# Polysémie régulière et néologie sémantique : constitution d'une ressource pour l'étude des sens nouveaux

Lucie Barque, Pauline Haas, Richard Huyghe

▶ **To cite this version:**

Lucie Barque, Pauline Haas, Richard Huyghe. Polysémie régulière et néologie sémantique : constitution d'une ressource pour l'étude des sens nouveaux. Neologica : revue internationale de la néologie, 2018, 12, pp.91-108. 10.15122/isbn.978-2-406-08196-8.p.0091 . hal-03823528v2

**HAL Id: hal-03823528**
**https://hal.science/hal-03823528v2**

Submitted on 21 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Publié dans *Neologica*, 12 (2018), pp 91-108**

# Polysémie régulière et néologie sémantique : constitution d'une ressource pour l'étude des sens nouveaux

Lucie Barque, Pauline Haas et Richard Huyghe

Université Paris 13 & UMR 7110 LLF (CNRS / Paris 7), Paris, France <lucie.barque@univ-paris13.fr>
Université Paris 13 & UMR 8094 Lattice (ENS / CNRS / Paris 3), Paris, France <pauline.haas@univ-paris13.fr>
Université de Fribourg, Fribourg, Suisse < richard.huyghe@unifr.ch>

## 1. **Introduction**

La néologie sémantique, définie comme l'ensemble des procédés permettant d'associer un sens nouveau à une unité existante (Bastuji 1974, Sablayrolles 2012), a fait l'objet de peu d'attention comparativement à la néologie formelle. Une des raisons de cette différence de traitement est peut-être que les néologismes sémantiques sont plus difficiles à collecter que les néologismes formels. Bien que se développent des techniques de repérage s'appuyant sur des indices contextuels (p. ex. Reutenauer *et al*. 2011, Falk *et al.* 2014), les systèmes de veille néologique ne détectent pour l'heure que très peu de néologismes purement sémantiques.
Le manque de données observables ne doit toutefois pas nous dissuader de nous intéresser à ce mécanisme fondamental du renouvellement du lexique qu'est la néologie sémantique. Une façon de le faire est de se concentrer sur les procédés à l'origine de la création de nouveaux sens. Si l'on s'en tient aux nouveaux sens produits par des extensions de sens récurrentes dans le lexique, la question du lien entre leur caractère plus ou moins néologique (entendu ici au sens de remarquable) et la nature de ces extensions de sens se pose. On peut par exemple faire l'hypothèse que les extensions de sens les plus régulières donnent lieu à des créations qui passent inaperçues auprès des locuteurs tandis que les extensions de sens peu régulières donnent lieu à des créations identifiées comme néologiques. L'emploi de DAHU en (1a), issu d'une métonymie quasi systématique en français (1a'), ne posera *a priori* aucun problème d'interprétation au contraire de l'emploi de GRANIT en (1b), issu d'une métonymie beaucoup moins régulière (1b').

(1) a. *Il a mangé du **dahu** hier*.
a'. Animal > Viande (p. ex. CANARD, POULET, LAPIN)
b. (?) *Il a acheté un **granit** au magasin*.
b'. Matière > Objet (p. ex. BRONZE, COTON, VERRE)

Les raisons qui font qu'une extension de sens est plus ou moins régulière dans une langue donnée ont bien été identifiées dans la littérature (p. ex. Lehrer 1990, Copestake et Briscoe 1995, Nunberg et Zaenen 1996, Kleiber 1999, Pethö 2007). Elles tiennent, entre autres, au type d'extension de sens en jeu (métonymie, métaphore, etc.), au caractère sémantiquement spécifié ou non du sens produit (p. ex. viande *vs* objet), au caractère nécessaire ou facultatif de la relation établie entre les référents en jeu (tout animal est constitué de chair mais toute matière n'est pas utilisée pour fabriquer des objets), ou encore à l'existence ou non dans le lexique d'unités concurrençant les sens issus d'une extension (p. ex. COL *vs* COU d'une chemise). L'étude des relations entre polysémie régulière et néologie sémantique requiert que

tous ces aspects soient pris en compte dans la description des patrons de polysémie afin d'évaluer leur degré de régularité respectif. Or, si l'on dispose bien dans la littérature d'inventaires de patrons de polysémie établis pour d'autres langues que le français (p. ex Apresjan 1974) et d'un certain nombre de descriptions de cas de polysémie régulière en français, tels que la polysémie des noms d'événement (p. ex Jacquey 2006, Barque *et al.* 2014) ou celle des noms d'affect (Goossens 2009), il n'existe pas, à notre connaissance, d'inventaire qui mette en regard les différents patrons du point de vue de leur régularité. Nous nous intéressons ici à la manière de constituer une telle ressource, qui pourrait fournir le matériel lexical nécessaire aux recherches psycholinguistiques portant sur la perception des sens nouveaux, et être utilisée pour détecter automatiquement en corpus les emplois non répertoriés de mots existants. La suite de l'article est structurée de la façon suivante : la section 2 est consacrée à la notion de polysémie régulière et aux difficultés relatives à sa description ; la section 3 décrit la méthode que nous avons adoptée pour constituer des patrons de polysémie nominale à partir de deux ressources lexicales du français librement disponibles, le *Wolf* (Fišer & Sagot 2015) et le *RL-fr* (Lux-Pogodalla & Polguère 2011). Enfin, la section 4 revient sur les relations entre polysémie régulière et néologie sémantique.

## 2. **La polysémie régulière**

On cite en général Bréal (1897) comme étant le premier à avoir posé les bases d'une étude proprement linguistique de la polysémie. On estime qu'il y a polysémie lorsqu'une forme lexicale est associée à différents sens, et que l'on peut établir un lien entre ces sens pris deux à deux — l'homonymie correspondant aux cas où une forme a différents sens non liés (p. ex. Martin 1972, ). À l'échelle du lexique, certains liens de polysémie entre paires de sens lexicaux paraissent récurrents et peuvent faire l'objet de généralisations. Selon Apresjan (1974), une polysémie mettant en jeu un rapport entre un sens A et un sens B est dite « régulière » dans une langue donnée si l'on observe au moins deux mots qui présentent cette polysémie. Ces deux mots ne doivent pas être synonymes, ce qui présuppose une généralisation du lien. En français, l'alternance consistant à dériver par métaphore un sens 'élément du corps d'une personne' à partir d'un sens 'élément du corps d'un animal' est ainsi régulière puisque, à l'instar du nom CROUPE (2), des noms comme BABINES, BAJOUE, CRINIERE, GUEULE, PATTE, etc. présentent cette polysémie.

(2) a. *Quand deux hommes se croisaient chacun tirant son cheval, l'un et l'autre donnait une petite tape de la main sur la croupe de la monture de l'homme de rencontre*.
b. *Un homme se retourne sur la croupe d'une jeune fille puis continue son chemin, un sourire aux lèvres*.

La régularité des patrons de polysémie est variable, suivant la proportion des mots qui, parmi ceux dotés d'un sens A, ont également le sens B. Selon Apresjan (1992), une polysémie régulière A > B est « productive » si tous les mots qui ont le sens A peuvent également être employés avec le sens B. Entendue dans ce sens, la notion de productivité s'observe dans des périmètres sémantiques très restreints, et demande de prendre en considération une granularité sémantique potentiellement très fine. Apresjan donne ainsi l'exemple des noms dénotant des types de voix (TENOR, ALTO, BARYTON) qui peuvent toujours être employés pour désigner la personne qui a ce type de voix. Pour Apresjan, la notion de productivité, synonyme de systématicité, est une notion discrète : une polysémie régulière est productive ou ne l'est pas. Il est possible toutefois d'envisager la productivité sémantique de manière graduelle, selon le degré de régularité d'application des schémas polysémiques. Evaluer cette régularité pose des problèmes théoriques et pratiques, qui peuvent rappeler ceux rencontrés dans l'évaluation de

la productivité des règles de construction morphologique (p. ex. Baayen 2009, Dal 2003, Plag 2006). Ils concernent aussi bien la question de la description des règles postulées que la question des modalités et précautions à adopter pour calculer leur productivité à partir de corpus.

### 2.1. *Evaluer la polysémie régulière : problèmes théoriques*

Évaluer la régularité d'un patron de polysémie nécessite que l'on distingue d'une part l'ensemble des mots qui présentent cette polysémie, d'autre part, l'ensemble des mots qui pourraient théoriquement la présenter. La circonscription de l'un et de l'autre de ces ensembles pose des problèmes d'ordres différents.
Commençons par la délimitation des instances attestées d'un patron donné. La question de savoir si un mot relève d'une polysémie régulière identifiée peut s'avérer moins triviale qu'il n'y paraît. Si l'on part du principe qu'un patron de polysémie représente une extension de sens, qui est donc par définition orientée, il peut être difficile de déterminer, pour un mot donné présentant la polysémie décrite, lequel des deux sens est dérivé de l'autre. Si l'on considère par exemple uniquement les sens 'meuble' et 'lieu' des noms BUREAU, BIBLIOTHEQUE, BAR et CUISINE, en faisant abstraction de leurs propriétés morphologiques et de leur structure polysémique globale, il semble *a priori* difficile de postuler une même orientation du lien pour ces différents noms. Le figement dénominatif corrélé à la lexicalisation peut également rendre délicate l'identification des instances d'un patron de polysémie. Le nom VERRE dénote aujourd'hui un contenant qui peut ou non être en verre, tandis que le nom BRONZE dénote une œuvre d'art qui est nécessairement en bronze. Aussi la question se pose-t-elle de savoir si ces noms doivent tous deux être considérés comme des instances attestées d'un même patron de métonymie produisant, à partir d'un sens de matière, un sens d'objet fait à partir de cette matière (Nunberg & Zaenen 1997 : 13). On voit que l'exercice consistant à identifier les instances d'une polysémie régulière donnée peut conduire à des analyses lexicales qui ne correspondent pas aux constructions sémantiques effectives.
Venons-en maintenant au problème de l'identification des mots qui ne présentent pas la polysémie décrite mais qui pourraient théoriquement la présenter dans la mesure où l'un de leur sens répond aux contraintes du sens de base de la dérivation sémantique postulée. Par exemple, des noms comme CORNE, GARROT ou SABOT, parce qu'ils dénotent un élément du corps d'un animal, pourraient théoriquement dénoter un élément du corps d'une personne, selon le même patron que celui illustré en (2) avec CROUPE. Le problème qui se pose ici est qu'il est très difficile de déterminer si un mot est ou non effectivement employé dans le sens second visé. Il faut se résoudre à fixer une limite (par définition arbitraire) entre sens attesté et sens non attesté. On peut pour ce faire recourir à la lexicographie et distinguer entre les sens qui sont répertoriés dans un ou plusieurs dictionnaires et ceux qui ne le sont pas. Les inconvénients et les limites de cette méthode sont connus, les dictionnaires n'offrant qu'une photographie partielle et partiale du lexique d'une langue (Sablayrolles 2008). Une méthode plus satisfaisante serait d'avoir, en plus des descriptions disponibles dans les dictionnaires, un accès aux sens en corpus. Toutefois, comme il a déjà été mentionné plus haut, les méthodes de détection automatique des sens lexicaux fondée sur leurs propriétés distributionnelles ne sont pas, à ce stade, assez performantes pour produire des résultats sur lesquels s'appuyer.

### 2.2. *Evaluer la polysémie régulière : problèmes pratiques*

La polysémie régulière est généralement considérée comme un phénomène central dans la constitution du lexique (p. ex. Nunberg & Zaenen 1997 : 13). Cependant, peu d'études se sont attachées à quantifier réellement ce phénomène en mesurant la part des mots polysémiques

dont les sens, considérés deux à deux, relèvent d'un patron de polysémie. Cette lacune est due en partie au fait que la quantification des cas de polysémie nécessite le recours à des ressources lexicales couvrantes et explicitement structurées, ce qui existe pour peu de langues, à l'exception de l'anglais. La ressource *WordNet*, réseau lexical de l'anglais (Miller *et al.* 1990), a de fait été exploitée dans l'optique d'en extraire des patrons de polysémie. Les méthodes utilisées divergent mais toutes reposent sur le regroupement de noms polysémiques qui partagent, pour deux de leur sens, une même paire d'hyperonymes (p. ex. Buitelaar 1998, Peter & Killgariff 2000, Peters 2006, Barque & Chaumartin 2008). Par exemple, les noms anglais ASSEMBLAGE ('assemblage'), ASSORTMENT ('assortiment'), COLLECTION ('collection'), ont chacun deux sens qui ont comme hyperonymes *act* et *group*, ce qui permet de postuler un patron de polysémie mettant en jeu ces deux sens. Le résultat de ce type d'extraction automatique laisse bien sûr apparaître de nombreuses erreurs, s'agissant aussi bien des patrons postulés que des mots supposés relever de ces patrons. Par exemple, le nom anglais ADMINISTRATION a un sens d'action et un sens de groupe mais le lien entre les deux n'est pas du même type que celui qui lie les sens d'ASSEMBLAGE, ASSORTMENT : le groupe résulte de l'action dans le cas d'ASSEMBLAGE et ASSORTMENT alors que dans le cas d'ADMINISTRATION, il est celui qui l'accomplit. Des heuristiques, reposant sur l'utilisation des gloses associées aux « synsets » (groupements de synonymes au sein du réseau lexical) ou sur la prise en compte de la profondeur de la hiérarchie, ont été proposées pour raffiner l'extraction. Toutefois, les données extraites automatiquement à partir des différentes méthodes exposées dans ces études n'ont pas, à notre connaissance, fait l'objet d'une validation manuelle extensive.

## 3. **Inventaire des patrons de polysémie du français**

Le français ne dispose pas à l'heure actuelle de ressources lexicales librement disponibles qui soient comparables à *WordNet* en terme de couverture et de précision des données. Il existe néanmoins deux ressources qui peuvent s'avérer complémentaires de ce point de vue : le *Wolf* (Fišer & Sagot 2015) et le *Réseau lexical du français* ou *RL-fr* (Lux-Pogodalla & Polguère 2011). Le *Wolf* est une ressource dérivée de *WordNet* qui reprend la structure hiérarchique de la ressource anglaise en complétant les synsets avec des mots du français. Cette ressource, développée automatiquement et seulement partiellement validée, est donc très couvrante mais également très bruitée. Le *RL-fr* à l'inverse est une ressource développée manuellement qui fournit des informations précises, notamment sur les liens de polysémie, mais pour comparativement moins de données.

| | *Wolf* | *RL-fr* |
|---|---|---|
| Noms | 25 420[1] | 8 055 |
| Noms monosémiques | 19 676 | 5 234 |
| Noms polysémiques | 5 744 | 2 821 |
| Sens liés par polysémie | 16 783 | 8 213 |
| Taux de polysémie (par N polysémique) | 2,9 | 2,9 |
| Liens de polysémie encodés | - | 4 324 |

**Tableau 1 : informations sur la polysémie nominale dans le *Wolf* et le *RL-fr***

Avant d'expliquer comment nous avons exploité le *Wolf* et le *RL-fr* pour en extraire des patrons de polysémie, disons deux mots du type de classification sémantique proposé dans ces deux ressources. Dans le *Wolf*, comme dans le *WordNet* dont il est issu, les sens lexicaux sont

[1] Les unités multi-mots du Wolf (qui incluent de nombreuses entités nommées) ne sont pas prises en compte ici.

regroupés dans des synsets qui forment les nœuds d'une hiérarchie. Un mot pris dans un sens donné appartient donc à un synset et, pour le regrouper avec un nombre plus important de mots dans une classe sémantique plus générale, il faut remonter la branche de ses synsets hyperonymes. A côté de la branche des hyperonymes, *WordNet* propose un étiquetage sémantique des synsets à gros grain reposant sur un ensemble de 26 étiquettes métalinguistiques appelées *Unique Beginners* [désormais UB] et dont la liste complète est donnée ci-dessous (Miller *et al.* 1990, Fellbaum 1998).

(3) Act, Animal, Artifact, Attribute, Body, Cognition, Communication, Event, Feeling, Food, Group, Location, Motive, Object, Person, Phenomenon, Plant, Possession, Process, Quantity, Relation, Shape, State, Substance, Time, Tops

Il est important de rappeler que ces étiquettes permettaient à l'origine aux lexicographes de *WordNet* de travailler sur des sous-ensembles de la hiérarchie et qu'elles n'ont pas été créées dans un but de classification sémantique à proprement parler. Nous avons néanmoins choisi d'exploiter ce système d'étiquetage plutôt que la hiérarchie des synsets pour trois raisons principales. Tout d'abord, *WordNet* est connu pour son découpage de la polysémie extrêmement fin et l'utilisation des UB présente l'intérêt de limiter cette polysémie en regroupant sous un même sens les synsets qui sont associés au même UB. D'autre part, le *Wolf* étant bruité, il nous a semblé exclu d'exploiter des chaînes d'hyperonymes comme l'ont proposé les études susmentionnées sur l'extraction de liens de polysémie à partir du *WordNet* anglais. Enfin, le degré de généralité des sens visés par les UB nous paraît globalement suffisant pour un premier dégrossissement des patrons de polysémie.
Dans le *RL-fr*, les unités lexicales qui forment les nœuds du réseau sont associées à l'une des 970 étiquettes sémantiques organisées au sein d'une hiérarchie dont les principes de constitution sont décrits dans Polguère (2011). Il est ainsi possible d'associer aux unités lexicales des chaines d'étiquettes sémantiques, comme illustré ci-dessous pour trois unités lexicales du *RL-fr* (4).

(4) VEAU-I.2 : viande > chair animale > aliment > qqch que l'on consomme > entité > Qqch
DINER-I : repas > fait de se nourrir > activité [de qqn] > fait > Qqch
FIERTE-I.1 : sentiment > état psychique > état > fait > Qqch

Il apparaît donc que les deux ressources que nous mobilisons, le *Wolf* et le *RL-fr*, reposent sur des formes d'inventaire lexical différentes et manipulent des jeux d'étiquettes sémantiques bien distincts. Avant d'aborder la question de leur compatibilité, nous présentons le type de matériel extrait des deux ressources, permettant la constitution de patrons de polysémie pertinents.

### 3.1. ***Extraction et analyse des données du Wolf***

La première étape d'extraction des données du *Wolf* a consisté à générer une liste de paires candidates à partir des 26 UB (3), soit 325 paires $UB_1$-$UB_2$ (p. ex. Act-Attribute, Act-Plant, Body-Location). Nous avons ensuite extrait automatiquement, pour chacune de ces paires, l'ensemble des noms du *Wolf* dont deux des sens sont associés aux UB ciblés, comme illustré en (5) pour la paire Body-Location.

(5) Body-Location : *aire, antre, arbrisseau, arbuste, avant, bassin, bout, buisson, canal, ceinture, cercle, corne, coupe, côlon, côte, côté, derrière, droit, entrailles, espace, extrémité, face, flanc, front, gauche, grand, gros, hémisphère, liaison, main, milieu,*

*mont, mur, nombril, nœud, œil, ombilic, orbite, origine, page, partie, passage, pied, pierre, plaque, point, racine, radius, rayon, région, stratum, trou, tête, ventre, zone*

L'analyse des données extraites consiste d'une part à déterminer si les paires $UB_1$-$UB_2$ abritent ou non des patrons de polysémie, c'est-à-dire des cas de polysémie régulière au sein du *Wolf*, et d'autre part à identifier, parmi les noms automatiquement associés à ces paires, lesquels sont des instances effectives des patrons postulés. Si l'on observe par exemple les données de (5), on voit se dégager un patron de métaphore qui lie un sens de partie du corps à un sens de localisation spatiale. Parmi les noms extraits automatiquement pour la paire Body-Location, on identifie PIED, ŒIL, FACE, FLANC etc. comme des instances de ce patron. Les autres noms sont considérés comme de faux candidats. De manière générale, on peut distinguer, en dehors du bruit inhérent au *Wolf* et des cas de polysémie distincts entre l'anglais et le français, quatre grands types de faux candidats pour un patron de polysémie A > B donné.

(6) a. Homonymie : aucun lien entre A et B
b. Résultat de deux dérivations morphologiques : $X >_{Morph} A$ et $X >_{Morph} B$
c. Résultat de deux dérivations sémantiques : $X >_{Sem} A$ et $X >_{Sem} B$
d. Résultat d'une dérivation sémantique en chaîne : $A >_{Sem} X >_{Sem} B$

Dans le premier cas (6a), les deux sens A et B d'un nom donné relèvent de l'homonymie. Par exemple, le nom COUSIN peut désigner un insecte ou un membre de la famille mais ces deux sens ne sont pas liés entre eux. Dans les trois autres cas mentionnés en (6), il y a bien un lien de sens entre A et B mais ce lien n'est qu'indirect. Il peut tout d'abord résulter d'une dérivation morphologique « en parallèle » (6b). Par exemple, le nom ACCORDEUR peut dénoter un artefact ou une personne mais ces deux sens ne sont pas directement liés l'un à l'autre puisqu'ils sont tous les deux dérivés morphologiquement du verbe ACCORDER. On ne peut donc pas postuler à partir de ce type de noms un éventuel patron de polysémie 'artefact' / 'personne'. De la même façon, les mots dont les deux sens considérés sont issus sémantiquement d'une tierce acception sont considérés comme de faux candidats (6c). Par exemple, le très polysémique nom CŒUR peut dénoter une localisation (ex. *le cœur de la ville*) et une propriété (ex. *un homme de cœur*) mais il ne sera pas considéré comme une instance d'un patron Attribute-Location dans la mesure où ces deux sens viennent l'un et l'autre de l'acception première du mot, à savoir celle qui dénote l'organe (ex. *un arrêt du cœur*). Il se peut enfin que, des deux sens considérés, l'un soit bien dérivés sémantiquement de l'autre mais que cette dérivation implique le passage par un sens intermédiaire (6d). Le nom TAUREAU figure par exemple dans la liste des noms candidats pour la paire Animal-Person. Cependant, l'acception de TAUREAU étiqueté comme 'Person' ne désigne pas un individu qui ressemble à un taureau mais un individu qui est du signe astrologique du taureau. La dérivation sémantique dans ce cas est donc : *animal > constellation dont la forme évoque l'animal > personne née sous le signe astral associé à la constellation*.

La description d'un patron de polysémie identifié à partir d'une paire $UB_1$-$UB_2$ requiert l'indication d'un certain nombre de propriétés du lien postulé, parmi lesquels son orientation (*i.e.* quel est celui des deux sens qui est à l'origine de l'extension), son type (*i.e.* métonymie, métaphore, etc.) et les conditions sémantiques liées au sens de base, au sens dérivé et à la relation qu'ils entretiennent (voir p. ex. Martin 1972). Une paire $UB_1$-$UB_2$ peut ainsi abriter des patrons qui s'opposent par leur orientation. Par exemple, les données extraites pour la paire Animal-Person regroupent des instances de deux patrons de métaphore se distinguant par leur orientation (7).

(7) a. Métaphore : Animal > Person (p. ex. AGNEAU, BLAIREAU, MOUTON)
b. Métaphore : Person > Animal (p. ex. CARDINAL, EMPEREUR, MANDARIN)

Une paire $UB_1$-$UB_2$ peut également abriter des patrons de polysémie qui, bien que reliant des paires de sens de même type, s'opposent sur le type d'extension de sens en jeu. C'est le cas par exemple du lien entre 'artefact' et 'personne', qui peut être métonymique lorsque la personne est symbolisée par l'artefact (8a) ou métaphorique lorsque la personne a une propriété similaire à celle de l'artefact (8b).

(8) a. Métonymie : Artifact > Person (p. ex. FLUTE, FUSIL)
b. Métaphore : Artifact > Person (p. ex. AS, MACHINE, PILIER)

Enfin, la relation entre le sens A et le sens B peut devoir être précisée. Par exemple, la paire Act-Artifact regroupe un ensemble de patrons qui sont tous de type métonymique, qui tous consistent à dériver un sens d'artefact à partir d'un sens d'action, mais qui se distinguent par le rôle joué par l'artefact dans la situation dénotée par Act (9a-c).

(9) a. Métonymie-Résultat : Act > Artefact (p. ex. AGRANDISSEMENT, IMPRESSION)
b. Métonymie-Instrument : Act > Artefact (p. ex. AÉRATION, CANDIDATURE)
c. Métonymie-Lieu : Act > Artefact (p. ex. PASSAGE, PROMENADE)

L'analyse manuelle des 5 744 noms polysémiques du *Wolf* associés aux 325 paires $UB_1$-$UB_2$ a donné lieu à la description de 44 patrons de métonymie et de 16 patrons de métaphore dont certains sont illustrés en (10) et (11).

(10) a. Métonymie : Attribute > Act (ex. AMABILITÉ, BRUTALITÉ, NÉGLIGENCE, RUSE)
b. Métonymie : Act > Possession (ex. AUGMENTATION, DÉPOT, INDEMNISATION)
c. Métonymie : Artifact > Location (ex. BAR, BUREAU, BIBLIOTHÈQUE)

(11) a. Métaphore : Body > Artifact (ex. BASSIN, BOUCHE, ANTENNE, AILE)
b. Métaphore : Animal > Person (ex. AGNEAU, BALEINE, BÂTARD, MOUTON)
c. Métaphore : Person > Animal (ex. EMPEREUR, CARDINAL, MANDARIN)

### 3.2 ***Extraction et analyse des données du RL-fr***

Le *RL-fr* se distingue du *Wolf* en ce qu'il encode explicitement les relations de polysémie au moyen d'un certain nombre de types listés en (12)[2].

(12) Codérivation, extension, intersection, métaphore, métonymie, spécialisation

Seuls les liens de métonymie et de métaphore, qui regroupent à eux seuls les deux tiers des liens de polysémie nominaux actuellement encodés dans le *RL-fr* (2905/4298), ont été pris en compte ici (voir tableau 2 ci-dessous). Comme indiqué plus haut au sujet des données du *Wolf*, les liens de co-dérivation morphologique ne nous intéressent pas puisqu'ils résultent de

[2] Seuls les types de polysémie qui concernent les noms dans le *RL-fr* sont consignés dans cette liste. Par ailleurs, le modèle du *RL-fr* prévoit d'encoder des sous-types pour certains types de polysémie, tels que « fonction », « forme », « comportement » pour la métaphore, et « ensemble de », « partie de » pour la métonymie. Ces sous-types n'étant à l'heure actuelle pas encore systématiquement encodés, nous les avons ignorés dans notre étude.

dérivations distinctes (p. ex. ASPIRATION ‘fait d’aspirer à qqch’ et ASPIRATION ‘fait d’aspirer qqch’). Les liens d’extension, de spécialisation et d’intersection, qui relèvent quant à eux de liens directs, nous ont paru sources d’un plus grand nombre d’erreurs d’encodage et laissent apparaître moins de régularité. Du reste, le grain choisi pour le découpage de la polysémie étant très fin dans le *RL-fr*, il nous a semblé opportun de nous concentrer sur les liens de métonymie et métaphore, qui impliquent souvent des changements de classes sémantiques et qui laissent, de ce fait, apparaître les cas de polysémie les plus clairs.

| Type | Liens | Exemples de liens |
|---|---|---|
| Métaphore | 1823 | MATRAQUAGE-I (action) > MATRAQUAGE-II (communication langagière)<br>VEAU-I.1 (animal) > VEAU -II (individu qui a un certain comportement)<br>NOISETTE-I (fruit) > NOISETTE-II (petite quantité) |
| Métonymie | 1082 | CYCLISTE-I (individu qui a une certaine activité) > CYCLISTE-II (vêtement)<br>ELOIGNEMENT-I (action) > ELOIGNEMENT-II (situation)<br>PORCELAINE-I (matière) > PORCELAINE-II (objet ayant une fonction esthétique) |

**Tableau 2 : liens de métaphore et de métonymie dans le *RL-fr***

La méthode adoptée pour extraire automatiquement des patrons de polysémie du *RL-fr* a suivi trois étapes. Les étiquettes sémantiques associées aux unités lexicales liées par polysémie étant issues d’une hiérarchie, nous avons généré, à partir d’un lien représenté par sa paire d’étiquettes, des patrons plus généraux en remontant les branches de chacune des étiquettes. La régression s’arrête lorsque l’on arrive à l’une des filles des étiquettes *fait* ou *entité* situées tout en haut de la hiérarchie. Considérons par exemple le lien de métaphore entre NOISETTE-I et NOISETTE-II donné en exemple dans le tableau 2. On postule un patron de polysémie à partir des deux étiquettes sémantiques des deux unités lexicales, puis on postule un nouveau patron plus général à chaque fois que l’on peut remonter dans la branche de l’une ou l’autre des étiquettes, comme illustré en (13). A ce stade, le lien entre NOISETTE-I et NOISETTE-II est considéré comme un lien relevant de chacun des trois patrons.

(13) fruit>petite quantité → aliment>quantité → qqch. que l’on consomme>quantité

La seconde étape consiste d’une part à sélectionner les patrons qui sont effectivement classifiants, c’est-à-dire qui sont associés à au moins deux liens distincts, d’autre part à ne conserver parmi les patrons généralisés que ceux qui sont plus classifiants que les patrons à partir desquels ils ont été généralisés, contrainte vérifiée lorsque le patron généré regroupe plus de noms que le patron de départ. Si l’on reprend notre exemple, le patron *fruit>petite quantité* postulé à partir du lien entre deux des sens de NOISETTE est conservé car on trouve dans le RL-fr un autre lien relevant du même patron (celui qui lie deux sens de NOIX). Le patron *aliment>quantité* issu de la généralisation est également conservé car il regroupe d’autres noms que NOISETTE et NOIX, en l’occurrence PAIN et TARTINE. La dernière étape consiste à valider manuellement les patrons conservés à l’étape précédente. Le processus de généralisation des patrons conduit en effet dans certains cas à mettre en relation des étiquettes sémantiques trop générales qui perdent leur pouvoir descriptif. Nous avons retenu, à l’issue de cette phase de validation manuelle, 50 patrons de métonymie et 53 patrons de métaphores, dont certains sont en relation hiérarchique.

### 3.3. ***Discussion***

Les deux ressources mobilisées dans notre étude l’ont été pour leur complémentarité. Le *Wolf*, bien que bruité, contient un grand nombre de noms polysémiques tandis que le *RL-fr*, parce qu’il encode explicitement les liens de polysémie entre unités lexicales, permet d’accéder

avec un bon taux de précision à des patrons de polysémie. Sans surprise, bon nombre de patrons extraits sont communs aux deux ressources (22 patrons de métonymie et 9 patrons de métaphore). Quatre d'entre eux sont illustrés ci-dessous dans le tableau 3. Certaines instances de ces patrons sont présentes dans les deux ressources (p. ex. AILE, ANTENNE, TALON pour le premier patron de métaphore), mais de nombreuses unités lexicales exemplifiant un patron donné ne figurent que dans l'une des deux ressources.

| Patrons | Instances du RL-fr | Instances du Wolf |
|---|---|---|
| Métaphore : élément du corps <> objet / partie d'un objet | AILE, ANTENNE, BARBE, BEC, BOUCHE, BRAS, CADAVRE, COU, DENT, DOS, FACE, FESSE, GENOU, GORGE, GUEULE, HANCHE, INDEX, JAMBE, MAIN, OREILLE, PATTE, PIED, POIL, TALON, TETE | APPENDICE, AILE, ANTENNE, ARTICULATION, BASSIN, BOUCHE, BRAS, CELLULE, DISQUE, TALON |
| Métaphore : animal > objet | SOURIS, PIGEON | BOA, BELIER, CHENILLE, PUNAISE |
| Métonymie : matière > objet | PORCELAINE, CAOUTCHOUC, PAPIER | ALPAGA, AQUARELLE, ARGENT, ARDOISE, BRONZE, CACHEMIRE, CAOUTCHOUC |
| Métonymie : plante > substance | CACAO, MOUTARDE, VANILLE | AMBROISIE, CANNABIS, CAOUTCHOUC, CHANVRE, RAPHIA |

**Tableau 3 : patrons communs aux deux ressources**

Les patrons présents seulement dans une des deux ressources le sont en raison de la nature des ressources (bruitée *vs* peu bruitée, couvrante *vs* peu couvrante) et du choix de classification sémantique adopté pour l'extraction des patrons. En effet, la classification à gros grain adoptée pour la description de la polysémie dans le *Wolf* implique d'une part que les deux sens en jeu dans un patron appartiennent à des classes différentes (p. ex. les noms polysémiques qui auraient deux sens de type Person n'ont pas été pris en compte), d'autre part que les patrons ne sont pas hiérarchisées (les UB définissant des classes qui s'excluent mutuellement). Les patrons présentant des sens issus d'une même classe générale, cas de figure théoriquement peu plausible pour la métonymie mais attendu pour la métaphore (14), ont donc été extraits uniquement du *RL-fr*.

(14) a. Métaphore : individu > individu (ex. FRÈRE, PARRAIN, PATRONNE)
b. Métaphore : caractéristique (physique) > caractéristique (ex. DISPOSITION, FLOU, FORCE, GOÛT, RONDEUR, SAVEUR)

De même, la méthode d'extraction adoptée pour le *RL-fr* a permis l'extraction de patrons de polysémie organisés hiérarchiquement. Les données en (15) présentent un ensemble hiérarchisé de patrons dans lesquels le sens dérivé varie en termes de précision sémantique.

(15) Métaphore : animal > individu
Métaphore : animal > individu de forte taille (ex. VACHE, BALEINE, ÉLÉPHANT, PACHYDERME)
Métaphore : animal > individu qui a un certain comportement (ex. CAMÉLÉON, COCHON, MOUTON)

A l'inverse, le caractère couvrant des données du Wolf a permis d'extraire des patrons qui ne ressortent pas à l'heure actuelle des données du *RL-fr* (16).

(16) a. Métaphore : Plant > Body (ex. BULBE, IRIS, GLAND)
b. Métonymie : Act > Food (ex. ASSAISONNEMENT, LIBATION, RAFRAICHISSEMENT)

En résumé, les patrons de polysémie et leurs instances nominales extraits du *RL-fr* et du *Wolf* vont permettre l'enrichissement mutuel des deux ressources, en peuplant la première de nouvelles paires d'unités lexicales liées par polysémie et en validant des données de la seconde. L'inventaire des patrons de polysémie initié ici devra encore faire l'objet d'un travail de validation et d'uniformisation des données, notamment pour ce qui est de la délimitation des classes sémantiques pertinentes pour la description de la polysémie régulière.

## 4. **Perspectives : utilisation des patrons pour l'étude de la néologie sémantique**

Nous avons exposé une méthode de description de patrons de polysémie à partir de données extraites de ressources lexicographiques. Ces patrons pourraient être utilisés dans le cadre d'une étude de la néologie sémantique, de différentes façons. D'une part, ils pourraient, à partir de l'établissement de schémas de polysémie régulière A > B, fournir des listes de lexèmes potentiellement candidats à la néologie, constituées des mots dotés lexicalement du sens A et non du sens B. L'observation en corpus de ces instances potentielles et de leur comportement distributionnel pourrait contribuer au repérage précoce des néologismes sémantiques.
D'autre part, les patrons de polysémie régulière pourraient être mobilisés dans l'étude du sentiment néologique. On s'interroge en effet sur la corrélation entre le degré de régularité d'une polysémie donnée et le sentiment de nouveauté sémantique que suscite auprès des locuteurs l'apparition du nouvel emploi d'un mot conforme à cette polysémie. Il est connu que le sentiment néologique varie non seulement d'un locuteur à l'autre, mais aussi d'une forme à l'autre (cf. Gardin *et al.* 1974, Sablayrolles 2003). S'agissant de la néologie sémantique, on peut faire l'hypothèse que plus une extension de sens est régulière et avérée, moins un sens nouveau produit par cette extension sera perçu comme néologique. Cette hypothèse implique que dans les exemples (1), repris ici en (17), l'interprétation de DAHU pose moins de problèmes au locuteur que l'interprétation de GRANIT — le transfert sémantique de 'viande' à 'animal' étant beaucoup plus régulier que celui de 'matière' à 'objet'.

(17) a. *Il a mangé du **dahu** hier*.
a'. Métonymie : animal > viande (p. ex. CANARD, POULET, LAPIN)
b. *Il a acheté un **granit** au magasin*.
b'. Métonymie : matière > objet (p. ex. BRONZE, COTON, VERRE)

Pour évaluer de manière expérimentale l'hypothèse d'un lien direct entre sentiment néologique et productivité polysémique, il faut disposer de patrons de polysémie et d'une indication de leur degré de régularité, calculé à partir des instances observées et des instances potentielles (cf. section 2.1). En interrogeant les locuteurs sur le caractère plus ou moins naturel de phrases comprenant des néologismes sémantiques réguliers, comme cela a été testé dans les expériences multilingues de Srinivasan & Rabagliati (2015), il est possible de déterminer quels schémas d'extension de sens produisent des acceptions susceptibles d'apparaître comme néologiques. On pourrait de la même manière examiner le lien entre les procédés de polysémie et le sentiment néologique, afin de déterminer si certains de ces procédés (p. ex. la métaphore) sont plus propices que d'autres (p. ex. la métonymie) à la création de sens perçus comme néologiques.
L'ensemble de ce traitement expérimental requiert une bonne connaissance des schémas de polysémie régulière et des instances lexicales qui les actualisent. La construction d'une ressource recensant l'ensemble des cas avérés de polysémie régulière semble ainsi trouver

toute sa pertinence dans un projet de dépistage de la néologie sémantique, et d'évaluation du sentiment néologique relatif aux emplois non lexicalisés de vocables existants.

## 5. **Références**


APRESJAN Jury (1974), « Regular Polysemy », *Linguistics*, 12(142), p. 5-32.

APRESJAN Jury (1992), *Lexical Semantics*, Ann Arbor, Karoma Publisher.

BAAYEN Harald (2009), « Corpus linguistics in morphology: morphological productivity », In A. Lüdeling and M. Kyto (eds) *Corpus Linguistics. An international handbook*. Mouton De Gruyter, Berlin, 900-919.

BARQUE Lucie et CHAUMARTIN François-Régis (2008), « La polysémie régulière dans WordNet », *Actes de TALN 2008*, Avignon.

BASTUJI Jacqueline (1974), « Aspects de la néologie sémantique », *Langages*, 36, p. 6-19.

BREAL Michel (1897), *Essai de sémantique (science des significations)*, Paris, Hachette.

BUITELAAR Paul (1998) « Corelex : An ontology of systematic polysemous classes », dans *Formal Ontology in Information Systems*, N. Guarino, Ed. IOS Press, Amsterdam, The Netherlands, p. 221–235.

COPESTAKE Ann et BRISCOE Ted (1995), « Semi-productive Polysemy and Sense Extension », *Journal of Semantics*, 1, p. 15-67.

DAL Georgette (2003), « Productivité morphologique : définitions et notions connexes », *Langue Française*, 140, p. 3-23.

FALK Ingrid, BERNHARD Delphine et GERARD Christophe (2014), « De la quenelle culinaire à la quenelle politique : identification de changements sémantiques à l'aide des Topic Models » dans *21ème conférence sur le traitement automatique des langues naturelles*, p. 525-530.

FELLBAUM Christiane (ed.) (1998), *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

FIŠER Darja et SAGOT Benoît (2015), Constructing a poor man's wordnet in a resource-rich world. *Language Resources and Evaluation*, 49(3), p. 601-635.

GARDIN Bernard, LEFEVRE G., TARDY Michel et MORTUREUX Marie-Françoise (1974), « À propos du *sentiment néologique* ». *Langages* 36, p. 45-52.

GOOSSENS Vannina (2009), « La polysémie des noms d'affect », *Tranel, 50*, p. 145-158.

KLEIBER George (1999). Problèmes de sémantique. La polysémie en question. Presses Universitaires du Septentrion, Villeneuve d'Ascq.

LEHRER Adrienne (1990), « Polysemy, conventionality, and the structure of the lexicon », *Cognitive Linguistics*, 1/2, p. 207-246.

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*, Ljubljana, Slovenia, p. 54-61.

MARTIN Robert (1972), « Esquisse d'une analyse formelle de la polysémie », *Travaux de Linguistique et de Littérature*, 10, pp 125-136.

MILLER George, BECKWITH Richard, FELLBAUM Christiane, GROSS Derek et MILLER Katherine (1990), « WordNet: An online lexical database », *International Journal of Lexicography*, 3(4), 235-244.

NUNBERG Geoffrey et ZAENEN Annie (1997), « La polysémie systématique dans la description lexicale », *Langue française*, 113, p. 12-23.

OSTLER, Nicholas et ATKINS Sue (1991), « Predictable Meaning Shift : Some Linguistic Properties of Lexical Implication Rules » dans Pustejovsky, J. et Bergler, S., éditeurs,

*Lexical Semantics and Knowledge Representation : First SIGLEX Workshop Proceedings*. Berlin, Springer-Verlag.

PETERS, Wim (2006), « In Search for More Knowledge : Regular Polysemy and Knowledge Acquisition » dans *Proceedings of GWC2006*, Jeju Island, Korea, p. 245–251.

PETHÖ Gergely (2007), « On Irregular Polysemy », dans P. Marina, P. Gergely & R. Csilla (Eds.). *The cognitive basis of polysemy : New sources of evidence for theories of word meaning*, Frankfurt, Peter Lang, p. 123-156.

PLAG Ingo (2006), « Productivity », dans Aarts, B. and A. McMahon (eds.), *Handbook of English linguistics*, Oxford, Basil Blackwell, p. 537-556.

POLGUERE Alain (2011), « Classification sémantique des lexies fondée sur le paraphrasage », *Cahiers de lexicologie*, 98, p. 197–211.

REUTENAUER Coralie, JACQUEY Evelyne et OLLINGER Sandrine (2011), « Néologismes de sens : contribution à leur caractérisation dans un corpus autour du thème de la crise financière » Dans *II Congrès International de Néologie des Langues Romanes (Cineo2011)*, São Paulo, Brazil, décembre 2011. 194 p.

SABLAYROLLES Jean-François (2003), « Le sentiment néologique », dans J.-F. Sablayrolles (éd.), L'innovation lexicale, Paris, Champion, p. 279-295.

SABLAYROLLES Jean-François (2012), dir., « Néologie sémantique et analyse de corpus », *Cahiers de lexicologie*, 100.

SRINIVASAN Mahesh et RABAGLIATI Hugh (2015), « How concepts and conventions structure the lexicon: Crosslinguistic evidence from polysemy », Lingua, 157, p. 124-152.

Résumé

Cette étude s'intéresse au lien entre la régularité polysémique et le sentiment néologique suscité par l'emploi nouveau de mots existants. Partant de l'hypothèse que la perception d'un sens nouveau dépend en grande partie du patron qui l'a produit, nous décrivons le développement, à partir de deux ressources du français, d'un inventaire de patrons de polysémie nominale qui prend en compte les types d'extension en jeu, leur orientation et leur degré de spécification sémantique. Ce travail de description a pour visée le repérage de candidats à la néologie sémantique et l'étude de la perception des sens nouvellement produits.

Mots clefs : sémantique nominale, polysémie régulière, néologie sémantique

Abstract

This paper deals with polysemy patterns and their relation to neological perception. Assuming that the latter is largely dependent on the former, we discuss a method for listing nominal polysemy patterns from two lexicographical resources of French. This method takes into account the relation between polysemous meanings, the orientation of semantic extensions, together with the more or less fine-grain semantics required to describe polysemous meanings. This work is intended to contribute to the evaluation of how semantic neologisms are perceived and to their detection in corpus.

Keywords : lexical semantics, regular polysemy, semantic productivity