âmes grises*, Éditions Stock, 2003, p. 31-32. Tout au long de cet article, on se référera à cette même édition du roman.

2 Il s'agit de la traduction par H. Rogers, *By the Slow River*, New York, Anchor Books, 2006.

3 L. Talmy, *Toward a Cognitive Semantics. Concept Structuring System*, vol. 1, Harvard, MIT Press, 2000, p. 185-186.

4 En analyse linguistique de discours, l'anaphore désigne non pas une figure de style, mais une organisation particulière de la chaîne de référence (ces maillons sont indiqués en gras dans l'exemple qui suit). Dans la séquence *Mon voisin élève des carpes, elles sont très grasses*, le pronom *elles* est une reprise anaphorique du SN *des carpes* (son antécédent). C'est grâce à l'antécédent *carpes* que le destinataire identifie le référent de *elles*. Un roman peut être considéré comme un genre par excellence anaphorique, car le plus souvent tous les éléments nécessaires pour trouver les référents des expressions employées sont fournies dans le texte qui les précède.

5 N. Goodman, *Languages of Art*, 2^e éd., Indianapolis, Hackett Publishing Co, 1968, traduction française par J. Morizot, *Langages de l'art*, Nîmes, Éditions Chambon, 1990.

Si la question de la référence dans un univers romanesque fictionnel trouve une robuste réponse philosophique, de nombreuses difficultés méthodologiques dans le traitement de la matière textuelle elle-même obligent à avancer avec prudence. Premièrement, *Les Âmes grises* n'est pas disponible en version numérisée en corpus : ce roman est notamment absent de la base textuelle Frantext⁶. On ne peut dès lors que repérer et décortiquer des phénomènes linguistiques « à la main », au fil des lectures et de relectures, avec un risque d'omissions et d'oublis inhérent, lorsque délaissant sa casquette de chercheur, le linguiste redevient lecteur pris dans l'action.

Deuxièmement, les moyens que l'auteur-scripteur met en œuvre par exemple pour représenter l'espace, relèvent rarement d'un seul niveau d'analyse. C'est dire que leur analyse doit naviguer entre la sémantique lexicale, la syntaxe, des procédés discursifs et les multiples voix résonnant dans le texte d'un roman. Cette multitude de faits de langue qui interagissent d'une part, et les fonctions esthétiques que ces procédés contribuent à assurer d'autre part, rendent difficile l'application de tests – courants en linguistique – destinés à vérifier les propriétés des lexèmes et des constructions étudiés. En effet, dans *Ils vont dans les Vosges*, l'impossibilité de détacher le complément en tête de phrase montre son caractère essentiel, contrairement à *Ils campent dans les Vosges* où son déplacement facile montre qu'il s'agit d'un ajout. Or, les manipulations de ce type sont pour le moins problématiques dans les exemples littéraires, notamment ceux qui apparaissent dans cette étude.

Consciente de ces écueils, et en soupçonnant bien d'autres, cette réflexion illustre une approche qualifiée par Oswald Ducrot de coopérative⁷ : « Une interprétation exige toujours que l'on mette en œuvre des connaissances que l'on possède indépendamment de l'objet à interpréter, et qu'on les applique à cet objet⁸ ». Notre objet, le texte des *Âmes grises* est ainsi exploré à l'aide de notions et d'outils conceptuels élaborés grâce à de nombreux travaux menés depuis des années en linguistique pour vérifier et expliquer le caractère éminemment spatial du roman.

ORGANISATION GLOBALE DE L'ESPACE DANS *LES ÂMES GRISSES*

Et pourtant le monde n'était pas loin : il suffisait pour le voir de monter sur le coteau⁹.

Avant d'analyser les moyens linguistiques spécifiques, observons comment s'organise l'espace dans l'univers romanesque ou, plus précisément sa représentation partagée par le narrateur et le lecteur, certes, mais également, selon l'hypothèse qui sera discutée plus loin, par l'auteur-scripteur.

Trois lieux-repères dans l'univers du roman

Dans le roman de Claudel, l'espace s'organise autour des trois lieux suivants :

- i P., « la petite ville », « notre ville » est située dans un territoire plus large évoqué par le complément localisateur « chez nous ».
- ii V., présentée comme une grande ville où les habitants de P. se rendent parfois en raison de diverses activités professionnelles et sociales.
- iii La ligne de front et le terrain d'opérations militaires de la Première Guerre mondiale, qui fait partie d'un espace plus large désigné comme « le monde ».

Les extraits (1) à (8) ci-dessous permettent de repérer les paramètres sur lesquels s'appuie l'organisation globale de l'espace des *Âmes grises* :

- 1) *V. est distant de chez nous d'une vingtaine de kilomètres*. Une vingtaine de kilomètres en 1917, c'était un monde déjà, surtout en hiver, surtout avec cette guerre qui n'en finissait pas et qui nous amenait un grand fracas sur les routes, de camions et de charrettes à bras, et *des fumées puantes* ainsi que des coups de tonnerre par milliers car *le front n'était pas loin*, même si de là où nous étions, c'était pour nous comme un monstre invisible, un pays caché¹⁰.
- 2) *Notre ville* n'est pas très grosse. Ce n'est pas V., loin s'en faut. Pourtant, on peut s'y perdre. J'entends par là qu'elle comporte suffisamment de coins d'ombre et de belvédères pour que tout un chacun puisse y trouver de quoi flatter sa mélancolie¹¹.

6 Voir : Base textuelle Frantext (www.frantext.fr) dans son état de développement d'août 2019.

7 O. Ducrot, « Sémantique linguistique et analyse de textes », *Littérature* n° 115 (3), 1999, p. 104-25.

8 *Ibidem*, p. 106.

9 Ph. Claudel, *op. cit.*, p. 83.

10 *Ibid.*, p.12. Tous les soulignements dans les extraits cités sont de mon fait.

11 *Ibid.*, p. 28-29.

- 3) Elle [Lysia] regardait droit devant elle, ce que nous autres ne voulions jamais voir. (...) Elle regardait la large plaine, brune et infinie, tremblotante sous *les fumées lointaines des explosions dont la fureur venait à nous* comme amortie et décantée, irréaliste pour tout dire¹².
- 4) *Au loin*, la ligne de front se confondait avec celle du ciel, si bien que par moments, on aurait cru que de multiples soleils se levaient en même temps, puis retombaient dans un bruit de pétard avorté¹³.
- 5) [...] *à moins de quinze lieues* on s'égorgeait à l'arme blanche [...]¹⁴
- 6) [Lettre de Lysia à Bastien Francœur, Le 13 décembre 1914] : Je suis enfin *proche de toi*. Je suis arrivée aujourd'hui à P., une petite ville qui n'est qu'*à quelques kilomètres du front* où tu te trouves¹⁵.
- 7) Notre petite ville *entendit* la guerre mais ne la fit pas vraiment¹⁶.
- 8) Bien sûr, la guerre, on *l'entendait*. On *l'avait vue annoncée* sur les placards de la mobilisation. On la lisait dans les journaux. Mais au fond, on feintait, on s'arrangeait avec elle, comme on fait avec les mauvais rêves et les âpres souvenirs. Elle n'était pas trop de notre monde¹⁷.

Proximité et éloignement : distance géographique ?

Seule la distance entre P. et V., bien qu'indiquée de manière approximative, garde une apparence d'objectivité, puisqu'elle est exprimée en kilomètres (exemple 1). Les voix de différents énonciateurs discordent en revanche, lorsqu'il s'agit d'expliquer où se trouve la ligne du front : selon le narrateur dans les extraits (3) et (4), elle se situe au loin, alors qu'au même moment du temps du récit, Lysia Verharaine s'en dit très proche, à quelques kilomètres à peine.

À y regarder de plus près, les informations fournies par le narrateurs sont contradictoires : tantôt, la guerre est près (extrait 1), tantôt elle est présentée comme lointaine (3, 4). Loin d'éclaircir la situation, les données quantifiées brouillent encore les pistes à cause de l'unité de mesure employée (« à moins de quinze lieues » dans 5). En effet, la distance qui correspond à une lieue n'est pas la même selon l'époque et la région : elle peut aller de 3,3 (lieue de Beauce) à 5,8 kilomètres (lieue de Gascogne)¹⁸. Il s'agit donc d'une mesure à valeur locale. Faute d'indication plus précise dans le roman, un simple calcul s'appuyant sur la lieue de Beauce conduit à situer le front à une cinquantaine de kilomètres de P. – nettement plus loin que les « quelques kilomètres » mentionnés par Lysia. Qu'elle soit considérée comme proche ou lointaine, la guerre reste néanmoins perceptible aux habitants par les bruits et les odeurs qui parviennent jusqu'à la petite ville (extraits 3, 7 et 8).

Ce jeu de pistes brouillées indique que la proximité et l'éloignement ne sont pas, et ne doivent pas être évalués à partir de paramètres topologiques objectifs. Les repères qui servent à configurer l'espace proximal et distal¹⁹ dans le roman sont indissociables du personnage du narrateur et de ceux avec qui il s'associe en recourant régulièrement au pronom « nous » dans ses différentes formes. Or, comme on le sait depuis Émile Benveniste²⁰, « nous » offre plusieurs configurations de la référence plurielle, toujours organisées autour de « je » : je + tu, je + vous, je + ils, etc. C'est donc l'inclusion de « je » qui confère à « nous » son caractère déictique. Dans la mesure où le « je » impliqué dans « nous » et « notre ville » est celui du narrateur, habitant « du pays » et acteur des événements, on peut admettre que ce pronom est employé pour désigner « je » et ses compatriotes. Ainsi, le « je » du narrateur, associé à l'ensemble des habitants de P., à leurs activités et aux événements qui s'y déroulent permet d'identifier l'origo²¹, c'est-à-dire le point de référence essentiel de l'organisation spatiale du roman : elle a à la fois un caractère déictique et actoriel, car elle est structurée autour du narrateur-acteur.

La question qui émerge dès lors est celle des paramètres selon lesquels s'organise l'espace proximal actoriel du roman. Les données géographiques étant peu fiables, nous l'avons vu, il convient d'identifier celles qui dans l'univers romanesque sont décisives pour considérer les trois lieux repères – les villes de P. et de V., ainsi que la ligne de front – comme proches ou distants. Le tableau 1 ci-dessous réunit les paramètres dont le narrateur semble tenir compte.

12 *Ibid.*, p. 84.

13 *Ibid.*, p. 84-85.

14 *Ibid.*, p. 57.

15 *Ibid.* p. 248.

16 *Ibid.*, p. 48.

17 *Ibid.*, p. 135.

18 fr.wikipedia.org/wiki/Lieue (consulté le 20 septembre 2018).

19 H. Diessel, *Demonstratives: form, function, and grammaticalization*, Amsterdam et Philadelphia, John Benjamins, 1999.

20 É. Benveniste, « De la subjectivité dans le langage », dans *Problèmes de linguistique générale*, vol. 1, Gallimard, 1966, p. 258-266.

21 K. Bühler, *Sprachtheorie. Die Darstellungsfunktion der Sprache*, Iena, Verlag von Gustav Fischer, 1934.

Tableau 1: Organisation déictique de l'espace dans *Les Âmes grises*.

| | | P. (<i>notre ville</i>) | V. | front |
|---|------------------------|---------------------------|------|-------|
| 1 | proximité géographique | origo | loin | ? |
| 2 | visibilité | oui | non | oui |
| 3 | perceptibilité | oui | non | oui |
| 4 | présence physique | oui | oui | non |
| 5 | expériences vécues | oui | oui | non |

La ville de P., lieu de vie du narrateur, constitue l'origo ou le centre déictique. Même si le théâtre des opérations militaires se trouve suffisamment près pour être perçu par les sens – outre la vue des explosions, ses bruits et ses odeurs parviennent bien jusqu'à P. – aucun de ses habitants ne s'y est jamais rendu, n'y a été physiquement présent. La ligne de front et la guerre n'ont fait l'objet d'aucune expérience personnelle de la part des habitants. En revanche, bien que considérée comme éloignée, la ville de V. constitue la destination des déplacements réguliers, souvent longs et pénibles, des habitants de P., y compris du narrateur. Les deux localités leur sont donc également familières et, par conséquent, perçues comme proches. Chacun des membres de la communauté désignée par « nous » y a été physiquement présent et a pu les connaître par expérience directe.

Un proximal aux contours flous

Le territoire proximal dans lequel évoluent les habitants de P. est présenté comme petit, resserré autour de ses occupants, renfermé sur lui-même (voir l'extrait 10 ci-dessous). On se pose dès lors la question de ses limites : où, à quel endroit précis le proximal cesse de l'être pour devenir distant ? En effet, la seule frontière explicitement tracée entre *chez nous* et le territoire distal de la guerre passe par la colline (chapitre IX).

L'espace proximal est évoqué dans le roman par des syntagmes nominaux (*notre ville, notre petite ville, le pays*) et prépositionnels (*chez nous, par ici, dans notre pays*). Ce répertoire relativement limité est intéressant par le jeu des limites qu'il offre au lecteur en raison du contraste sémantique entre les lexèmes VILLE I d'une part et PAYS II.1 ('partie familière d'une région'), ICI de l'autre²².

Dans son sens lexical, VILLE I renvoie à un espace contenu dans des limites précises. Il est possible de le vérifier en combinant ce lexème avec toute une série d'expressions (présentées en 9a) impliquant une frontière. De telles constructions semblent problématiques pour PAYS II.1. S'il est possible de dire *dans le pays*, les locutions prépositionnelles EN DEHORS, À L'EXTÉRIEUR, À L'INTÉRIEUR, ainsi que les verbes tels que ENTRER et SORTIR marquant une frontière nette²³ ne sont pas acceptables (voir 9b). De même, si *en dehors de chez nous* est un syntagme parfaitement acceptable, *chez nous* résiste aux autres expressions testant les limites. ICI, quant à lui, n'autorise aucune d'entre elles.

- | | | | |
|----|--|---------------|--|
| 9) | a. en dehors de la ville à l'extérieur de la ville à l'intérieur de la ville entrer dans la ville sortir de la ville | <i>versus</i> | b. ? en dehors de notre pays (II.1) ? à l'extérieur de notre pays (II.1) ? à l'intérieur de notre pays (II.1) ? entrer au/dans notre pays (II.1) ? sortir de notre pays (II.1) |
|----|--|---------------|--|

Ainsi, dans le proximal aux contours indéterminés se découpe le territoire de la ville de P. aux frontières bien nettes. Ce découpage se poursuit par une partition de la ville et par un resserrement des distances. Le chapitre IX apporte la séquence descriptive suivante :

22 Comme dans un dictionnaire, le numéro précise l'acception dont il s'agit. Cette analyse repose sur la description lexicographique réalisée dans le *Réseau Lexical du Français (RL-fr)*, où PAYS II.1 dénote non pas l'unité administrative (PAYS I.1), mais partie familière d'une région. Les conventions typographiques propres à cette approche linguistique sont adoptées : en tant qu'objets de descriptions, les lexèmes sont indiqués en petites majuscules. Les syntagmes, qui sont des combinaisons de lexèmes, évoqués dans notre réflexion sont marqués en italiques. Le recours à des conventions typographiques est systématique en sciences du langage, où il permet de rendre compte de statuts linguistiques différents des éléments étudiés.

23 Voir *verbes de franchissement de frontière* dans D. I. Slobin, « The Many Ways to Search for a Frog: Linguistic Typology and the Expression of Motion Events », dans S. Strömquist et L. Verhoeven (éds) *Relation Events in Narrative: Typological and Contextual Perspectives*, Mahwah (NJ): Lawrence Erlbaum Associates, 2004, p. 219-257.

10) Un peu en contrebas, il y avait *la petite ville, ramassée sur elle-même, maison contre maison*, avec en retrait *la masse entassée des bâtiments* de l'Usine et ses cheminées de brique qui semblaient rentrer dans le ciel (...). Un paysage de fumée et de travail, *une sorte de coquille avec plein d'escargots dedans* qui ne se souciaient pas du reste du monde²⁴.

Cependant, des séquences descriptives qui contribuent à construire la représentation d'un territoire proximal resserré et confiné dans ses limites sont relativement rares. Si le lecteur le perçoit comme tel, c'est en raison des partitions successives effectuées à travers des choix lexicaux et syntaxiques, des lexèmes et des constructions qui tout au long du roman servent à évoquer les lieux familiers aux protagonistes. En effet, sur le plan lexical, dans le périmètre proximal qui inclut aussi bien P. que V. éloignée pourtant d'une vingtaine de kilomètres – « tout un monde » (voir : exemple 1 ci-dessus), les deux localités, ainsi que leurs environs familiers sont évoqués à l'aide des lexèmes qui forment un groupe sémantiquement cohérent que l'on pourra qualifier de noms de lieux fonctionnels. Ils sont employés dans des syntagmes nominaux définis de forme générale *Art_{def} SN*. Nous verrons dans la section qui suit comment ce procédé, impliquant l'organisation lexicale et syntaxique du texte, permet d'effectuer des divisions de l'espace familial pour y délimiter des emplacements du territoire proximal concentrant des activités qui préoccupent les habitants de « notre petite ville ».

NOMS DE LIEUX FONCTIONNELS ET PARTITION DE L'ESPACE PROXIMAL

Le territoire proximal incluant la ville de P. bien intégrée dans le *pays* et celle de V., est évoqué par des lexèmes que nous qualifions ici de *noms de lieux fonctionnels* (désormais NLF). Les plus fréquents d'entre eux sont réunis dans le tableau 2 ci-dessous.

Tableau 2 : Noms de lieux fonctionnels dans l'espace proximal.

| NLF de <i>chez nous</i> , P. | NLF de V. |
|-------------------------------------|---------------------------------|
| la capitainerie | le Rébillon |
| le Château | la cathédrale |
| les dépendances | le Tribunal |
| la petite maison | la prison |
| le parc | le marché |
| le port d'embarquement | le bordel |
| l'Usine | l'octroi |
| le quartier cossu | l'évêché (siège institutionnel) |
| les anciennes fermes | |
| l'église | |
| l'école | |
| le cimetière | |
| le bistrot de Fermilien | |
| la quincaillerie de Quentin Thierry | |
| la mercerie d'Augustine Marchoprat | |
| la ferme des Mureaux | |
| la maison, notre maison | |

Les NLF possèdent trois caractéristiques sémantiques qui conduisent à délimiter dans le territoire proximal restreint du roman des parcelles plus petites encore, contigües et serrées les unes contre les autres, dans lesquelles se concentrent les activités des habitants. C'est par leur usage systématique que l'image de « masse entassée », de « coquille avec plein d'escargots dedans » apparaît dans la représentation des lieux avant son évocation explicite dans le chapitre IX, pour être relayée ensuite tout au long du roman. Cet effet de sens est lié aux trois caractéristiques lexico-sémantiques suivantes :

- i. Les NLF dénotent des entités concrètes.
- ii. Même lorsqu'ils autorisent des recatégorisations par des termes génériques parfois très différents, ils sont tous hyponymes des noms généraux d'espace (NGE)²⁵ LIEU et ENDROIT.

24 Ph. Claudel, *op. cit.*, p. 82-83.

25 R. Huyghe, *Les noms généraux d'espace en français. Enquête sur la notion de lieu*, Bruxelles, De Boeck Duculot, 2009.

- iii. Les NLF sont des unités lexicales de type quasi-prédicats²⁶ : ils dénomment des entités qui, tout comme les prédicats verbaux, engagent des actants-participants.

NLF : des noms concrets

En tant que noms concrets²⁷, les NLF dénotent des entités dont on peut spécifier la matière (modificateur *en briques* dans ma phrase 11) et les trois dimensions (attributs des exemples 11 et 12). Leur matérialité implique à son tour qu'elles sont accessibles aux sens, ce qui se manifeste dans les collocations que ces noms peuvent former avec des prédicats interactionnels tels que VOIR, TOUCHER, CONSTRUIRE, RÉPARER, etc. (cf. l'exemple 13). Précisons que sans être des citations, les phrases (11) à (13), tout comme celles en (14) et (15) utilisent des NLF du roman pour les besoins de cette démonstration.

- 11) Le Château en briques est immense, il n'est pas à taille humaine.
12) La quincaillerie de Quentin-Thierry est très large et haute de plafond.
13) Après la guerre, on a reconstruit l'Usine en la modernisant.

Notons que le caractère concret s'unit dans le cas des NLF à leur nature comptable, ce qui implique non seulement une localisation dans l'espace, ce qui est une propriété bien connue de tout objet²⁸, mais également des contours spatiaux bien déterminés. Cette caractéristique sémantique se manifeste entre autre dans la possibilité de les recatégoriser entre autres par par le NGE LIEU.

Recatégorisation par les NGE LIEU et ENDROIT

Pour comprendre l'organisation spatiale du roman, il est intéressant d'observer les possibilités de recatégorisation par hyperonymes. La manipulation consistant dans une reprise par un terme superordonné permet de mettre en avant des propriétés sémantiques des hyponymes, en l'occurrence de nos NLF.

Cette opération peut paraître légèrement déconcertante à cause des différences notables : CHÂTEAU, ÉGLISE, DÉPENDANCE, CAPITAINERIE peuvent par exemple être repris par BÂTIMENT ; ÉCOLE, RESTAURANT, BORDEL – par ÉTABLISSEMENT. Or, ni l'un ni l'autre de ces termes superordonnés ne convient pour CIMETIÈRE, MARCHÉ ou QUARTIER.

Les seuls hyperonymes communs aux NLF du tableau 2 sont des noms généraux d'espace ENDROIT et LIEU. Observons en effet les phrases (14) et (15), extraites respectivement des chapitres III et V. Dans les séquences en a-c, puis d-f, ajoutés pour les besoins de la discussion, les NLF *le Château* et *la mercerie d'Augustine Marchoprat*, ainsi que quelques autres NLF sont anaphoriquement repris par ENDROIT.

- 14) Le Château est immense, il n'est pas à taille humaine²⁹.
a) Cet endroit me donne des frissons.
b) L'endroit a une longue histoire.
c) C'est un endroit très particulier.
15) Son petit trajet [de Lysia] qui se déroula loin de nos regards la mena jusqu'à la mercerie d'Augustine Marchoprat³⁰.
a) L'endroit est bien connu de tous les habitants.
b) La mercerie d'Augustine Marchoprat est un endroit très fréquenté.
c) La mercerie, l'école, la mairie, la maison du maire – tous ces endroits ont un charme désuet.

Selon Andrée Borillo, suivie par Richard Huyghe³¹, ENDROIT décrit des localisateurs partitifs, c'est-à-dire des sites qui découpent des portions dans un espace plus vaste. La possibilité de placer les NLF dans la catégorie d'endroits s'explique certes par leur caractère concret, comptable et spatialement défini. Cependant, cette recatégorisation met aussi en évidence le rôle qu'ils jouent dans l'organisation de l'univers de référence : ils servent à découper et à isoler à l'intérieur du territoire proximal, des portions d'espace concrètes, enfermées dans des limites bien définies. Or, identifier un territoire revient à l'isoler (voir le constat de L. Talmy dans

26 I. Mel'čuk et A. Polguère, « Prédicats et quasi-prédicats sémantiques dans une perspective lexicographique », *Revue de linguistique et de didactique des langues (Lidil)* 37, 2008, p. 99-114.

27 M. Galmiche et G. Kleiber, « Sur les noms abstraits », dans N. Flaux, M. Glatigny et D. Samain (éds) *Les noms abstraits*, Lille, Presses Universitaires du Septentrion, 1996.

28 A. Quinton, « Objects and Events », *Mind* 88 (350), 1979, p. 197-214.

29 Ph. Claudel, *op. cit.*, p. 30.

30 *Ibid.*, p. 53.

31 A. Borillo, « Partition, répartition, balayage de parcours » dans G. Kleiber, C. Schnedecker et A. Theissen (éds) *La relation partie-tout*, Louvain, Éditions Peeters, 2006, p. 17-30, R. Huyghe, *op. cit.*

l'Introduction) et à le singulariser par rapport au reste. Ainsi, l'évocation de l'espace proximal de P. et de V. à travers des NLF conduit à le réduire à des portions bien déterminées.

Il est également possible de classer les NLF dans la catégorie des lieux. Reprenons les extraits (14) et (15), suivis par des séquences d) à f), dans lesquels *le Château* et *la mercerie d'Augustine Marchoprat* sont anaphoriquement repris par l'hyperonyme LIEU.

- 14) Le Château est immense, il n'est pas à taille humaine³².
- d) Ce lieu a une longue histoire.
 - e) C'est un lieu historique.
 - f) Le lieu est très impressionnant.
- 15) Son petit trajet qui se déroula loin de nos regards la mena jusqu'à la mercerie d'Augustine Marchoprat³³.
- d) Ce lieu a toujours attiré beaucoup de monde.
 - e) Le lieu est bien connu dans le pays.
 - f) C'est un lieu de rencontres et d'échanges de potins.

Les noms des entités que l'on peut inclure dans la catégorie de lieux héritent de leur hyperonyme un certain nombre de propriétés. Premièrement, comme l'observe Laure Vieu³⁴, elles apparaissent comme prédéterminées, incluses topologiquement dans le tout ; dans *Les Âmes grises*, ce sont les villes de P., de V. et *le pays* qui forment des tous. Deuxièmement, nous pouvons désigner par le lexème LIEU uniquement des entités ancrées au sol³⁵. Non seulement l'espace qu'elles occupent est isolé de son environnement immédiat selon le contour de leur limites matérielles, mais elles se trouvent fixées au sol, donc inséparables de la ville : elles sont là de manière permanente. Troisièmement enfin, un lieu permet de localiser des procès, des activités, des actions³⁶ : cette propriété sémantique est mise en avant en (15f). Autrement dit, la reprise des NLF par cet hyperonyme est liée également à leur caractère fonctionnel de quasi-prédicats : les habitants de P. et de V. s'y rendent pour y accomplir des activités d'un certain type.

NLF : des quasi-prédicats

Les entités concrètes dénommées par le NLF sont toutes des artefacts conçus et réalisés dans leur matérialité pour remplir des fonctions particulières, c'est-à-dire pour être utilisés d'une certaine façon et servir à des finalités précises. Autrement dit, ils sont destinés à accueillir des activités et des actants qui les réalisent. Ce sont des composantes sémantiques que l'on retrouve systématiquement dans leur sens lexical. Ainsi, en expliquant le sens d'ÉCOLE, on décrira certes les aspects matériels du bâtiment, mais la paraphrase définitoire devra préciser également les individus que l'on y trouve (enseignants, élèves) et leurs activités. En effet, sur le plan lexical, le lexème ÉCOLE est relié à ENSEIGNANT, ENSEIGNANTE, ÉLÈVE_{N masc}, ÉLÈVE_{N fém}, APPRENDRE. De même, RESTAURANT reste connecté à RESTAURATEUR, RESTAURATRICE, SERVEUR, SERVEUSE, REPAS, CLIENT, CLIENTE. Autrement dit, c'est en explicitant les rapports entre X (SERVEUR, SERVEUSE), Y (REPAS) et Z (CLIENT, CLIENTE) que l'on parviendra à construire une paraphrase du sens lexicalisé dans RESTAURANT : établissement où X sert Y à Z.

C'est justement le statut lexical des quasi-prédicats qui confère aux noms réunis dans le tableau 2 leur caractère fonctionnel : l'emploi d'un NLF conduit à évoquer, explicitement ou non, l'activité et les actants que ce lieu concret est destiné à accueillir.

L'image évoquée explicitement dans la séquence descriptive du chapitre IX (l'extrait 10 ci-dessus), celle d'un espace cloisonné et ramassé sur lui-même, est donc systématiquement construite tout au long du roman par le recours aux NLF. Leurs propriétés sémantiques sont mises à contribution pour découper dans le territoire proximal des portions bien délimitées, fixes, où les chemins des habitants se croisent régulièrement et où se concentrent leurs activités.

32 Ph. Claudel, *op. cit.*, p. 30.

33 *Ibid.*, p. 53.

34 L. Vieu, *Sémantique des relations spatiales et inférences spatio-temporelles. Une contribution à l'étude des structures formelles de l'espace en langage naturel*, Thèse de Doctorat de l'Université Paul Sabatier, Toulouse, Université Paul Sabatier, 1991.

35 R. Huyghe, *op. cit.*

36 *Ibidem.*

Les caractéristiques des NLF analysées dans la section précédente relèvent de la sémantique lexicale : elles sont codées dans la langue, dans les unités lexicales qu'elle met à disposition des locuteurs. Dans l'univers romanesque, elles sont mises à contribution pour diviser l'espace proximal en des portions plus petites, renfermées dans des limites bien précises. Dans le lexique français, les NLF sont des dénominations des entités relevant de certaines catégories d'objets. Mais c'est seulement lorsqu'un NLF (ou toute autre dénomination) est employé en discours qu'il réfère à un objet particulier.

Au-delà de leur sens lexical, les structures syntaxiques dans lesquelles apparaissent régulièrement les NLF dans le texte du roman sont également très importantes pour la représentation de l'espace construit pour et par le lecteur, car sans elles, l'énonciateur ne pourrait renvoyer à l'objet de son discours. Dans cette section, nous nous arrêterons sur les syntagmes nominaux déterminés gouvernés par les NLF (desormais SN_{LF} tels que *la prison de V.*, *le bistrot de Fermilien*, *la cathédrale*, *la capitainerie*, etc.). En effet, les NLF sont régulièrement employés avec un article défini. Cette organisation syntaxique repose en réalité sur des présupposés d'un certain type, ce qui permet ensuite de développer des procédés discursifs particulièrement intéressants du point de vue de la représentation de l'espace dans *Les Âmes grises*. De fait, c'est par le recours fréquent à l'anaphore associative dont les SN_{LF} font partie, que l'organisation polyphonique du discours romanesque non seulement contribue à la représentation spatiale, mais y place le lecteur.

Deux types de syntagmes nominaux définis, une même présupposition d'unicité existentielle

Lorsqu'un énonciateur mentionne une entité pour la première fois dans un discours, il place son nom dans la portée d'un article indéfini : c'est ainsi que procède le narrateur pour *un petit sentier communal* et *un pont japonais* dans l'extrait (16). Un SN indéfini sert à catégoriser le référent de N, c'est-à-dire à indiquer au destinataire du message dans quelle catégorie d'objets il faut chercher celui dont il est question. C'est seulement lorsqu'il est employé une deuxième, troisième fois que N apparaît dans un SN défini : c'est ce qui se passe pour *le petit canal dont j'ai déjà parlé*. De même, on pourrait par la suite évoquer *le petit sentier communal*, *le pont japonais*.

Georges Kleiber³⁷ indique que l'article défini véhicule une présupposition existentielle d'unicité à propos du référent du nom qui fait objet de détermination. Autrement dit, lorsqu'un énonciateur emploie un article défini, c'est parce que dans son esprit, il n'y a qu'un seul référent possible et il présuppose que le destinataire du message dispose du même savoir. La condition d'unicité est remplie, lorsque pour l'auteur et pour son interlocuteur, il n'y a qu'un seul référent – objet nommé – dont il peut être question. Dans l'affirmation *Le restaurant est fermé*, l'article défini n'est possible que si toutes les personnes impliquées dans l'échange possèdent les informations nécessaires pour l'identifier parmi les restaurants plus ou moins concurrents.

Or, tout au long du roman, dès la première mention d'un objet, le narrateur y renvoie par des SN_{LF} définis. Les extraits (16) à (18) en offrent un échantillon.

- 16) *Le Château* est immense, il n'est pas à taille humaine. (...) *Le parc du Château*, un régiment entier pourrait s'y cantonner sans se gêner aux entournares. Il est bordé d'eau : au fond, il y a *un petit sentier communal* qui sert de raccourci entre *la place de la Mairie* et *le port d'embarquement*, puis *le petit canal* dont j'ai parlé et sur lequel le vieux avait fait jeter *un pont japonais* (...) ³⁸.
- 17) Destinât, on l'appelait différemment selon les endroits et selon les gens. À *la prison de V.*, la plupart des pensionnaires l'appelaient Bois-le-sang. Dans une cellule, j'ai même vu un dessin au couteau sur une grosse porte en chêne qui le représentait. C'était d'ailleurs assez ressemblant. Il faut dire que l'artiste avait eu tout le temps d'admirer le modèle durant ses quinze jours de grand procès³⁹. [...] Après chaque procès, Destinât allait déjeuner au Rébillon, face à *la cathédrale*⁴⁰.
- 18) Premier lundi de décembre. *Chez nous*. 1917. Un temps de Sibérie. (...) Un peu en retrait il y avait le fils Bréchet, un gaillard pansu, les cheveux raides comme des poils de balai (...). Il faisait des écritures à *la capitainerie*⁴¹.

Ces SN_{LF} définis qui mentionnent pour la première fois un lieu particulier de l'espace proximal se répartissent en deux types selon leurs formes syntaxiques. Le premier pourra être représenté par la formule SN_{LF} de SN_2 : un NLF en position de gouverneur syntaxique y est suivi d'un complément prépositionnel introduisant un autre

37 Voir : G. Kleiber, « Article défini, théorie de la localisation et présupposition existentielle », *Langue française* 57, 1983, p. 87-105, G. Kleiber, « Anaphore associative et relations partie-tout : condition d'aliénation et principe de congruence ontologique », *Langue française* 122, 1999, p. 70-100, G. Kleiber, *L'Anaphore associative*, Presses Universitaires de France, 2001.

38 Ph. Claudel, *op. cit.*, p. 30-31.

39 *Ibid.*, p. 12-13.

40 *Ibid.*, p. 15.

41 *Ibid.*, p. 18.

syntagme nominal (SN₂) : *la prison de V.* dans (16), *la mercerie d'Augustine Marchoprat* (voir l'extrait 15 ci-dessous), *la quincaillerie de Quentin Thierry*, *le bistrot de Fermilien*, *le parc du Château*, *la maison du maire*, etc. Le SN₂ est le plus souvent alors un nom propre ou, comme dans le dernier cas, un autre SN défini (*le maire*)⁴².

La condition d'unicité existentielle est alors satisfaite par la connaissance présupposée du lieu, de l'individu dont le nom est introduit par la préposition *de* ou, le cas échéant, par le caractère défini d'un syntagme tel que *le maire* : on le sait, il ne peut y en avoir qu'un seul dans une commune. Pour dire les choses simplement, l'énonciateur de ces SN définis présuppose que tout le monde connaît la ville de V., Quentin Thierry, Augustine Marchoprat, Fermilien, le Château et le maire aussi bien que lui-même et ses compatriotes.

Le second type des SN_{1F} définis qui sont une première évocation d'un lieu est autrement plus intrigant, puisqu'une aucune spécification ne vient s'y ajouter : c'est le cas de *la cathédrale* dans (17), de *la capitainerie* (18). Dans l'extrait (16), *le port d'embarquement* est un type de port, sans que le complément *de SN₂* permette une identification comparable à celle qu'assure un nom propre. En vertu de la condition d'unicité, l'énonciateur de ces portions de discours présuppose que son ou ses interlocuteurs disposent de connaissances nécessaires pour identifier le référent. Autrement dit, parmi tous les ports d'embarquement, toutes les cathédrales et capitaineries possibles, le destinataire-lecteur saura repérer celui ou celle dont il s'agit.

On peut admettre que le lecteur identifie ces référents, car le cotexte gauche fournit leur emplacement : dans (17) la ville de V. est mentionnée dans la partie qui précède l'évocation de la cathédrale, et *chez nous* permet de comprendre où se trouve la capitainerie. Ce mode d'organisation textuelle relève de l'anaphore : l'interprétation d'une expression B (pour nous, celle d'un NFL telle que *la cathédrale* ou *la capitainerie*) se fait par rapport à une autre expression A (son antécédent), qui se trouve dans la séquence qui précède (cotexte gauche). Francis Corblin⁴³ caractérise cette relation en termes de dépendance, puisque B (en l'occurrence, notre NFL) est nécessairement interprété en fonction de son antécédent A (V., *chez nous*). Plus précisément, il s'agit de ce que G. Kleiber⁴⁴ appelle *anaphore associative locative* : l'antécédent A permet de localiser l'objet le lieu évoqué par l'expression anaphorique B.

Néanmoins, l'antécédent est souvent repris par des anaphoriques situés bien loin dans le cotexte à droite. Ainsi, en (17), l'anaphore associative s'étend depuis la page 12 où le nom de V. est cité, jusqu'à la page 15 qui voit apparaître le SN défini *la cathédrale*. Sa dépendance par rapport au nom de la ville repose donc entièrement sur la mémoire du destinataire du texte. Sa connaissance des configurations des lieux dans l'univers romanesque est suffisamment forte pour assurer l'unicité référentielle indispensable pour le fonctionnement de l'enchaînement anaphorique, par exemple celui entre la cathédrale et la ville de V. (où se trouve par ailleurs également la prison locale).

Cette connaissance des lieux repose-elle exclusivement sur l'attention et la bonne mémoire du lecteur ? La réponse à cette question nous conduira, dans la section suivante, à essayer d'identifier les différents énonciateurs et ceux auxquels ils s'adressent.

UN JEU DE POLYPHONIE : LE LECTEUR DANS L'ESPACE PROXIMAL DES *ÂMES GRISES*

Résumons : les repères spatiaux dans *Les Âmes grises* sont déictiques. L'espace de l'univers romanesque est organisé par rapport au narrateur, témoin et acteur des événements, qui se déclare néanmoins appartenir à une communauté des habitants : « nous ». Pour ceux qui, comme lui, en font partie, est proche tout ce qu'ils fréquentent et connaissent. Ce qui se déroule tout près de chez eux, mais dont ils n'ont pas fait l'expérience est considéré comme lointain.

Leur espace proximal est représenté comme resserré et renfermé sur lui-même. En l'évoquant, le narrateur recourt systématiquement aux NFL, dont les propriétés sémantiques permettent de diviser le territoire familier et connu en des parcelles plus petites encore, dans lesquelles se concentrent les activités des habitants.

42 Les SN_{1F} de SN₂ *le parc du Château*, *les bâtiments de l'Usine* (l'extrait 10) mériteraient une étude plus approfondie dans la mesure où, si CHÂTEAU et USINE sont bien des noms communs, leur emploi dans le roman avec le recours constant à la majuscule semble indiquer qu'ils y sont traités comme des noms propres, du moins pour la communauté des habitants désignée par « nous ».

43 F. Corblin, « Les chaînes de référence : analyse linguistique et automatique », *Intellectica* 1 (1), p. 123-143.

44 G. Kleiber, *op. cit.*, 2001.

L'énonciateur du discours, message textuel que forme le roman, désigne les lieux fonctionnels (*la mercerie, la capitainerie, la cathédrale* dans les exemples ci-dessus et bien d'autres) par des SN_{LF} définis, parce qu'il présuppose que pour le destinataire, comme pour lui-même, il n'y a qu'un seul référent possible. Ce présupposé assure également l'identification du lieu désigné par un SN_{LF} défini suivi d'un complément prépositionnel de SN₂ nom propre (*la prison de V., la mercerie d'Augustine Marchoprat*, etc.). Celui à qui on s'adresse est supposé connaître la ville de V. et savoir qui est Augustine Marchoprat, ce qui le conduit à comprendre immédiatement de quelle prison et de quelle mercerie il est question. Autrement dit, les deux interlocuteurs disposent de mêmes connaissances et d'une même expérience, indispensables pour trouver le référent, en l'occurrence le lieu fonctionnel.

La question se pose alors de savoir qui sont l'énonciateur et le destinataire du message textuel. Un texte romanesque étant par nature polyphonique, plusieurs voix s'y font entendre. Tout en s'adressant à ses lecteurs, l'écrivain-scripteur laisse la parole au narrateur. La narration est organisée selon un point de vue actoriel : le narrateur est l'un des habitants de P. Ses activités professionnelles le conduisent régulièrement à V. et l'amènent à prendre part dans les événements qui s'y déroulent. Son discours s'adresse à sa femme défunte, qui avait appartenu, elle-aussi à la communauté locale désignée par « nous ». De ce fait, elle possédait les connaissances nécessaires pour identifier les personnes et les lieux : le présupposé existentiel d'unicité est ainsi pleinement justifié.

Comment s'explique-t-il cependant, lorsque l'on pense que ce même discours est celui que l'auteur-scripteur adresse à ses lecteurs, qu'une distance géographique et temporelle sépare des « nous » du roman et de leur univers ? Sans jamais avoir fréquenté ni P. ni V. (admettons l'espace d'un instant qu'il soit possible de rejoindre un univers fictionnel), les lecteurs doivent identifier les lieux fonctionnels en interprétant les SN_{LF} définis destinés à des membres de la communauté locale. Ainsi, en tant qu'énonciateur premier du discours romanesque, l'écrivain-scripteur présuppose chez ses lecteurs les mêmes connaissances et une même expérience que celles dont disposent les habitants de P., comme si l'espace décrit dans le roman, le *chez nous*, leur était tout aussi familier. Ils s'y trouvent impliqués, absorbés de la même façon que les personnages du roman. Cette intégration des lecteurs dans « la petite ville ramassée sur elle-même, maison contre maison », en leur présupposant les savoirs et expériences communes avec les habitants de P. ne serait-elle pas, de la part de l'auteur un moyen de les compter eux-aussi parmi ces « âmes joliment grises » qui peuplent le roman ? Celles-là mêmes qui regardent sans voir et entendent sans écouter les échos de ce qui se déroule en dehors de leur espace proximal.