



HAL
open science

UNITÉ ET DIVERSITÉ DE LA LINGUISTIQUE

Francis Manzano

► **To cite this version:**

Francis Manzano. UNITÉ ET DIVERSITÉ DE LA LINGUISTIQUE. Publications du CEL, Nouvelle série 2011-1, 2011, 978-2-36442-013-7. hal-02558785

HAL Id: hal-02558785

<https://hal.science/hal-02558785>

Submitted on 29 Apr 2020

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.



UNITÉ ET DIVERSITÉ DE LA LINGUISTIQUE



Sous la direction de Francis MANZANO

Publications du Centre d'Etudes Linguistiques
[Linguistique, dialectologie]

Nouvelle série 2011-1



UNITÉ ET DIVERSITÉ DE LA LINGUISTIQUE

© *Francis Manzano*

**Travaux à l'initiative du Centre d'Etudes Linguistiques de l'Université
Lyon 3 Jean Moulin
(Université de Lyon, France)**

**« Les mercredis du CEL »
2008-2009**

[13 textes, langue française et langue anglaise]

ISBN : 978-2-36442-013-7

Les données et positions intellectuelles des chercheurs appartiennent à ces chercheurs comme au terrain dont proviennent ces données. Vous pouvez utiliser et citer librement ce document mais en ayant la courtoisie d'en citer les sources exactes. Chaque texte reste la propriété scientifique exclusive de son auteur.

francis.manzano@univ-lyon3.fr

UNITÉ ET DIVERSITÉ DE LA LINGUISTIQUE

UNITÉ ET DIVERSITÉ DE LA LINGUISTIQUE

(CEL, 2008-2011)

Table des matières

- Francis Manzano
Unité et diversité de la linguistique
[pages 07 à 12]
- Peter Stockwell
Accommodating identity in cognitive sociolinguistics
[pages 13 à 33]
- Mariette Meunier & Luu Dinh Phuc
Présentation d'une recherche jurilinguistique sur le Code des délits et des peines -1810- et sur le Code pénal -1994-
[pages 35 à 43]
- Denis Jamet
Le repérage des énoncés métaphoriques (anglais – français)
[pages 45 à 65]
- François Maniez
Quelques tendances récentes de la linguistique de corpus
[pages 67 à 85]
- Dan McIntyre & Brian Walker
Annotating a corpus of Early Modern English writing for categories of discourse presentation
[pages 87 à 107]
- Laure Gardelle
Que faut-il entendre par « genre grammatical » en anglais moderne ?
[pages 109 à 124]
- Veronika Koller
Analyzing metaphor and gender in discourse
[pages 125 à 158]

Martina Pitz

La genèse de la frontière linguistique germano-romane à la lumière des recherches archéologiques et toponymiques

[pages 159 à 188]

Manuel Jobert

Estuary English (EE) : rupture ou continuité ?

[pages 189 à 209]

Brigitte Horiot

Autour du francoprovençal : intérêt diachronique et repères dialectologiques

[pages 211 à 232]

James Costa

Du local au global : essai de clarification idéologique préalable. Discours concurrents et revitalisation linguistique en Provence

[pages 233 à 255]

Francis Manzano

Patrimoines langagiers méditerranéens : modèles de gestion et perspectives

[pages 257 à 298]

Page de couverture :

Interactions - Banquet en pays gallo (Bretagne)
(collection personnelle, Marie-France Manzano, 2005)

Francis Manzano
Université Jean Moulin Lyon 3 – CEL, EA 1663

Unité et diversité de la linguistique

Durant les années 2008 et 2009, le Centre d'Etudes Linguistiques de l'université Lyon 3 procédait à une mise à plat de ses directions scientifiques en favorisant des échanges nouveaux entre chercheurs de cette équipe d'accueil de l'université (EA 1663) et en invitant à plusieurs reprises des chercheurs issus du PRES lyonnais (alors en cours de constitution) ou d'universités étrangères, notamment britanniques. Cette liaison scientifique avec différents centres de recherche d'outre Manche est à mettre à l'actif du Groupe de Recherche en Linguistique Anglaise (GRLA), composante du CEL dirigée et dynamisée par Denis Jamet et Manuel Jobert¹.

Durant la même période, l'équipe engageait une réflexion sur ses activités originelles en linguistique et en dialectologie romanes. Il apparaissait, de plus en plus clairement, que le cœur des activités du CEL devrait se situer dans les domaines fondamentaux de la dialectologie et de la sociolinguistique, ce qui fut particulièrement pointé par le colloque international « Mémoires du terrain : enquêtes, matériaux, traitement des données »², organisé à la même époque. Dans cet esprit, les composantes anglo-germanique d'une part, romane de l'autre, devraient par la suite poursuivre leur convergence autour de problématiques qui ont toutes en commun la relation fondamentale entre langage et société : variation géo-diachronique des langues, sociolinguistique et contact des usages contemporains, pragmatique et anthropologie des productions langagières. Par cette démarche

¹ Le GRLA réunissait des enseignants-chercheurs, notamment Denis Jamet, Manuel Jobert, Laure Gardelle, des doctorants ou post-doctorants, comme Lucille Bordet et Rémi Dignonnet. De nouveaux jeunes chercheurs doctorants les rejoignent depuis l'année universitaire 2010-2011.

² Colloque tenu les 12 et 13 mars 2009 (à paraître). Sur ce point et d'autres, on pourra se reporter au site Web du CEL : <http://facdeslangues.univ-lyon3.fr/recherche/centres/cel/>, onglet Archives. Travaux à paraître.

d'extension programmatique raisonnée se définissaient de mieux en mieux les contours d'une action commune durant une période où, par ailleurs, différentes voies de rapprochement des centres de recherche apparaissaient et se concrétisaient au sein du paysage lyonnais.

Ces différents ajustements et rapprochements ont logiquement concerné les chercheurs en tant que personnes individuelles. Au début du quadriennal universitaire 2011-2015, le CEL peut être ainsi caractérisé par un noyau actif de chercheurs opérant dans les directions de la dialectologie, de la sociolinguistique et de l'onomastique (langues romanes et zones méditerranéennes, domaines anglo-saxon et langues germaniques). Ceci concerne autant la description des variétés en présence (dialectologie historique et descriptive) que les modalités de contact au sein des paysages linguistiques contemporains (recul ou progression des langues, mixages, implications identitaires).

Un second pôle fédérateur rassemble des recherches d'anthropologie linguistique et de linguistique cognitive. Ce pôle rejoint régulièrement les pistes connexes de la stylistique et de la pragmatique des langues. Le GRLA y consacre nombre de ses activités et certaines recherches théoriques diachroniques sur d'autres terrains émergent plus globalement au sein du CEL (mécanismes de la production et de la diversification diachronique des langues, étiologie et épistémologie du langage).

De ces deux directions principales, *lato sensu* celles de la sociolinguistique et de l'anthropologie linguistique, amenées à se rapprocher plus encore dans l'avenir, la présente publication porte un premier témoignage.

Au-delà des chercheurs, développant individuellement leurs propres dynamiques scientifiques, on doit également évoquer les mutations en cours au sein du PRES lyonnais, car ces mutations ne sont pas sans rapport et sans incidences avec les mouvements qui viennent d'être rapidement ébauchés à propos du Centre d'Etudes Linguistiques.

En 2011 a été approuvé le développement d'un laboratoire d'excellence sur le site de Lyon, autour du projet ASLAN (*Advanced Studies on Language, Etudes avancées sur la complexité du langage*). Ce projet, initialement proposé par nos collègues des laboratoires UMR 5191 ICAR [Interactions, Corpus, Apprentissages, Représentations] et UMR 5596 DDL [Dynamique du Langage], a

reçu dès sa conception certaines propositions thématiques complémentaires du CEL. Les questions précédemment évoquées pourraient trouver naturellement leur place dans un ensemble scientifiquement cohérent si l'avenir est favorable aux rapprochements universitaires : dialectologie et sociolinguistique des langues d'origine de la région Rhône-Alpes (francoprovençal, occitan ou langue d'oc), contacts entre ces langues/cultures d'origine et les langues et cultures issues de l'immigration dans un vaste processus de fusion socio-identitaire en situation principalement francophonique.

Dans un cadre de ce type, chaque laboratoire pourrait sans doute développer ses spécificités tout en contribuant à une modélisation scientifique d'ensemble portant sur les contacts, les processus d'affirmation et de mouvement des langues.

Le CEL poursuivra donc ses recherches collectives sur les langues régionales de France et/ou d'Europe et du bassin méditerranéen. Dans une période où la plupart de ces langues d'origine (souvent romanes), le plus souvent dites « régionales » ou « minoritaires », rencontrent des difficultés croissantes, il est particulièrement important d'apprécier la part des segments identitaires qu'elles nous transmettent alors même qu'elles semblent le plus généralement vouées à régression ou disparition pure et simple dans les faits de parole.

Dans le même temps, qu'advient-il des langues et cultures issues d'une immigration contemporaine plongeant elle-même ses racines dans la longue histoire coloniale de la France ? Comment ces langues pourraient-elles mécaniquement échapper à l'immense chaudron de la francisation et de l'uniformisation centrifuge que semblent impliquer par ailleurs les mouvements de la « mondialisation » ? Sur ce point aussi l'une des tâches de la sociolinguistique est de tenter de définir les étapes logiques de cette fusion dont beaucoup de monde parle mais sur laquelle en réalité nous sommes peu renseignés, en repérant les traces formelles (internes, phonologiques, morpho-syntaxiques, lexicales, discursives) que les langues et cultures qui « cèdent » laissent à la koinè qui subsiste, laquelle se reconditionne elle-même de manière incessante, ce que l'on perd de vue parfois.

C'est sur ce point de vue général et cette zone de recherche que l'EA 1663 fondera ses axes de travail, en fonction de ses directions principales. L'équipe le fera au titre de ses programmes propres ou au sein du Labex ASLAN, elle pourra le faire également au sein de la

structure fédérative de recherche « Migration et Citoyenneté : Mutations, Métissage, Multilinguisme » (MC3M), à l'initiative de l'IETT³ et de l'université Lyon 3.

*

La publication qui suit comporte une bonne partie des interventions faites soit par des chercheurs du CEL, lors du séminaire « Les mercredis du CEL » programmé en 2008-2009, soit par différents invités au cours de rencontres organisées par les chercheurs romanistes ou anglicistes de l'équipe.

Ont été mises à contribution diverses manifestations organisées par le Département d'études anglophones, en accord et avec le soutien du CEL et du GRLA. C'est notamment de la sorte que sont intervenus Peter Stockwell, David Mac Intyre et Veronika Koller autour de thématiques sociolinguistiques et stylistiques⁴.

Le titre de la publication reprend l'intitulé synthétique des « mercredis du CEL » en 2008 : *Unité et diversité de la linguistique*. L'identité de la linguistique, en effet, soulève toujours quelques problèmes de fond que nous évoquerons brièvement pour finir.

La diversité est réelle. Partie des reconstructions historiques et typologiques du XIX^e siècle, la linguistique a fait sienne la reconnaissance et la description programmatique des langues du monde. Dans cette entreprise, les dialectologues, les ethnolinguistes, particulièrement les romanistes, ont occupé longtemps des positions épistémologiques fortes, c'est là un héritage naturel du CEL, que nous assumons car nous ne le considérons pas comme périmé. Savoir d'où l'on vient et une clef essentielle pour savoir où l'on va.

Mais en même temps le curseur scientifique a bougé au cours du XX^e siècle. La linguistique est devenue un modèle pour le reste des sciences humaines, du fait notamment de ses orientations structurales, systémiques et internes. Dans cette évolution, la collectivité des

³ Institut d'Études Transtextuelles et Transculturelles (EA 4186), Université Lyon 3, Faculté des langues. La SFR *Migration et citoyenneté*, initiée par Grégory Lee, directeur de l'IETT (actuellement détaché à l'université de Hong-Kong), est dirigée par Florent Villard, Directeur du département des Études chinoises.

⁴ On peut accéder au détail des interventions sur le site du CEL déjà mentionné : http://facdeslangues.univ-lyon3.fr/recherche/centres/cel/archives/cel-archives-2008-2009-299927.kjsp?RH=INS-LANcel_arch.

Les textes de nos collègues britanniques sont présentés ici dans leur langue d'origine conformément à leur présentation orale.

linguistes a certainement gagné en méthodologie, en capacités de description des corpus, dans des champs vedettes comme la phonologie ou la syntaxe. Mais à mesure qu'étaient pointées les faiblesses des prédécesseurs (diachroniciens, philologues ou dialectologues), remontaient les limitations naturelles d'une entreprise structurale isolée des fondements sociaux du langage, pour aussi saine et nécessaire qu'elle paraisse à première vue.

Ce sont ces limites qui ont déterminé l'émergence de la sociolinguistique depuis une quarantaine d'années, et nous n'y reviendrons pas une fois de plus. La publication d'aujourd'hui révèle bien quelques unes de ces entrées qui réaffirment les conditionnements sociétaux des mouvements de langues ou de la mise en discours.

Pour prendre un peu de hauteur et réutiliser la belle formulation de William Labov, le champ de la linguistique ne doit plus être ramené à deux camps qui s'affronteraient en bataille rangée : celui des linguistes « sociaux » (sociolinguistes *lato sensu*) et celui des linguistes « a-sociaux » (généralistes, structuralistes, générativistes)⁵. Après une tension très vive durant les dernières décennies du XX^e siècle (qu'on pouvait alors percevoir comme un schisme au sein des sciences du langage), les faits, la demande générale, laissent apparaître que la linguistique dite trop rapidement « générale » et les différents courants de la sociolinguistique aplanissent les aspérités et entreprennent très naturellement leur rassemblement, au bénéfice de l'ensemble des sciences humaines.

Il est en effet intéressant de voir que les linguistes, longtemps (seulement) réputés « techniciens » habiles des humanités, descripteurs modèles, sont à nouveau sollicités par différentes sciences humaines ou biologiques, et plus globalement par la société. Ce qui est écrit plus haut à propos des évolutions au sein du paysage universitaire lyonnais le souligne.

Ce n'est peut-être pas (en tout cas espérons-le) un énième phénomène de mode, mais plutôt la démonstration d'un malentendu profond que les sociolinguistes, précédés et accompagnés en cela par les dialectologues, ont clairement dissipé par leurs travaux tenaces. Au

⁵ Les propos de William Labov ne sont pas exactement cités mais évoqués à partir de son ouvrage fondateur *Sociolinguistique*, paru en 1976 aux Editions de Minuit (traduction française de *Sociolinguistics patterns*, 1972. University of Pennsylvania Press).

cœur de ce malentendu se trouvait la dimension intrinsèquement sociale du langage. Certes de nombreux linguistes généralistes, Saussure en premier lieu, l'avaient évoquée ; mais sans en faire, paradoxalement, la colonne vertébrale de leur action.

Il n'existe pourtant pas de société humaine sans langage. Truisme apparent qu'il faudrait pourtant répéter inlassablement et que cette publication redit encore. Le langage n'est pas un simple élément définitoire de la société parmi d'autres, il est, radicalement, basiquement, le principe même de la société humaine et de son identité. Tout le reste fondrait comme neige au soleil sans ce fondement que redécouvrent à leur manière les préhistoriens, les biologistes, les acteurs politiques de tous ordres.

Pour finir, c'est donc bien sur fond d'unité de la linguistique qu'un temps de la recherche collective lyonnaise est ici évoqué, au delà de la diversité apparente qu'apportent les marques historiques, dialectologiques, pragmatiques et cognitives, sociologiques ou même politologiques des textes proposés.

Francis Manzano
Directeur du Centre d'Etudes Linguistiques (EA 1663),
Université de Lyon
Directeur de la publication « Unité et diversité de la linguistique »

Peter Stockwell
University of Nottingham (United Kingdom)

Accommodating identity in cognitive sociolinguistics

For most of its history, modern linguistic study has maintained a divergence between mentalism on the one hand and social study on the other. The former domain encompasses the fields of pure and theoretical linguistics, formal syntax and semantics, and pragmatics in a philosophy-of-language tradition; the latter domain encompasses the fields of applied linguistics and sociolinguistics, diachronic variation studies, and pragmatics in a social context. To adapt a terminological dualism from the former domain, theorists of language have focused attention either on competence or performance, but have rarely integrated the two as a matter of principle. More recently, a cognitive linguistics has emerged as part of a general intellectual shift towards cognitivism across the sciences, arts and humanities. Many of the early figures in cognitive linguistics evolved their thinking out of the generative, mentalist approach; many of the later figures in the field came to cognitive linguistics from a functional and applied linguistic tradition. It was a matter of certain predictability that the two practices would produce some first steps in an integrated *cognitive sociolinguistics*. This paper is a contribution to that emerging field.

Most of the focus of Kristiansen and Dirven's (2008) collection entitled *Cognitive Sociolinguistics* lies in macro-sociolinguistic matters: dialectal variation and change, code-switching as national and regional identity, metaphors in culturally diverse models, and a critical discourse analytical application of cognitive linguistics at the level of social policy. These are recent innovations, but in fact there has long been a psychological and cognitively-sensitive facet in sociolinguistics, especially in the areas where it overlaps with pragmatics, such as in the exploration of implicature, or politeness

norms and deviations, or the analysis of social power along gendered, racial, ethnic and other dimensions. Indeed, the notion of *identity* has been a key foundational concept in sociolinguistics since the inception of the field in its modern form over the last 50 years; almost every discussion of the term has involved a recognition of the cline between social identity and individual membership and enactment of that social feature (see Antaki and Widdicombe 1998; Bucholtz, Liang and Sutton 1999; Dyer 2007).

Identity in sociolinguistics

Sociolinguistic accounts of identity over the years have mirrored general developments in the field. In the period of early quantitative variationism, identity was matched with the social variables of geography, gender, age, race, ethnicity, class, politics, occupation, family type, hostility to out-groups, and so on (see Stockwell 2007 for a summary). Later, influenced by ethnographic approaches, identity comes to be defined in terms of social network structures, speech communities, communities of practice, and a nexus of analysis (see Le Page and Tabouret-Keller 1985; Johnstone 1996). Latterly, under a social constructivist influence, the pragmatics of politeness and power once again are redefining identity (see Eckert 2000; Eckert and Rickford 2001; and Llamas, Mullany and Stockwell 2007 for an overview).

Much depends on the working definition of ‘identity’, of course. Some sociolinguists take a narrow compass and others a broader one: Bauer (2008) and Schneider (2008) take Trudgill (2008) to task for his narrow view of identity. Trudgill regards identity as ‘irrelevant’ in sociolinguistic study because it is a consequence of other factors; Schneider claims that it is identity which is the motivating factor behind the language convergence and divergence of individuals in groups. However, there is no question for me that there is a mutually reinforcing relationship between broad social identity (the group-feelings of speech communities, nationalities, tribal communities or political factions, and so on) and a more personal view of individual singularity and attachment to these larger groupings. One area of longstanding interest in sociolinguistics where social practice and

individual perception and behaviour collide is in the management and negotiation of identity through *accommodation* (see Tuten 2008, who also indicates this connection).

Accommodation refers to the ways in which speakers are observed adjusting their speech patterns either towards (converging) or away (diverging) from their interlocutors. Early work (as ‘speech accommodation theory’) focused on the shifts in accent patterns and some dialectal lexicogrammatical choices, and also on shifts in register to mirror or establish difference between interlocutors. Later work (as ‘communication accommodation theory’) has explored more performative aspects of accommodation, for example where men are recorded using some of the norms of apparent feminine discourse patterns when contributing online to cancer support websites (Seale 2006).

Auer (2007: 109) makes the point that accommodation is used in two ways in sociolinguistics: as an interpersonal adjustment, for example in accent features, where interlocutors will shift their pronunciations towards each other in the course of a conversation; and as long-term dialect adjustment, where large-scale convergence is more permanent. Though Auer insists these two applications of the term are different, his own work suggests that individual shifts form the rich and complex basis for group shifts when repeated and diffused throughout a speech community (Auer 1998; Auer, Barden and Grosskopf 2000; Auer and Hinskens 2005). The earliest work (Giles and Powesland 1975; Giles 1970, 1973, 1977, 1979) deals with the social psychology of interpersonal adjustments, and later applications (Coupland 1984; Giles, Coupland and Coupland 1991) extend these insights into an account of macro-level linguistic change. It is clear, whether implicitly or sometimes explicitly, that a cognitive psychological dimension has been involved throughout.

Identity in psychology

Of course, identity has been a key notion in the field of psychology also for most of its history. There are many schematic definitions of identity but the notion of personality traits and types has become

paradigmatic, and the belief that personality remains fundamentally unchanged over an individual's life has been longstanding. One of the most established schemes of personal identity is characterised as the 'OCEAN' model, on the basis of the traits of Openness to new and varied experience; Conscientiousness in the form of organisation, self-discipline, dutiful and moral behaviour, planned action rather than spontaneity; Extraversion as enthusiasm, drivenness, energy and optimism, positiveness and sociability; Agreeableness in terms of compassion, sympathy, empathy, and co-operation; and Neuroticism manifest as pessimism, anger, anxiety, depression and vulnerability (Goldberg 1993). The contrary traits respectively are closedness, capriciousness, introversion, solipsism and emotional stability (note that the positive polarity of the last category is the wrong way round to maintain the 'OCEAN' acronym). These 'Big Five' traits are claimed to encompass just about all other personality attributes (see John and Srivastava 1999).

Though personality traits have increasingly been used in commerce and business for profiling potential employees, for example, the scheme has been criticised within the field of psychology (see Block 1995).

One of the main problems with a trait view of human personality and identity is its inflexibility. People are seen as immovable personality types rather than as possessing identities that can be adapted and reconfigured over time. More challengingly, some within the field of psychology have argued that personality can change and that the evidence for fixed traits is in fact a side-effect of the methodology for their investigation (Heatherton and Weinberger 1994; Caspi and Roberts 2001). Some of this challenge comes from an approach based in *narrative identity* (see McAdams et al 2006), which suggests that personality and identity are narrativised constructs based on the life-stories that individuals value and preserve. In this view, identity is not so much made manifest from a fixed pattern as performed according to the context at hand. This perspective accords with a more cognitively psychological view of identity.

Insights from cognitive science

One of the first insights developed within cognitive linguistics was the development of *prototypicality* in the processing of categorisation (see Lakoff 1987 for a review). Prototype effects undermine the classical view of discrete categorisation, since every attribute (or ‘trait’) can be regarded as being on a radial continuum with other, less good examples of the category in hand. For example, a screwdriver is usually a very good example of a carpentry tool, and a spoon is a less good example. But both can become very good examples of a paint-pot opener, given different circumstances. The essential point here is that every example of a concept can be adapted in different contexts, so that categories blend into each other without edges. It also entails a view of memory and reference in which referents are idealised cognitive models – schematic abstractions – rather than concrete memories of material objects and their properties.

Lakoff (1987) is careful always to talk about ‘prototype effects’ rather than ‘prototypes’ as objects. Gibbs (2003) goes further and notes that, though linguistic reference is abstract and generalised, in any given set of circumstances a referent becomes manifest as a prototype in mind for the purpose at hand. He cites much evidence for the provisionality, plasticity, contingency and temporary nature of prototypes, and argues that both linguistic understanding and language performance are ‘fundamentally adaptive to the moment-by-moment dynamics that arise in any interpretive situation’ (Gibbs 2003: 39).

Gibbs’ (2003) particular concern is with literary reading, but it is clear that his argument in general applies to all categories, including the sense of personality and identity set out above. Far from being fixed and immutable, a cognitive psychological view of personality must regard identity as multiply configurable and adaptive to the immediate physical context. Personality and identity, in Gibbs’ (2003: 37) terms, are ‘soft-assembled’ online and dynamically. Of course, it seems to be the case that individuals maintain a reasonably coherent and consistent sense of their own identities, but the argument here is that this consistency is performed and actively maintained rather than pre-existing. The consequence of this view is that the observable performance of identity might be expected to shift, given variable

circumstances – and my argument in this paper is that this is exactly how we can best account for the process of interpersonal accommodation in language.

A second useful insight from cognitive linguistics concerns our understanding of *deictic braiding* (Stockwell 2009: 127-31). Developed for the complex ontological relationships that exist in literary reading, where author, narrator, character, and reader are displaced from each other through written discourse, the connection between the deictic stance adopted by the reader and the deictic centre established linguistically by the other represented mind draws, of course, on natural, face-to-face patterns and processes. In a conversational dyad (and in multi-person discussions too), an individual adopts a cognitive stance appropriate to his sense of the language situation. This adoption is both a perceptual recognition on the one hand and a template for performativity on the other, with both aspects informing and reinforcing each other according to the fundamental cognitivist principle of embodiment (see, for example, Turner 1996, 2006). Drawing on the modelling projection potential of our human ability to imagine a ‘theory of mind’ in other humans (see Zunshine 2003, 2006; Belmonte 2008), we project a deictic centre for the other person that is richly motivated and complex. The mechanism for this projection can be regarded for analysis as the six ‘braids’ of deictic realisation and deictic centre maintenance which connect speaker and hearer:

- perceptual deixis (pronouns, demonstratives, definite articles and definite reference, and verbs of mental states)
- spatial deixis (locatives, spatial adverbs, distal demonstratives, verbs of motion)
- temporal deixis (locatives, temporal adverbs, tense and aspect)
- relational deixis (encoding of social position through lexicogrammar)
- textual deixis (self-referential textuality, iconicity, sense of texture)
- compositional deixis (interpersonal extratextual features).

The important point to recognise here is that an analytical description must proceed on the basis of one deictic centre or another

– in other words, the description must focus on what an individual perceives to be happening. Global knowledge of the entire situation – the normal viewing point of a sociolinguist looking at a transcription – will not provide an accurate approach.

An individual keeps a coherent deictic centre in mind by processing together all the different deictic braids that are being produced by his interlocutor: she in turn – at the other end of the thread, as it were – makes sociolinguistic choices in presenting her own cognitive stance along each dimension of deixis as above. According to work done within *deictic shift theory* (Duchan, Bruder and Hewitt 1995; Stockwell 2002), the deictic centre of all the interlocutors are kept consistently in mind by a range of maintenance devices and attention-sustaining features. Non-maintenance by relative quietness, moving away, or the non-performance of deictics which would otherwise sustain the activation of the other's deictic centre, leads gradually to a sense of deictic *decay*. For the discussion in the rest of this paper, the important consequence of this is that an individual might be presented with a concrete and deictically performed set of linguistic features initially, but these quickly become memorised and schematised as working memory moves on to other incoming matters.

The social psychology of accommodation

In almost all work on accommodation in sociolinguistics, there is an implicit assumption that both the convergence and divergence of linguistic features at the individual level is based on direct imitation and modelling of the interlocutor. Even, as briefly mentioned above, where macro-sociolinguistic shift is explored, there is a derived assumption that group behaviour is driven by a complex bundling of individuals' behaviour. The initial imitation is material and face-to-face, and though the subsequent modelling of the interlocutor's sociolinguistic positioning represents a certain level of abstraction away from this actualised material, the analysis is still usually conducted within the ongoing transcription or in the same close discourse situation.

This concern to remain close to the transcription data is, of course, not surprising, given the methodological paradigm within sociolinguistics for empirical and naturally-occurring discourse in real settings. However, I think that an unintended and unfortunate side-effect of this proper concern is that the materiality of the interlocutor is given undue emphasis in an understanding of what the individual in focus is doing in the situation. The observable and recorded speech of the interlocutor is an easy and empirically verifiable source for the linguistic behaviour of the focused individual. The interlocutor's speech patterns provide a simple benchmark for analysing and evaluating the relativity of the individual's shift either towards or away from that established and apparent norm. Later accommodation theory studies (see Ochs and Capps 2001; Auer and Hinskens 2005; Tuten 2008) explore not so much a simple convergence or divergence but the ways in which interlocutors establish their differential power 'pegs', but even these more complex studies rely on a concrete and immediate model for relative linguistic performance.

I would like to argue for a more cognitively sensitive analysis of these situations. In order to demonstrate this, I will firstly draw on a couple of typical examples of a sociolinguistic treatment of accommodation (from Bucholtz and Hall 2005), and then on a more extreme example (briefly commented on by Cameron 2008).

In a paper which takes a diametrically opposite view from my current position, Bucholtz and Hall (2005: 585) argue that 'identity is the product rather than the source of linguistic and other semiotic practices and therefore is a social and cultural rather than primarily internal psychological phenomenon'. They review several examples of transcriptions from a range of sociolinguistic studies, in order to demonstrate a sociocultural linguistic approach to identity: most of these examples feature what would be classified as instances of interlocutor accommodation. Bucholtz and Hall (2005: 607) believe that because 'identity is emergent in discourse and does not precede it, we are able to locate identity as an intersubjectively achieved social and cultural phenomenon'. I agree that identity is an emergent property of discourse, but I would argue that discourse in their perspective is skin-deep and focused shallowly on the materiality of the transcription, rather than encompassing essential 'internal

psychological phenomen[a]’. In what follows, I will re-examine two of their cited examples in order to demonstrate this.

The following short transcription is taken from Besnier (2004). The setting is a second-hand market in Tonga, a kingdom of 170 islands in the south Pacific just over 1500 miles from New Zealand. Tongan language transliteration is marked in italics and an English translation follows in square brackets.

- Seller: *Sai ia kia koe, Sonia.*
[‘Looks good on you, Sonia.’]
- Customer: Yeah- if it fits =
- Seller: ((ignoring customer’s contingency)) = Ni::ce. (10.0)
What size is it? (2.0)
- Customer: Eight. (3.0)
- Seller: Ohh. (4.0) Too small. (2.0)
‘E hao ‘ia Malia. (2.0) ‘Ia me’a. (2.0)
[‘It’ll fit Malia. I mean, what’s-her-name.’]
It’s might fit you, cuz it looks big!
- Customer: *‘Io?*
[‘Yes?’]
- Seller: Yeah! (2.0) The waist, look!
- Customer: I know-
- Seller: I think it’s one of those one that it has to show the bellybutton.
- Customer: No way!
- Seller: Aaaha-ha-haa!
- Customer: Haa-ha-hah!
- Seller: That’s the in-thing in New Zealand now. Even my kids say,
‘Mummy, see, it has to show the b-!’ Huh! I say, ‘No:::., no::!’ Ahahahuh-hh! Cuz that’s the look now!

(Besnier 2004: 29-30)

In their review of this study, Bucholtz and Hall (2005) focus on the use of code-switching to establish an identity in the marketplace. Under pressure from New Zealand, English is regarded in this commercial public setting as emblematic of a modern and

cosmopolitan lifestyle, and this aspirational identity is further instantiated by the particular topic of buying a piece of fashionwear. Aside from the examples of full inter-language code-switching, the passages in English are also characterised by a high degree of centralization (strongly associated with the New Zealand accent) of the vowels in ‘eight’, ‘fit’ and ‘mean’.

My central point in relation to this transcription derives from the fact that neither of the participants here are native English speakers nor from New Zealand – both are Tongans. Whatever model of New Zealand English they are imitating, it is clearly not present in the immediate setting, but instead is drawn from their existing and established senses of how an English-speaker would speak. The sociolinguistic transcript labels the participants by role (‘Seller’ and ‘Customer’), even though one of them is named (‘Sonia’), and while this is of course disciplinarily conventional, it also points to and reinforces the sociolinguist’s primary and exclusive interest in social role construction rather than personal self-awareness.

In fact, the transcript captures the moments in which the two individuals shift from their prior identities into performing their social roles as customer and seller. They begin by direct naming (‘Sonia’), other-oriented tokens (‘Looks good on you’, ‘might fit you’), and a question-answer structure all of which signal a relationship of personal familiarity and ease. The examples of Tongan-English code-mixing are all located in this first half of the transcript. The moment of transition into their performed roles comes as the customer’s questioning ‘*Io?*’ is met with the directly comparable ‘Yeah!’ that is switched and diametrically assertive. From then on, English predominates exclusively, the question-answer structure disappears, and the seller’s token-orientation shifts from other- to self-oriented (‘Even my kids say...’). All of these are examples of the establishment of a deictic centre that is characterised primarily by commercial role and by aspiration to modernity – a constructed persona to which the other interlocutor responds in kind.

I would like to term this shift in persona-performance a shift in identity – this is possible only under a definition of identity that is plastic and provisional. As the two participants perform their shift,

they are not moving towards each other in any material sense, but both to a mental representation of an aspirational prototype that they project as being shared by the other. Their language is governed by a cognitive recipient design, and reinforced both positively by the echoing strategies of the other and negatively by the absence of any signals of confusion or eccentricity. The only stall on the process is the relative disfluency in English which is signaled in the grammatical slips ('it's one of those one'), the infilling with sound-effects and gestures ('Aaaha-ha-haa!', "'No:::'"), and the quotation of a readymade piece of English in the form of a quoted insert ('Even my kids say, "Mummy...").

In the process of this shift from prior to public identity, the prior identity suffers from decay and is gradually effaced by the public identity. This, of course, will be temporary and provisional, and the customer at least no doubt reverted back to performing her prior identity configuration shortly after the transcript ends.

The features I have briefly set out above are examples of social deictic markers, and it is along the braid of social deixis that most of the shift is effected, of course. The two participants lose their prior sense of self as they reconfigure themselves in the new setting. The customer even draws attention to her plasticity by evoking social deictic shifts in her idiomatic use of New Zealand English ('That's the in-thing in New Zealand now'), immediately followed by her re-designation from her projected children's deictic viewpoint as 'Mummy', and their coy elision of 'bellybutton' ('the b-!') as a semi-taboo term, and then her own relative disfluency in English ('Huh! I say') to signal that the modern cosmopolitan image of herself just evoked to the seller must be understood against the real modernity of her children.

Throughout the transcript, the model being accommodated is complex, non-present and non-actual.

In this second example reviewed by Bucholtz and Hall (2005), a group of Korean American young men are discussing race issues in their lives and neighbourhood.

Jin: i think white people just don't keep it real and that's why
Dave: that is = that's true man?
Jin: cause that's why they always back stabbin like my
roommate who
wasn't gonna pay the last month's // rent
JH: white.
Jin: he kicks us out [of
Eric: [the prototypical *whitey*.
Jin: ye:::ah ma::n?
JH: no social skills.
Jin: but that's not true for everyone i don't think.
EC: uh huh
Jin: cause all those ghetto whiteys in my neighborhood i think
they're cool
(Chun 2001: 60)

In the US, Asian-Americans do not have the same high-profile English discourse as, for example, African-Americans, but in this example they are seen to draw on the latter's symbolic value as a resistant discourse to white American English. There are several clear uses of the features of African-American English, but it is noticeable that the particular features that are performed here are from a non-localised, generic set of prototypical 'black' features:

- idiomatic phrases ('keep it real')
- tags ('man')
- lexis ('whitey', 'cool')
- pronunciation ('ye:::ah ma::n', 'uh huh')
- copula deletion ('they always back stabbin')
- contraction ('gonna')
- clefting and fronting ('all those ghetto whiteys I think they're cool').

Other prominent and stereotypical features are not used, even where there was an opportunity. For example, a thorough final-consonant reduction or deletion does not appear in: ('just(t)', 'that(t)'s why', 'that(t)'s true', 'las(t) months'). Similarly, the copula deletion is not thorough: "'they'(re) cool'.

As with the Tongan example, the accommodation is clearly not towards the discourse being modelled by any of the participants who are actually there. Instead, a schematised mental model of African-American English in its prototypical set is being targeted, assimilated and assembled by these speakers. Again, the assembly is contingent and temporary, soft-assembled for this specific collision of participants and topic. Jin, in particular, becomes identifiably black (at least linguistically) as the passage develops, ending up in the final utterance of the transcript adopting the perspective of African-Americans on the ‘ghetto whiteys’ in black idiom at the same time as his self-reference (‘in my neighborhood’) aligns strongly with this identity.

Bucholtz and Hall (2005: 591), aside from other arguments, endorse the provisionality and constructed nature of identity: its ‘actual manifestation in practice is dependent on the interactional demands of the immediate social context. Such interactions therefore highlight what is equally true of even the most predictable and non-innovative identities: that they are only constituted as socially real through discourse, and especially interaction.’ Identity, in their terms, is an emergent property rather than a definitional source. However, this formulation, in adopting a directional metaphor of cause-and-effect, does not capture the mutually reinforcing nature of a *strong sense* of identity with its linguistically performed features. To demonstrate this point, I will finish with an extreme example of accommodation that defies traditional sociolinguistic description.

Extreme accommodation and cognitive sociolinguistics

One of the most popular clips on the online video-site YouTube in the summer of 2008 was an excerpt of the former England football manager, Steve McClaren, being interviewed on Dutch television (see McClaren 2008).

Extraordinarily, the British interviewee affected a Dutch accent throughout, as reported on the *Sky News* website at the time:

Schtop! Mac's Gone Double Dutch

Steve McClaren ditched his Yorkshire accent and began speaking with a Dutch twang during bizarre television interviews.

The former England boss, now manager of FC Twente in the Netherlands, adopted broken English while answering questions.

Among the gems were: 'We are not just, what you call, underdogs, but massive underdogs. Everything is there for a surprise.'

At one point he said mid-sentence 'The only thing you can-a-do...'

And he added: 'On the night, maybe we get the luck.'

This was despite the Dutch interviewer speaking in perfect English - and the fact he has only been in the job a few weeks.

McClaren - speaking ahead of his team's Champions League defeat to Arsenal - was sacked as England boss after failing to qualify for the European Championships.

His new accent had puzzled viewers on YouTube, who have delivered their own harsh verdict.

However, it is not unusual for footballers to eventually shed their native tongue after spells in England.

Dane Peter Schmeichel famously spoke with a Mancunian accent during his time with Manchester United and compatriot Jan Molby sounded like a scouser while with Liverpool.

Sky News website (Friday August 15, 2008)

The report goes on to cite some examples from McClaren's speech that exemplify his 'Dutch' accent:

Arsenal becomes 'Arrshenal', and the Champions League fixture against them 'is big newsh now in Englisch media.'

'Championsh League, Liverpool or Arshenal, I thought one of them we would draw and it is Arshenal I think.'

'To experiensh big gamesh, Championsh League... Arshenal... The Emiratesh... will be fantashtic for the playersh, not just for now but for the future ash well. I shay I think we are not just... what you call?... underdogsh but mashive underdogsh.'

Here the report describes (and in fact slightly exaggerates) some of the accent features prototypically associated with Dutch non-native speakers of English, particularly the affricating of sibilants. However, it also captures some of the other lexicogrammatical features that are

less central examples: the non-fluent syntax involving clefting, and the listing of noun-phrases typical of non-native speakers who lack verb-phrase knowledge.

Here is my transcription of the interview between Steve McClaren (SM) and the female Dutch television interviewer (TV). :

SM ha – I sort of knew ah when I came here ah Champion’s league ah Liverpool or Arsenal ah I thought maybe one of them we would draw and ah it is Arsenal I think [whew] one of the toughest teams - in the – draw – and I think it will be very very difficult for our players – but also a great experience ah we have a young team – and to experience ah big games – Champion’s league – Arsenal – at home – the Emirates – will be fantastic for the players – not just for now but in the future as well

TV It must be very special for you personally

SM Yeah, I errr, ah – of course – it’s one of the reasons I came because – a good team and they qualify – Champion’s league so it’s one of the reasons I came to draw Arsenal one of the big teams [whew] is ah exciting for everybody

TV Arsenal is a big team – what are your chances against Arsenal?

SM I say I think we ah are not just what do you call underdogs but massive underdogs ah I think there is no expectations for FC Twente so ah everything is for a surprise ah we have nothing to lose so we go into the game nice relaxed – enjoy the experience

TV But you know the English football by heart

SM I do I know the team very well and ah the Arsenal very well so ah it is nothing new and I think the players they know all the Arsenal players – they watch them – em Premier league ah I think they are one of the top teams in Europe and it will be good for our players to play against that quality because – ah anything after that may be a little easier

TV You have a very talented group but they do not have a lot of experience in big European games – is it going to be a problem?

SM Ahm we will see on the night ah – some people – [whew] they grow – with the experience and I hope the players will take the opportunity to use it and to grow to become better players

Aside from the affrication, McClaren also adopts a rhotic pronunciation (not native to his own Yorkshire accent), a Dutch intonation pattern, and a careful enunciation that articulates every consonant where he would perhaps naturally elide some sounds. He also adopts a syntactic hedge ('I sort of know') and clefting ('one of them we would draw') that are typical of non-native English speakers. From time to time he intersperses his language with '[whew]' – a puffing out of breath accompanied by a shrug and open-armed gesture that British people associate with continental Europeans (actually, more usually regarding it as a stereotypically French mannerism).

Cameron (2008) suggests that what McClaren is doing is closest sociolinguistically to what Rampton (1995) calls 'crossing': the adoption of speech markers of another group in order to lay claim to some of that group's attributes. The most well-known example that Cameron cites is that of white youths affecting black English patterns as a form of coolness or toughness. Rampton's notion of crossing is psychologically plausible because it involves a contingent and soft-assembled sense of identity: 'crossers' know they are not actually becoming either physically or ethnically black, and McClaren denied afterwards that he considered himself to have 'gone native' as a Dutch football manager.

However, what marks this performance out from standard accommodation or crossing is the degree of travel beyond the target that McClaren effects. He produces idioms that are not naturally idiomatic in English ('everything is for a surprise') that have a foreign-sounding flavour but in fact are never produced by Dutch people either. He aims not towards the relative fluency of the physically-present television interviewer, but instead towards a model of someone who is a rather disfluent English speaker, one who needs to list noun-phrases for lack of vocabulary or clausal syntax, one who omits prepositions because of uncertainty in selecting the correct one ('they qualify – Champion's league'), one who does not display normal native-speaker elisions (of 'it is', or 'I will' or 'we are') but

articulates them fully, and one who replaces his British cultural paralinguistic sounds and gestures with foreign ones. Most strikingly of all, at one point he appears to efface his own deictic centre so effectively that he forgets that he is British, when he declares: ‘I say I think we ah are not just what do you call underdogs but massive underdogs’. Of course, it is not ‘you’ Dutch who have this idiom but ‘you’ British, whom McClaren here deictically marks as other.

Such a performance of identity might be an extreme example, but it serves to demonstrate on a continuum from more standard examples that the traditional approach to accommodation cannot account for the phenomenon satisfactorily enough. A measure needs to be taken of personal identity formation based on cognitive psychological insights *as well as* social practice and performativity, and the relationship between the two is mutually reinforcing and embodied rather than causal or motivational. Identity is certainly performed and of course there are consistencies across an individual’s performative repertoire that they and observers will regard as consistent and stable. However, even generic or familiar circumstances are locally variable, and it is one of our most human capacities that we are adaptable in different situations. This adaptability of identity – made manifest through patterns of accommodation – can be accounted for within a thoroughly integrated cognitive sociolinguistics where subjective consciousness and social practice collide in the field of language. As I hope very briefly to have indicated, there are positive benefits from an integrated approach, in making us realise that we are both more peculiar and more adaptable than we think.

References

- Antaki, C. and Widdicombe, S. (eds) (1998) *Identities in Talk*. London: Sage.
- Auer, P. (2007) 'Mobility, contact and accommodation' in C. Llamas, L. Mullany and P. Stockwell (eds) *The Routledge Companion to Sociolinguistics*, London: Routledge, pp.109-15.
- Auer, P. (ed.) (1998) *Code-Switching in Conversation: Language, Interaction, and Identity*. London: Routledge.
- Auer, P. and Hinskens, F. (2005) 'The role of interpersonal accommodation in a theory of language change', in P. Auer, F. Hinskens and P. Kerswill (eds), *Dialect Change*, Cambridge: University Press.
- Auer, P., Barden, B. and Grosskopf, B.E. (2000) 'Long-term linguistic accommodation and its sociolinguistic interpretation: evidence from the inner-German migration after the *Wende*', in K. Mattheier (ed.) *Dialect and Migration in a Changing Europe*, Frankfurt: Peter Lang, pp.79-98.
- Bauer, L. (2008) 'A question of identity: a response to Trudgill', *Language in Society* 37 (2): 270-3.
- Belmonte, M.K. (2008) 'Does the experimental scientist have a "Theory of Mind"?' *Review of General Psychology* 12 (2): 192-204.
- Besnier, N. (2004) 'Consumption and cosmopolitanism: practicing modernity at the secondhand marketplace in Nuku'alofa, Tonga', *Anthropological Quarterly* 77 (1): 7-45.
- Block, J. (1995) 'A contrarian view of the five-factor approach to personality description', *Psychological Bulletin* 117: 187-215.
- Bucholtz, M. and Hall, K. (2005) 'Identity and interaction: a sociocultural linguistic approach', *Discourse Studies* 7 (4-5): 585-614.
- Bucholtz, M., Liang, A. and Sutton, L. (eds) (1999) *Reinventing Identities: The Gendered Self in Discourse*. Oxford: Oxford University Press.
- Cameron, D. (2008) 'Going Dutch', *Critical Quarterly* 50 (4): 90-93.
- Caspi, A. and Roberts B.W. (2001) 'Personality development across the life course: the argument for change and continuity' *Psychological Inquiry* 12 (2): 49-66.

- Chun, E.W. (2001) 'The construction of White, Black, and Korean American identities through African American Vernacular English', *Journal of Linguistic Anthropology* 11 (1): 52–64.
- Coupland, N. (1984) 'Accommodation at work: some phonological data and their implications', *International Journal of the Sociology of Language* 46: 49-70.
- Duchan, J.F., Bruder, G.A. and Hewitt, L.E. (eds) (1995) *Deixis in Narrative: A Cognitive Science Perspective*. Hillsdale: Lawrence Erlbaum.
- Dyer, J. (2007) 'Language and identity', in C. Llamas, L. Mullany and P. Stockwell (eds) *The Routledge Companion to Sociolinguistics*. London: Routledge, pp.101-8.
- Eckert, P. (2000) *Language Variation as Social Practice*. Oxford: Basil Blackwell.
- Eckert, P. and Rickford, J.R. (eds) (2001) *Style and Sociolinguistic Variation*. New York: Cambridge University Press.
- Gibbs, R.W. (2003) 'Prototypes in dynamic meaning construal', in J. Gavins and G. Steen (eds) *Cognitive Poetics in Practice*, London: Routledge, pp. 27-40.
- Giles, H. (1970) 'Evaluative reactions to accents', *Educational Review* 22: 211-27.
- Giles, H. (1973) 'Accent mobility: a model and some data', *Anthropological Linguistics* 15: 87-105.
- Giles, H. (1977) *Language, Ethnicity and Intergroup Relations*. New York: Academic Press.
- Giles, H. (1979) 'Ethnicity markers in speech', in K. Scherer and H. Giles (eds) *Social Markers in Speech*, London: Cambridge University Press, pp. 251-89.
- Giles, H. and Powesland, P.F. (1975) *Speech Style and Social Evaluation*. London: Academic Press.
- Giles, H., Coupland, J. and Coupland, N. (eds) (1991) *Contexts of Accommodation: Developments in Applied Sociolinguistics*. Cambridge: Cambridge University Press.
- Goldberg, L.R. (1993) 'The structure of phenotypic personality traits', *American Psychologist* 48: 26-34.
- Heatherton, T.F. and Weinberger, J.L. (eds) (1994) *Can Personality Change?* Washington, DC: American Psychological Association.
- John, O.P. and Srivastava, S. (1999) 'The Big-Five trait taxonomy: history, measurement, and theoretical perspectives', in L.A. Pervin

- and O.P. John (eds) *Handbook of Personality: Theory and Research* (Vol. 2), New York: Guilford Press, pp.102–38.
- Johnstone, B. (1996) *The Linguistic Individual: Self-Expression in Language and Linguistics*. New York: Oxford University Press.
- Kristiansen, G. and Dirven, R. (eds) (2008) *Cognitive Sociolinguistics: Language Variation, Cultural Models, Social Systems*. Berlin: de Gruyter
- Lakoff, G. (1987) *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- Le Page, R.B. and Tabouret-Keller, A. (1985) *Acts of Identity: Creole-Based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.
- Llamas, C., Mullany, L. and Stockwell, P. (eds) (2007) *The Routledge Companion to Sociolinguistics*. London: Routledge.
- McAdams, D.P., Bauer, J.J., Sakaeda, A.R., Anyidoho, N.A., Machado, M.A., Magrino-Failla, K., White, K.W. and Pals, J.L. (2006) ‘Continuity and change in the life story: a longitudinal study of autobiographical memories in emerging adulthood’, *Journal of Personality* 74 (5): 1371-400.
- McClaren, S. (2008) ‘Shteeve McClaren’, *YouTube* several postings, for example <<http://www.youtube.com/watch?v=2ZnoP4sUV90>>, and <<http://www.youtube.com/watch?v=VWptErAWrmQ>>, last accessed February 2010. This clip is also available from Peter Stockwell’s personal website at the University of Nottingham: follow <<http://www.nottingham.ac.uk/english>>
- Ochs, E. and Capps, L. (2001) *Living Narrative: Creating Lives in Everyday Storytelling*. Cambridge, MA: Harvard University Press.
- Rampton, B. (1995) *Crossing: Language and Ethnicity among Adolescents*. London: Longman.
- Schneider, E.W. (2008) ‘Accommodation versus identity? A response to Trudgill’, *Language in Society* 37 (2): 262-7.
- Seale, C. (2006) ‘Gender accommodation in online cancer support groups’, *Health* 10 (3): 345-60.
- Stockwell, P. (2002) *Cognitive Poetics: An Introduction*. London: Routledge.
- Stockwell, P. (2007) *Sociolinguistics* (2nd edition), London: Routledge.
- Stockwell, P. (2009) *Texture: A Cognitive Aesthetics of Reading*. Edinburgh: Edinburgh University Press.

- Trudgill, P. (2008) 'Colonial dialect contact in the history of European languages: on the irrelevance of identity to new-dialect formation', *Language in Society* 37 (2): 241-54.
- Turner, M. (1996) *The Literary Mind: The Origins of Thought and Language*. New York: Oxford University Press.
- Turner, M. (2006) *The Artful Mind: Cognitive Science and the Riddle of Human Creativity*. Oxford: Oxford University Press.
- Tuten, D.N. (2008) 'Identity formation and accommodation: sequential and simultaneous relations', *Language in Society* 37 (2): 259-62.
- Zunshine, L. (2003) 'Theory of mind and experimental representations of fictional consciousness', *Narrative* 11: 270–91.
- Zunshine, L. (2006) *Why We Read Fiction: Theory of Mind and the Novel*. Columbus: Ohio State University Press.

Mariette Meunier-Crespo & Luu Dinh Phuc
Université Jean Moulin Lyon 3 – CEL, EA 1663

Présentation d'une recherche jurilinguistique sur le Code des délits et des peines (1810) et sur le Code pénal (1994)

Le langage du droit existe parce qu'il n'est pas compris (Gérard Cornu)¹

Préliminaires

Ces quelques lignes veulent préciser le point de vue choisi pour examiner notre objet d'étude : un corpus constitué de deux textes législatifs français majeurs, le Code des délits et des peines (1810) et le Code pénal (1994). Ces textes ne sont un objet linguistique pertinent qu'en fonction du point de vue qui les déclare tels. Nous allons donc nous livrer à un salutaire exercice de modestie (bien réelle !) qui consiste à passer en revue nos compétences et incompétences de linguiste désireux d'observer et d'analyser de tels textes. Que l'on ne se s'étonne pas de la fragmentation des activités du linguiste. Elle n'a de valeur qu'heuristique. Il ne saurait y avoir de cloison étanche entre la terminologie, la stylistique et la syntaxe, toute étude un peu poussée dans l'un ou l'autre de ces domaines finissant par rejoindre les deux autres.

¹ Gérard Cornu, *Linguistique juridique*, Montchrestien, Paris, 1990, préface.

Le point de vue terminologique

Il semble évident que les recherches linguistiques concernant la langue juridique se concentrent d'abord du côté du signifié, sur la terminologie. La boutade de Gérard Cornu « *le langage du droit existe parce qu'il n'est pas compris* » trouve sans peine un large écho dans le public. Comme Cornu l'explique, « *il y a un langage du droit parce que le droit donne un sens particulier à certains termes* »².

L'une des caractéristiques de la terminologie juridique par rapport aux autres terminologies spécialisées est en effet l'importance de la polysémie. Le caractère souverain des textes législatifs leur donnant le pouvoir de nommer, le législateur crée donc les termes qui correspondent aux notions juridiques nouvelles. Il apparaît que c'est de création sémantique qu'il s'agit le plus souvent : au lieu d'inventer de toutes pièces un nouveau terme, ou d'en forger un par composition ou dérivation, la création sémantique consacre dans un emploi nouveau des mots de la langue usuelle. Ces mots sont ainsi transformés en termes juridiques et deviennent de ce fait des polysèmes. La polysémie est considérée comme un obstacle majeur à la clarté du langage juridique et c'est donc un terrain de choix pour l'activité des linguistes terminologues.

Les juristes n'ont cependant pas attendu l'essor de la terminologie au XX^e siècle comme branche de la linguistique pour repérer les problèmes de polysémie, élaborer leurs définitions et proposer des solutions. Il est remarquable que les lexiques et les dictionnaires de droit soient rédigés par des professionnels du droit et, sauf exception, pas par des linguistes³. Il n'y a rien d'étonnant à cela car le propre de la terminologie, c'est sa démarche sémasiologique : il s'agit de partir d'une organisation conceptuelle pour aller vers son expression linguistique. Il en résulte que les travaux concernant la terminologie juridique ne se limitent généralement pas à de simples définitions. Ils cherchent à présenter les notions dans leur aspect technique bien sûr, mais aussi dans leur perspective historique et avec leur dimension actualisée. C'est ainsi qu'ils rendent le plus de services aux praticiens du droit et aux étudiants. Seuls des juristes peuvent réellement mener

² Gérard Cornu, 1990, p. 20.

³ Pour ne citer qu'un exemple, il est significatif que l'ouvrage de référence en matière de terminologie juridique soit l'ouvrage d'un professeur de droit qui s'est intéressé à la linguistique (et non l'inverse) : Gérard Cornu, *Vocabulaire juridique*. 8^e édition mise à jour, Quadriga, 2007.

à bien l'intégralité de ce travail terminologique car ce sont eux qui maîtrisent le mieux les réseaux conceptuels de leur matière.

Le linguiste est cependant utile pour les traductions. Comme pour tout document technique qui lui pose des problèmes de terminologie, il se livre à une recherche documentaire et questionne les personnes expertes du domaine. Mais la traduction juridique pose des problèmes plus sérieux dans la mesure où, parfois, la traduction du document produit des effets juridiques. Le linguiste doit donc impérativement avoir recours en dernier ressort au juriste et pour établir le dialogue avec lui, une formation minimum en droit lui est indispensable. Tout enseignant universitaire en charge de l'enseignement de la terminologie juridique dans différentes langues est conscient de cette nécessité.

Le point de vue stylistique

Au vu de nos insuffisances dans le maniement des réseaux conceptuels du droit, nous n'envisageons pas un point de vue terminologique sur notre objet d'étude. Il en résulte qu'un point de vue stylistique n'est pas davantage dans nos compétences. Il est paradoxal pour un linguiste d'avancer un tel propos, attendu que la stylistique moderne englobe les champs connexes de la sémantique, de la rhétorique et de la pragmatique. Cette stylistique ou linguistique dite textuelle (relayant l'analyse du discours inaugurée dans les années 60) se charge d'étudier l'entrelacs des signifiés dans les énoncés longs. Elle repousse les limites fixées par le structuralisme au domaine du mot et de la phrase.

Mais l'une des plus remarquables particularités du discours juridique est que les juristes théoriciens produisent du métalangage sur leurs propres textes : ils les commentent, les expliquent, en retracent l'histoire, la formation, en soulignent les implications, en dévoilent les secrets de fabrication. Ils n'ont pas grand-chose à apprendre des dénotations et connotations qu'étudient les linguistes. Tout texte juridique renvoie à des niveaux de langue, des états de langue et des sources d'énonciation très variés. Ce sont les juristes qui, par leur formation, maîtrisent les univers de référence auxquels renvoient les signes de leurs textes.

Les linguistes analysent le signifiant, c'est-à-dire la matière même du texte. Cependant, la langue est l'outil de travail des juristes. Ils connaissent les règles de la rhétorique, ancienne et nouvelle. L'application des « séquences argumentatives » de la linguistique textuelle n'a guère de secrets pour eux. Ils connaissent l'art de persuader, savent l'expliquer à leurs étudiants, et sont à même d'écrire des traités sur le sujet. Les juristes étudient et utilisent leurs textes et les signes qui les composent dans leurs deux dimensions : signifiante et signifiée. Ils font même de la « pragmatique » au sens linguistique, sans le savoir, et bien avant que les linguistes n'y aient songé.

En effet, est juridique tout discours qui a pour objet la création du droit. En droit, dire, c'est donc en même temps faire. Les linguistes pensent aussitôt à la théorie des *speech acts*, ou actes de langage et au célèbre titre d'Austin : « *Quand dire c'est faire* », qui traduit l'anglais « *how to do things with words* »⁴. (Ex : je te baptise, je t'interdis de...). Cette théorie pragmatique du langage s'est édifiée sur le fait que l'on puisse agir par le moyen du langage. Bien avant la théorie des actes de langage d'Austin, Reinach⁵ avait déjà découvert la dimension pragmatique du langage en isolant les *actes sociaux* qu'Austin appellera les *actes illocutionnaires*. On parle donc de force illocutoire ou illocutionnaire d'un énoncé (traduction de l'anglais *illocutionary*). Les énoncés de notre corpus font typiquement partie de la catégorie des *énoncés performatifs* c'est-à-dire investis de force illocutoire, décrits par Austin : ce sont des règles qui ne décrivent pas le réel mais correspondent à l'exécution d'une action. Ce caractère performatif est explicite car il faut aussi que les circonstances de cette action soient appropriées, c'est à dire qu'il faut déterminer les conditions d'application de ces règles.

Le point de vue actuel de la pragmatique du langage semblerait un bon point de vue pour aborder notre corpus. Mais il ne suffit pas d'étudier cet aspect en soi, car les juristes le font déjà fort bien en produisant leurs textes ou en évoluant dans leurs méandres. La seule spécificité du linguiste reste en dernière analyse de retourner au signifiant, comme on retourne à la lettre : l'étude des structures syntaxiques.

⁴ John Langshaw Austin, *Quand dire c'est faire*, 1962, trad. fr. 1970, rééd. Seuil, coll. « Points essais », 1991.

⁵ Adolf Reinach, *Les fondements a priori du droit civil*, Vrin, « Bibliothèque des Textes Philosophiques », 2004.

Le point de vue syntaxique

La syntaxe du discours législatif appartient au français. Ses spécificités sont celles des langues de spécialité en général. On s'accorde pour relever une prédominance des tournures passives, une tendance à la nominalisation, et à l'effacement des marques de l'énonciateur, en particulier l'absence de déictiques et la simplification des formes verbales. Gérard Cornu (1990) a lui-même fait le tour de la question. Il faut chercher, au-delà de l'étude syntaxique en elle-même (le simple constat d'un fonctionnement), ce qui fait la spécificité de la syntaxe juridique. Cornu nous met sur la voie : l'intérêt de l'analyse syntaxique est, dit-il, de « *chercher dans les mots et les énoncés les traces que laisse la démarche de la pensée juridique* »⁶.

Nous nous intéresserons donc aux structures qui traduisent syntaxiquement leur démarche, pour montrer, selon le souhait de Cornu, ce que la forme peut apporter au fond. Depuis Jakobson, la poétique est l'étude de l'énonciation du point de vue de la forme du message. Nous n'oserons pas plus que Gérard Cornu, parler de poétique juridique, mais linguistiquement parlant, c'est bien de cela qu'il s'agit.

Le point de vue logique de la règle juridique

Pour un juriste, qu'est-ce qu'une règle juridique⁷ ? Voici une des définitions possibles du droit pénal : « *l'ensemble des règles juridiques qui organisent la réaction de l'Etat vis-à-vis des infractions et des délinquants* »⁸.

Cette définition réunit les deux éléments nécessaires à la règle pénale : l'infraction, et la peine. Gérard Cornu (toujours lui) parle de la double détermination de la règle juridique, dans le sens où une règle

⁶ « *L'analyse syntaxique cherche dans les mots et les énoncés les traces que laisse la démarche de la pensée juridique. Elle aide à reconstituer le processus de la pensée par le repérage, dans le discours, des marques qui le jalonnent. L'analyse linguistique est un auxiliaire de l'analyse logique* », Gérard Cornu 1990 p. 41.

⁷ « *La juridicité du discours tient à sa finalité. Est juridique tout discours qui a pour objet la création ou la réalisation du droit. Ce critère commande la logique et le ton du discours* », Gérard Cornu p. 21.

⁸ Roger Merle et André Vitu, *Traité de droit criminel, tome 1 : Problèmes généraux de la science criminelle. Droit pénal général*, 7e édition, Cujas, 1997.

est une règle en ce qu'elle détermine : « *Il n'y a de règle que s'il y a détermination du CAS et détermination de la solution qui lui est applicable, double détermination de la situation et de l'effet de droit qui lui est attaché* »⁹. Il s'agit pour le législateur d'établir une infraction (un texte d'incrimination), suivie de son effet juridique : une peine (un texte de pénalité). L'ensemble acquiert la qualité d'énoncé performatif explicite dont il tire sa force illocutoire.

De façon plus épurée, il s'agit de définir un cas de figure, puis sa solution. La structure logique de l'article de loi pénale se présente donc majoritairement de la façon suivante : si A, alors B. A étant un préalable notionnel, nécessaire à la réalisation de B.

Une simple lecture des deux Codes montre une répartition fonctionnelle différente des structures syntaxiques réalisant ce schéma logique. Avant d'en esquisser un aperçu, précisons qu'en ce qui concerne le problème complexe (car fort ancien) de la terminologie et des catégories grammaticales nous nous référerons à la grammaire rénovée du français de Marc Wilmet¹⁰.

Aperçu des réalisations syntaxiques de la structure logique : si A, alors B

On distingue trois structures syntaxiques, classées ici par ordre de fréquence :

1 – A est un « complément de temps, de circonstance, de but » (complément adverbial circonstanciel de la prédication¹¹)

Art. 148 : Dans tous les cas exprimés au présent paragraphe, celui qui aura fait usage des actes faux sera puni des travaux forcés à temps. (1810)

Art. 131-25 : En cas de condamnation à une peine de jours-amende, le montant global est exigible à l'expiration du délai correspondant au nombre de jours-amende prononcés. (1994)

⁹ Gérard Cornu, 1990, p. 288.

¹⁰ Marc Wilmet, *Grammaire rénovée du français*, De Boeck Université, Bruxelles, 2007.

¹¹ Marc Wilmet propose de réduire le foisonnement des circonstances de lieu, de temps, de cause, de manière, de moyen, de prix, de mesure, de but, d'accompagnement, de condition, de conséquence, d'opposition, de comparaison, etc. en ne parlant plus que de *complément adverbial circonstanciel de la prédication*. Au-delà du changement de terminologie, il s'agit de se dégager de la sémantique pour se concentrer sur la fonction syntaxique.

Art. 133-17 : **Pour** l'application des règles sur la réhabilitation, la remise gracieuse d'une peine **équivaut** à son exécution. (1994)

Ces occurrences sont très minoritaires dans les 2 corpus.

2 – A est introduit par des « propositions circonstancielles si..., quand..., lorsque... » (sous-phrases conjonctives)

On trouve dans le Code de 1810 des tournures plus variées que dans celui de 1994, les tournures avec « si » apparaissant majoritairement.

Art. 51 : **Quand** il y aura lieu à restitution, le coupable sera condamné en outre, envers la partie, à des indemnités (...)

Art. 94 : **Si** cette réquisition ou cet ordre ont été suivis de leur effet, le coupable sera puni de mort et ses biens seront confisqués.

Art. 178 : **Dans le cas où** la corruption aurait pour objet un fait criminel emportant une peine plus forte que celle du carcan, cette peine plus forte sera appliquée au coupable

Dans le Code de 1994, la tournure s'uniformise majoritairement avec « lorsque » :

Art.131-10 : **Lorsque** la loi le prévoit, un crime ou un délit peut être sanctionné d'une ou de plusieurs peines complémentaires (...)

Art. 131-32 : **Lorsque** l'interdiction de séjour accompagne une peine privative de liberté sans sursis, elle s'applique dès le commencement de cette peine (...)

Une première observation livre donc une tendance à la généralisation et à la simplification des sous-phrases conjonctives.

3 – A et B entrent majoritairement dans le schéma syntaxique de la prédication attributive : Sujet + copule + prédicat

En termes syntaxiques, A (le cas) est le sujet grammatical, et B (la solution) est la prédication attributive (la voix passive entrant dans ce type de prédication). Une première observation de la typologie des sujets livre deux modèles de construction :

- **Sujet sur le modèle : Nom + déterminant non phrastique**

Très peu d'occurrences sur ce modèle dans le Code de 1810 :

Art. 16 : *Les femmes et les filles **condamnées aux travaux forcés** n'y seront employées que dans l'intérieur d'une maison de force.*

Beaucoup plus d'occurrences dans le Code de 1994 :

Art. 131-9 : la peine **de jours-amende** ne peut être prononcée cumulativement avec la peine d'amende.

Art. 131-36-2 : Les mesures **de surveillance** applicables à la personne condamnée à un suivi socio-judiciaires sont celles prévues à l'article 132-44.

- **Sujet sur le modèle : Nom + déterminant phrastique**

Le déterminant phrastique est une relative (sous-phrase pronominale)¹².

Il s'agit là de la majorité des occurrences dans le Code de 1810, avec tendance à la thématization (Le prédicat apparaît en premier, ce qui a pour effet de mettre en valeur la peine).

Art. 196 : *Tout fonctionnaire public **qui sera entré** en exercice de ses fonctions sans avoir prêté le serment, pourra être poursuivi, et sera puni d'une amende de seize francs à cent cinquante francs.*

Art. 151 : *Sera puni de la même peine **celui qui aura fait usage** de la pièce fausse.* (thématisation du prédicat).

Ce type d'occurrences est minoritaire dans le Code de 1994 :

Art. 121-3 : ***Les personnes physiques qui n'ont pas causé directement le dommage, mais qui ont créé ou contribué à créer la situation qui a permis la réalisation du dommage ou qui n'ont pas pris les mesures permettant de l'éviter sont responsables pénalement.***

Art. 121-7 : *est complice d'un crime ou d'un délit **la personne qui sciemment, par aide ou assistance, en a facilité la préparation ou la consommation*** (thématisation du prédicat)

- **Et les définitions ?**

Signalons pour terminer le cas particulier où seul A apparaît (la simple définition du « cas »). Une définition est, pour le non juriste, *la détermination des limites de l'extension d'un concept* (Lalande, Dictionnaire critique). Ce type de définition est rarissime dans notre corpus car la définition juridique d'une incrimination ne s'établit que par la sanction qui l'accompagne. L'infraction, juridiquement, est ce qui est interdit par la loi sous la menace d'une peine.

Art. 37 : ***La confiscation générale** est l'attribution des biens d'un condamné au domaine de l'état.*

¹² Marc Wilmet (2007), p. 255.

Art. 132-74 : L'escalade est le fait de s'introduire dans un lieu quelconque, soit par-dessus un élément de clôture, soit par toute ouverture non destinée à servir d'entrée.

Le plus souvent, ce type d'occurrence qualifié de définition n'est que la première partie de la structure logique restée en suspens, et dont la suite apparaît plus loin. Le cas et la solution peuvent donc ne pas être formellement réunis dans le même article. La peine applicable se trouve plus loin :

Art. 331-1 : Le vol est la soustraction frauduleuse de la chose d'autrui

Art. 311-3 : Le vol est puni de trois ans d'emprisonnement et de 45000 euros d'amende (1994)

Conclusion

Notre remarque du début concernant l'aspect heuristique de nos distinctions prend toute sa valeur ici. Il est artificiel de séparer les trois domaines de la terminologie, de la sémantique et de la syntaxe. Notre point de vue syntaxique est un retour à la lettre qui s'adosse à une prise en compte de la finalité et de la logique propres aux textes de loi.

Un survol contrastif de la typologie des sujets dans la prédication attributive (majoritaire dans les deux Codes) montre en position sujet une progression de la lexicalisation et une tendance à la simplification des outils grammaticaux. Ces deux mouvements semblent inversement proportionnels : plus la terminologie se complexifie et plus la syntaxe s'uniformise. C'est l'hypothèse que nous formulons pour l'étude de notre corpus.

Denis Jamet
Université Jean Moulin Lyon 3 – CEL, EA 1663

Le repérage des énoncés métaphoriques (anglais – français)

On a tendance trop souvent à aborder les figures de style sous un angle étroitement lexical. Pourtant il apparaît qu'en déplaçant le sens des mots les locuteurs se donnent des balises qui sont d'ordre essentiellement syntaxique. Grâce à quoi ils/elles marquent clairement les frontières entre sens propre et sens figuré. Ce qui montre à quel point la créativité du discours se trouve inscrite au cœur même de la langue. [Yaguello 1998 : 106]

Introduction

Si la métaphore a longtemps été rejetée, car considérée comme une anomalie aussi bien au niveau conceptuel qu'au niveau linguistique, les linguistes ont, depuis quelques décennies, démontré qu'elle respectait en tous points les règles de construction morpho-syntaxiques de la langue. Serait-ce alors à dire que rien ne distingue morpho-syntaxiquement parlant les énoncés littéraux et les énoncés métaphoriques ? Si la réponse à cette question demande à être modérée, il convient néanmoins de remarquer que les énoncés métaphoriques ont des propriétés qui leur permettent d'être repérés.

Pour ce faire, on verra qu'il est nécessaire de passer d'une approche lexicaliste de la métaphore – c'est-à-dire fondée sur le mot – à une approche énonciative – c'est-à-dire fondée sur l'énonciation en discours d'une prédication.

On débutera cet article en rappelant dans quelle optique l'on se situe, et ce que l'on considère comme métaphores, ce qui ne sera pas sans répercussions sur l'optique d'analyse.

Dans le cadre de cet article, on ne considérera pas la métaphore comme une forme essentiellement, originellement linguistique, mais comme un phénomène avant tout conceptuel, c'est-à-dire se situant au niveau de la pensée, rejoignant par là les positions et analyses des linguistes cognitivistes américains tels Lakoff et Johnson, Turner, Sweetser, Gibbs, etc. Ce n'est que parce que la métaphore est avant tout un processus conceptuel, que les métaphores linguistiques peuvent exister. L'on se réclame ainsi d'une vision « interactive », conceptuelle ou cognitive de la métaphore, c'est-à-dire de la vision initiée par I. A. Richards (dès les années 1930), puis théorisée par Max Black (dans les années 1960), et popularisée par Lakoff et Johnson, entre autres, depuis les années 1970.

L'on définira ainsi la métaphore : la métaphore consiste en une *mise en relation de deux domaines (source et cible) et de l'interaction de ces deux domaines*. Les théories récentes vont parler d'intégration conceptuelle (*blending*), mais on ne le mentionnera pas dans ce travail ; cette théorie ne s'applique pas uniquement au processus métaphorique, mais le dépasse largement, et est censée rendre compte de l'ensemble des productions langagières. Plus que « métaphore », on utilisera souvent le terme de « processus de métaphorisation », afin d'insister sur le passage pensée – langue. Selon les linguistes cognitivistes, de nombreuses notions, essentiellement les notions abstraites, ne peuvent être verbalisées, et donc connues, sans avoir recours à la métaphore. Cette dernière joue donc un rôle cognitif, car elle seule permet de fournir une nouvelle vision, une nouvelle perception¹, en créant un nouveau lien entre deux domaines conceptuels assez éloignés, assez différents, c'est-à-dire sans lien apparent au moins au premier abord, avant que la métaphore ne soit venue les relier. Le but de la métaphore sera donc principalement la

¹ Voir à ce sujet D. Jamet (2008a).

connaissance d'un domaine cible par analogie avec un domaine source. Elle permet ainsi de saisir un domaine entier (le domaine cible) par analogie avec un domaine source, comme le montre l'exemple suivant :

- (1) **INT65** - She was a girl then, looking for someone to **take the pressure off**. She had run away from home when her mother's boyfriend smashed through the glass in the front door with his hands, and she was forced to hide in a cupboard. She was an unhappy and changeable girl, who often lost herself into inexplicable moods.

Ici, la métaphore linguistique est générée sur le modèle de la métaphore conceptuelle THE BODY IS A CONTAINER FOR THE EMOTIONS ; lorsqu'un récipient se remplit sans pouvoir se vider, la pression monte, d'où l'analogie possible sur laquelle se base cette métaphore, qui saisit le domaine cible /EMOTIONS IN THE BODY/ par le domaine source /PRESSURE IN A CONTAINER/.

A titre d'exemples, on citera les métaphores linguistiques que permettent les deux métaphores conceptuelles suivantes :

- (2) LOVE IS A JOURNEY permet entre autres les métaphores linguistiques suivantes² : *we'll have to go our separate ways, we're at a crossroads, we can't turn back now, our relationship is going nowhere, we're stuck, this relationship is a dead-end street, it's a long, bumpy road, our marriage is on the rocks, we've gotten off the track, we're just spinning our wheels, look how far we've come, etc.*
- (3) IDEAS ARE FOOD permet entre autres les métaphores linguistiques suivantes : *I just can't swallow that claim, that's food for thought, let me stew over that for a while, she devoured the book, I can't digest all these ideas, all this paper has in it are raw facts, half-baked ideas, and warmed-over theories, let me chew over the proposal, let's that idea simmer on the back burner for a while, he thrives on stuff like that, etc.*

En se basant sur deux domaines existants mis en relation, la métaphore ne crée pas tant du nouveau qu'une nouvelle perception du réel (voir son rôle dans les découvertes scientifiques, comme l'exemple des métaphores d'Internet³). La métaphore est donc une sorte de *réinterprétation*, et plus particulièrement de « réinterprétation

² Les exemples sont empruntés à G. Lakoff et M. Johnson (1980) et Z. Kövecses (2002).

³ Voir D. Jamet (2008b).

véridictionnelle, celle de la fonction référentielle et celle de la vérité », selon les termes de Robert Martin (1983 [1992] : 268).

La mise en relation de deux domaines nécessite la mise en relation par un individu (l'énonciateur dans notre cas). Cette mise en relation demeure assez subjective, car c'est l'énonciateur qui a perçu une analogie entre un domaine X et un domaine Y – contrairement à la métonymie et à la synecdoque qui se basent sur des contiguïtés référentielles. Cette mise en relation passe toujours par un phénomène d'identification – au sens large – car on identifie une partie des propriétés d'un domaine X avec une partie des propriétés d'un domaine Y, propriétés que l'on a estimé être partagées par ces deux domaines, selon le phénomène traditionnellement nommé « analogie ».

Par « métaphore », on entendra donc le « processus de superposition implicite de domaines notionnels différents », à l'instar de la définition proposée par G. Lakoff et M. Johnson (1980 : 5) :

The essence of metaphor is understanding and experiencing one kind of thing in terms of another. [...] The metaphor involves understanding one domain of experience [...] in terms of a very different domain of experience [...]. The metaphor can be understood as a mapping (in the mathematical sense) from a source domain [...] to a target domain.

Maintenant, si l'on accepte l'hypothèse qu'il y a prédication dans tout énoncé en vertu du principe selon lequel il n'y a pas d'énoncé sans prédication, et que la prédication fonde donc l'énoncé, l'on peut alors se demander si le type de prédication métaphorique est identique à celui à l'œuvre lors de l'énonciation littérale, et si la distribution des modalités syntaxiques est à même de couvrir tous les types d'énoncés. Pour tenter de répondre à cette question, on adoptera une analyse qui se veut cognitive, mais également résolument énonciative.

Après un court rappel historique, on montrera que l'on est en droit de parler de « prédication métaphorique », et que la métaphore est non seulement un objet d'étude pour la stylistique, mais également pour la linguistique. Reprenant les conclusions auxquelles on était arrivé dans un travail précédent (Jamet 2009a) sur la prédication métaphorique des énoncés métaphoriques prototypiques du type X BE/ETRE Y en anglais et en français, la seconde partie élargira le corpus à d'autres types d'énoncés métaphoriques afin de dégager les conditions

morphosyntaxiques nécessaires à la prédication métaphorique. Les conclusions de ce travail montreront que l'on est en droit de parler d'un « repérage des énoncés métaphoriques ».

I. Peut-on parler de prédication dans un énoncé métaphorique ?

I.1. De la lexie à la prédication

On sait que la prédication métaphorique est généralement qualifiée de « prédication contradictoire », à l'instar de Weinreich (1976 : 308). Cela signifie-t-il pour autant que les énoncés métaphoriques sont anomaux syntaxiquement, morphologiquement, sémantiquement et logiquement parlant ?

La vision classique du processus de métaphorisation remonte à Platon, et considère que la métaphore peut être remplacée sans changement de sens par une dénomination littérale, d'où son appellation de « théorie ou modèle par substitution »⁴. On parle alors de « tropologie » pour l'étude de la métaphore, cette dernière étant réduite à une étude lexicale du mot et, comme le mentionne J.-H. Boisset (1979 : 79), il y a « hégémonie du mot », mot qui devient le « centre de la signification ». Il s'agit d'une théorie « lexicaliste » de la métaphore, le processus métaphorique ne reposant que sur un seul terme sans affecter l'ensemble de la prédication⁵.

Le corollaire de ce modèle par substitution est le principe d'« anomalie » (voir Katz, Fodor), de « déviance » (voir Weinreich) dans les énoncés métaphoriques qui sont considérés comme a-grammaticaux, « déviants » ou « anomaux ». Cependant, comme l'ont montré certains linguistes comme M. Beardsley (1978) et M. Reddy (1969), la plupart des énoncés métaphoriques ne violent aucune règle de sélection si on les replace dans le cadre du segment métaphorique

⁴ On retrouve cette conception chez Dumarsais (1730) ou Fontanier (1830).

⁵ Si cette théorie, à laquelle s'oppose vivement P. Ricoeur, est à la rigueur acceptable pour les métaphores mortes et les métaphores lexicalisées, elle ne fonctionne pas pour les métaphores vives. Et que dire des catachrèses, pour lesquelles aucun terme littéral n'existe, puisque le terme généré métaphoriquement n'existe que parce qu'il y a une absence de terme littéral pour désigner cette réalité extralinguistique ? De plus, il n'y a aucun embellissement du discours dans le cas des catachrèses.

dans lequel ils sont inclus. Ainsi, peut-on affirmer que non seulement les énoncés métaphoriques ne sont pas anomaux syntaxiquement, mais encore, comme l'écrit I. Loewenberg (1973 : 315) que :

It is an obvious fact that sentences in which metaphorical expressions occur are not distinguishable by syntactic form.

Les énoncés métaphoriques n'exhiberaient alors aucune différence qui permettrait de les identifier, voire de les différencier des énoncés littéraux, comme les approches contemporaines se plaisent à le répéter. Cependant, on estime que cette question n'est pas tranchée, et qu'il existe des propriétés linguistiques nécessaires lors de l'énonciation métaphorique. Avec la popularité des travaux plus ou moins contemporains sur la métaphore, celle-ci a fini par acquérir – essentiellement grâce aux travaux de la linguistique cognitive américaine et à l'ouvrage *La métaphore vive* de P. Ricœur (1975) – une place de choix dans la prédication. Replacé dans le cadre d'une théorie interactive, c'est-à-dire prédicative de la métaphore – ou plutôt de l'énoncé métaphorique – le foyer métaphorique va avoir une incidence sur toute la prédication, et P. Ricœur (1975 : 79) ira jusqu'à parler du « caractère quasi prédicatif de la métaphore ». Ce caractère prédicatif de l'énonciation métaphorique est également souligné par C. Franquart-Declercq et M.-D. Gineste (2001 : 746) lorsqu'elles définissent la métaphore comme « un acte de prédication ». On est alors en droit de parler de « prédication métaphorique », et on définira cette dernière comme le fait R. Jongen (1980 : 89-90) :

L'énoncé métaphorique est une affirmation implicite d'identité de propriétés par le biais d'une affirmation d'identité des supports de celles-ci. [...] En d'autres mots, la prédication métaphorique identifie p une propriété implicite d'un terme a avec une propriété implicite d'un terme b.

On s'éloigne ainsi d'une théorie lexicaliste de la métaphore, ainsi que d'une trop simpliste dichotomie entre « sens propre » et « sens métaphorique », et on estimera que l'énonciation métaphorique, tout comme l'énonciation littérale, est définie comme un acte de prédication. Néanmoins, il est nécessaire de s'interroger sur la construction prédicative métaphorique en termes d'opérations linguistiques, afin de voir si lors du passage langue – discours, des traces ne subsistent pas dans les énoncés.

I.2. Du transfert conceptuel à la mise en relation linguistique : la notion de « supplément de structure », de « prédication dépliée »

Michele Prandi (1999 : 188) définit l'opération d'énonciation métaphorique en parlant de « conflit conceptuel » :

Envisagée dans sa réalité grammaticale, la métaphore se présente comme l'articulation linguistique d'un conflit conceptuel dans une forme spécifique en fonction de sa localisation dans la structure grammaticale et fonctionnelle de l'énoncé.

Le même M. Prandi (1999 : 197) déclare que :

Si l'articulation linguistique des contenus incohérents présente des caractères exclusifs, ces caractères résident non pas dans un défaut de structure, mais dans *la présence d'un supplément de structure* : [...] la description est appelée à expliciter le surplus d'articulation qui caractérise la mise en forme d'un conflit conceptuel.⁶

Cette notion de « supplément de structure » paraît primordiale, et elle va guider l'intégralité de l'analyse de cet article, lorsque l'on va tenter de mettre au jour certaines structures aptes à véhiculer le sens métaphorique, et d'expliquer ce en quoi elles permettent à l'énonciation métaphorique d'avoir lieu. M. Prandi reste cependant au niveau qu'il nomme sémantique, c'est-à-dire conceptuel, car il ne dégage nullement les structures linguistiques qui constitueraient ce « supplément de structure ». Notons que ce supplément de structure, cette complexification linguistique qui va apparaître sous la forme d'une prédication seconde, est obligatoire dans certains cas si l'on veut que l'énoncé revête une interprétation métaphorique, alors qu'elle ne sera que facultative dans d'autres, servant seulement à orienter le coénonciateur vers l'interprétation métaphorique. Le fait que le supplément structurel soit nécessaire ou facultatif dépendra du degré de figement de la métaphore, comme on le verra plus bas. Ce supplément de structure va servir à éviter l'ambiguïté en spécifiant le domaine cible par lequel le domaine source va être conceptualisé. Il semble ainsi raisonnable de faire l'hypothèse suivante : il existe un lien de réciprocité entre les structures conceptuelles et les structures linguistiques dans le cadre de l'énonciation métaphorique ; en d'autres termes, on postule que l'opération conceptuelle métaphorique laisse

⁶ Nous soulignons.

des traces discursives lors de la verbalisation, de l'énonciation. En effet, la métaphore exhibe un codage linguistique qui, s'il n'est pas différent de celui du langage littéral, possède néanmoins certaines caractéristiques. Comme le rappelle M. Prandi (1999 : 187) :

Dans la construction de la métaphore, il n'y a pas de transgression grammaticale. Loin d'être enfreinte, la structure grammaticale est au contraire valorisée dans son pouvoir de mise en forme. [...]

La réalité formelle de la métaphore renvoie à des facteurs essentiellement systématiques, grammaticaux et conceptuels.

C'est pour cette raison qu'à la question de M. Black (1962 : 25), « Are there any criteria for the detection of metaphors? », on répondra par l'affirmative, en gardant néanmoins à l'esprit que ces critères ne sont pas restreints aux seuls énoncés métaphoriques, mais apparaissent également dans les énoncés littéraux, même s'il nous semble indéniable que l'énonciation métaphorique privilégie certaines structures lexicales, morphologiques, syntaxiques, phonologiques (sous forme typographique), etc. Il existe donc des structures morphosyntaxiques plus aptes que d'autres à véhiculer la lecture métaphorique, et ce sont donc les traces, les marqueurs, les structures véhiculant le sens métaphorique que l'on va tenter de dégager, cette « logique » du langage que l'on va dévoiler dans le cadre de la prédication métaphorique.

Pour résumer, on peut dire que la métaphore n'utilise pas des structures morfo-syntaxico-lexicales déviantes, mais des structures langagières également à l'œuvre dans le langage littéral. Cependant, si les structures linguistiques utilisées par la métaphore ne lui sont pas spécifiques, un énoncé métaphorique exhibe des structures lexico-morfo-syntaxiques privilégiées, et il y a donc, selon les termes de J. Tamine (1979 : 78) une « localisation des métaphores » dans des cadres privilégiés. On montrera pourquoi ce sont ces structures – aux dépens d'autres structures possibles – qui sont à même de véhiculer le sens métaphorique, en se concentrant dans le cadre de cet article sur les traces morfo-syntaxiques, laissant volontairement de côté les traces lexicales, lexico-sémantiques, phonologiques et typographiques propices à l'énonciation métaphorique. L'on partira donc du postulat que l'énoncé de surface, l'énoncé linéaire, permet non seulement de construire du sens – de par la mise en relation des lexèmes et des opérateurs – mais encore de renseigner sur la structure linguistique

qu'il reflète. Il convient alors de se pencher minutieusement sur la structure linguistique si l'on postule qu'elle conserve des traces des opérations énonciatives métaphoriques qui l'ont permise.

II. Structures privilégiées de l'énonciation métaphorique

Dans Jamet (2009a), on avait montré que les énoncés métaphoriques en BE / ETRE partageaient une propriété remarquable, à savoir le fait que le foyer métaphorique (sous la forme de l'attribut du sujet) exhibait un supplément de structure qui pouvait prendre plusieurs formes : un syntagme prépositionnel jouant le rôle de postmodification / postqualification, un adjectif qualificatif, une proposition subordonnée relative, la répétition de la métaphore (à l'identique, par réinterprétation, ou par apposition littérale). Les conclusions de ce travail ont montré que les cas les plus fréquents exhibaient une combinaison de ces divers suppléments de structure, correspondant à des prédications secondes. Le travail effectué sur les énoncés métaphoriques en X BE / ETRE Y a montré que le Y (le foyer métaphorique) dans les métaphores vives n'apparaissait que très rarement sous la forme d'un SN simple, non-construit, du type dét.⁷ + nom. Il a besoin d'être complexifié par une prédication seconde afin que l'énoncé reçoive une interprétation métaphorique. C'est ainsi que dans le cadre de l'énonciation métaphorique, on est en droit de parler non pas de « prédication repliée », mais de « prédication dépliée », qui va servir à expliciter l'interprétation métaphorique. Par « prédication dépliée », on entend le fait que la prédication de base du type SVO n'est pas suffisante, mais a besoin d'être complexifiée, ce que l'on nomme « prédication dépliée ».

En partant des conclusions auxquelles on avait abouti dans Jamet (2009a), on proposera de nouvelles occurrences, et on vérifiera si cette notion de « prédication dépliée » a une quelconque validité pour l'analyse de l'énonciation métaphorique. La plupart des énoncés métaphoriques sélectionnés obéissent à un schéma morpho-syntaxique de type X BE / ETRE Y, même si pour les besoins de l'analyse, d'autres configurations grammaticales ont également été proposées.

⁷ Déterminant autre que les déterminants en TH- en anglais, et LE, LA, LES, CE, CET, CETTE et CES en français.

II.1. Prédication dépliée sous la forme d'un syntagme prépositionnel

On a remarqué qu'une postmodification sous la forme d'un syntagme prépositionnel, souligné dans les exemples ci-dessous, permettait la lecture métaphorique. Sans cet ajout structurel, l'énoncé soit devient incohérent, soit change de sens en recouvrant une acception littérale. Ceci confirme une idée fondamentale et bien développée⁸ : ce n'est pas seulement un mot ou un groupe de mots qui est métaphorique, même si ces derniers sont porteurs du foyer métaphorique, mais l'énoncé dans son intégralité. C'est donc la prédication qui est en jeu lors de l'énonciation métaphorique. Ainsi, de leur statut traditionnel de circonstants à l'intérieur de l'énoncé – c'est-à-dire d'éléments optionnels à la construction du sens « fondamental », ces syntagmes prépositionnels vont acquérir un statut d'arguments, en ce qu'ils deviennent nécessaires à la construction et à l'interprétation du sens. Les expansions qu'ils constituent sont donc nécessaires au décodage métaphorique de l'énoncé, et pas seulement du foyer métaphorique, en permettant de déplier la prédication. On débutera par des occurrences où c'est un substantif qui est postmodifié par un syntagme prépositionnel :

- (4) **INT16c** - My young gay friend Ian liked to stand with me outside tube stations where I would watch the **flocks of girls** in the summer, after I had finished work for the day, around lunch time.
- (5) **INT18d** - It was the late seventies, and relationships were nonchalant and easy, as if it had been agreed that the **confinement of regularity** made people mentally sick.
- (6) **INT45c** - All effort was a trouble and she lived on the **edge of panic**, as if everything was about to break down.
- (7) **DIA13b** - Il est rare qu'un cataclysme se produise sans phénomènes avant-coureurs. L'attentat autrichien, l'**orage du procès Caillaux** répandaient une atmosphère irrespirable, propice à l'extravagance.
- (8) **DIA61** - Un **parfum de provisoire** excitait mes sens. D'avoir goûté à des joies plus brutales, plus ressemblantes à celles qu'on éprouve sans amour avec la première venue, affadissait les autres.

⁸ Voir par exemple I. Tamba-Mecz (1981).

- (9) **DIA49b** - Dans toutes ces petites villes de banlieue ouvrière, sévissaient les mêmes passions, la même **soif de raconter** qu'en province.

On trouve également des exemples de métaphores verbales, rendues possibles là aussi par l'ajout d'une prédication dépliée sous la forme d'un syntagme prépositionnel :

- (10) **INT6b** - He is used to the way I **dash from subject to subject**, following the natural momentum of my mind.
- (11) **INT27** - Fear is something I recognize. My childhood still **tastes of fear**, of hours, days and months of fear.
- (12) **INT11a** - This confusion is not going to leave me alone. But by the morning my mind had better be made up about certain things. I must not **descend into self-pity**, at least not for longer than necessary. I have found that it is not my moods that frustrate me but the depth and indeterminacy of their duration. If I feel a bit low, I fear a year-long depression.
- (13) **DIA11a** - A une heure, j'avais supplié le directeur de ne rien dire à mon père ; à quatre, je **brûlais de lui raconter tout**.

Entre le fonctionnement nominal et le fonctionnement verbal, on va également trouver tous les cas de fonctionnements adjectivaux et adverbiaux, aussi bien en anglais qu'en français :

- (14) **INT18a** - Now I resent being **bombarded by vulgarity, emptiness and repetition**.
- (15) **DIA25b** - Nous restions immobiles. Ainsi doit être le bonheur. Un chien sauta du bassin et se secoua. Marthe se leva, comme quelqu'un qui, après la sieste, et le visage encore **enduit de sommeil**, secoue ses rêves.
- (16) **DIA60a** - Je n'avais jamais tel désir d'embrasser Marthe que lorsqu'un travail la distrayait de moi ; jamais tant envie de toucher ses cheveux, de la décoiffer, que quand elle se coiffait. Dans le canot, je me précipitais sur elle, la **jonchant de baisers**, pour qu'elle lâchât ses rames, et que le canot dérivât, prisonnier des herbes, des nénuphars blancs et jaunes.

Dans tous ces exemples, le supplément de structure sous la forme du syntagme prépositionnel joue le rôle de prédication dépliée obligatoire, et permet à l'énonciateur de frayer la lecture métaphorique pour le co-énonciateur. Cependant, certains énoncés conservent un sens si l'on enlève le syntagme prépositionnel, mais celui-ci est

différent, et il convient de se demander pourquoi il reste cependant interprétable, malgré l'absence de ce dernier :

- (17) **INT4c** - This morning, before I set out for the day, knowing I had to **settle** a few things in my mind, the elder boy, insisting on another kiss before I closed the door, said, 'Daddy, I love everyone'.
- (18) **INT93** - Nina used to tease me, saying my attitude towards drugs belonged in another era. It is true that when I was growing up, drugs were **fuel** for a journey into the self.
- (19) **EMM18** - On aurait dû vous prévenir, s'entêta Pierrot. Ma mère, elle est connue comme concierge ! C'est une **star** dans sa profession !

Tous les énoncés ci-dessus pourraient plus ou moins être acceptables sans la prédication dépliée que constitue le syntagme prépositionnel, mais il convient de remarquer que ce dernier apporte néanmoins une précision sémantique.

La relative acceptabilité des énoncés sans prédication dépliée peut s'expliquer par le fait que l'on est face à des métaphores lexicalisées. En effet, dès que la métaphore se lexicalise, c'est-à-dire qu'elle commence à être acceptée par les locuteurs natifs comme faisant partie intégrante de leur langue, ces derniers perdent peu à peu conscience de l'origine et de la motivation métaphorique de celle-ci. L'énoncé ne nécessite plus de supplément de structure, c'est-à-dire de prédication dépliée, comme dans les deux occurrences suivantes, où un locuteur natif ne ressent aucunement l'origine métaphorique :

- (20) **INT44** - What did Father's life show me? That life is a **struggle**, and that struggle gets you nowhere and is neither recognized nor rewarded.
- (21) **INT53b** - Then we went on the dole for five years in order to pursue our self-righteous politics, before starting to work in the media and making a lot of money. We weren't much restrained by morality or religion. Music, dancing and conscienceless fucking were our **totems**.

II.2. Prédication dépliée sous la forme d'un adjectif qualificatif

La prédication dépliée peut également être réalisée sous la forme d'une pré- ou post-modification par un adjectif qualificatif.

- (22) **INT10a** - I have to be aware of everything. I want to develop a mental picture I can carry around and refer to when I am at Victor's place.

Sans la prédication dépliée sous la forme de l'adjectif *mental*, l'énoncé retrouve un sens littéral, et l'énonciateur développerait une image qu'il porterait avec lui... Les occurrences en anglais (23) et (24), ainsi que les occurrences en français (25), (26) et (27) exhibent le même phénomène :

- (23) **INT80c** - I push down my trousers and look for a suitable lubricant. The last time I did this, when Susan had some friends round for dinner, I used my children's shampoo, and felt as if a wasp had been pushed into my urethra. I should have complained to the manufacturer and had them conduct an experiment on animals. Even masturbation can be a medical minefield.
- (24) **SHU187b** - He starts to see the publication of his memoirs in a quite different light.
- (25) **DIA32** - La joie de donner enfin satisfaction à mon père comblait un peu le **vide sentimental** dans lequel je me trouvais, car, si je croyais ne plus aimer Marthe, je la considérais du moins comme le seul amour qui eût été digne de moi.
- (26) **DIA80b** - Après avoir conservé à notre maison une **façade digne**, il perdait toute retenue, et, quand je n'étais pas rentré depuis plusieurs jours, envoyait la femme de chambre chez Marthe, avec un mot à mon adresse, m'ordonnant de rentrer d'urgence ; sinon il déclarerait ma fuite à la préfecture de police et poursuivrait Mme L. pour détournement de mineur.

Ce dernier exemple est intéressant en ce qu'il exhibe un jeu sur le sens littéral – sens figuré (« maison » + « façade »), mais également un phénomène de métonymie (« maison », le tout pour la partie).

- (27) **SIL43** - Ceux qu'il commande obéissent à la crainte et non plus à l'amour. Un chef qui n'a pas l'amour des siens est un bien misérable mannequin.

Si l'on revient sur ce dernier exemple, on remarque que l'absence du groupe adjectival « bien misérable » change complètement le sens : d'un sens positif (voir les dénotations et les connotations attachées au substantif « mannequin »), le chef se voit alors retirer tout pouvoir, mais n'est là que pour figurer le rôle de chef, sans n'en avoir aucun des attributs : pouvoir, respect, force, etc. C'est pour cette raison que l'adjectif joue également le rôle de prédication dépliée, en « contaminant » tous les membres de l'énoncé.

II.3. Prédication dépliée sous la forme d'une subordonnée relative (déterminative)

- (28) **BOD60a** - If these gratifications were intoxicating and almost mystical in their intensity, I learned, when something stranger happened, that indulgence wasn't a full-time job and reality was **a shore** where dreams broke.
- (29) **DIA12b** - Ainsi, mon travail de chaque jour était-il peu de chose, mais, comme, travaillant moins de temps que les autres, je travaillais en plus pendant les vacances, ce peu de chose était **le bouchon de liège** qu'un chat garde toute sa vie au bout de la queue, alors qu'il préférerait sans doute un mois de casserole.
- (30) **DIA18b** - Pour la première fois, nul jour des vacances de cette année ne me fut pesant. Je connus donc que personne n'échappe à son âge, et que mon dangereux mépris s'était fondu comme glace dès que quelqu'un avait bien voulu prendre garde à moi, de la façon qui me convenait. Nos communes avances raccourcirent de moitié **la route** que l'orgueil de chacun de nous avait à faire.
- (31) **ATT13a** - Des grands bois voisins, descendaient de **longues haleines** qui passaient sur les toitures comme des caresses.

Dans ces quatre exemples, la suppression de la relative – qui joue un rôle très proche de celui de l'adjectif – entraîne une irrecevabilité des énoncés, car le coénonciateur ne peut interpréter l'énoncé métaphoriquement sans la prédication dépliée que constitue la subordonnée relative. Il faut noter que seules les propositions relatives restrictives ou déterminatives permettent une interprétation métaphorique, car elles seules – contrairement aux relatives appositives ou explicatives – restreignent la portée sémantique du nom auquel elles se rapportent. Et qu'est-ce que la métaphore, si ce n'est

une restriction dans le programme sémantique du terme métaphorisé – au moins selon l’approche lexicaliste traditionnelle ? Les relatives dans les exemples ci-dessus ne peuvent être enlevées, car elles sont nécessaires à la référentiation de l’antécédent métaphorique. Là aussi, il y a une prédication seconde complètement dépliée en ce que la structure Sujet-Prédicat apparaît de façon explicite.

II.4. Prédication dépliée sous la forme de répétition de métaphores ou d’apposition littérale

Avant d’aborder les cas où les divers types de complexifications structurelles se combinent, on souhaite examiner les cas où l’interprétation métaphorique est rendue possible tout d’abord par une apposition littérale (33), mais aussi par une répétition de métaphores, cas traditionnellement connu sous le nom de « métaphore filée » (voir exemples suivants) :

- (32) **INT80a** - How weak the arc of my urine is, and how I strain to send a respectable semicircle into the pan. Even when my boys were tiny, and those round little **worms**, their penises, were no thicker than a cable, the arc of their urine had a magnificent velocity.
- (33) **INT38** - But to be able to bear one’s own mind, to wait while the inner **storm** of intolerable thoughts **blows itself out**, leaving one to contemplate the **debris** with some understanding – that is an enviable state of mind.

Cet énoncé (33) se compose de trois métaphores : la première est rendue possible par plusieurs procédés qui se combinent, et la seconde métaphore est rendue possible par la première. Il y a donc d’une part une raison lexicale – à savoir la reprise d’une partie du champ sémantique qui permet de créer la métaphore, ici *storm* → *blow out* – mais également une autre reprise lexicale des deux premières métaphores qui permet la troisième, *debris*, c’est-à-dire ce qu’il reste après une tempête qui a été très agitée à cause d’un vent violent.

- (34) **INT46-47** - My parents sat in separate armchairs, watching me carry out my records. What was left for them to do? Hadn’t I rendered them irrelevant? [...]

Victor says that once the **lights** on a love have **dimmed**, you can never **illuminate** them again, any more than you can reheat a soufflé. But my parents **went through the darkness** and discovered a new intimacy.

Plus encore que dans l'occurrence (33), en (34), on peut dire que la métaphore est filée sur le thème de LOVE IS LIGHT. Toutes les métaphores se fondent sur cette métaphore conceptuelle, et il s'agit donc ici plus d'une justification lexicale que d'une justification morpho-syntaxique.

- (35) **INT34b** - People don't want you to have too much pleasure; they think it's bad for you. You might start wanting it all the time. How **unsettling** is desire! That **devil** never sleeps or keeps still. Desire is naughty and doesn't conform to our ideals, which is why we have such a need of them. Desire **mocks** all human endeavour and makes it worthwhile. Desire is the original **anarchist** and **undercover agent** – no wonder people want it **arrested** and **kept in a safe place**. And just when we think we've got desire under control it **lets us down** or **fills us** with hope. Desire makes me laugh because it makes fools of us all. Still, rather a fool than a fascist.

La métaphore est ici filée sur le thème du désir. On notera que la métaphore *arrested and kept in a safe place* qui joue sur le sens propre – sens figuré, n'est rendue possible que par la métaphore précédente, *Desire is the original anarchist and undercover agent*.

- (36) **DIA89a** - Ma mère remarqua que j'avais les yeux rouges. Mes sœurs rirent parce que je laissais deux fois de suite retomber ma cuillère à soupe. Le plancher **chavirait**. Je **n'avais pas le pied marin** pour la souffrance. Du reste, je ne crois pouvoir comparer mieux qu'au mal de mer ces **vertiges** du cœur et de l'âme. La vie sans Marthe, c'était **une longue traversée**. Arriverais-je ?

On se trouve dans cet exemple face à une métaphore filée marine, chaque métaphore s'appuyant sur la précédente, et formant ainsi une sorte de résonance métaphorique à l'intérieur du texte.

II.5. Combinaison de divers types de prédictions dépliées

On terminera en notant un phénomène fréquent, à savoir la combinatoire de ces diverses prédictions dépliées, sans oublier d'autres types de marquage métaphorique, non abordées dans cet

article (voir Jamet, 2009b). Il apparaît que plus la métaphore est vive, c'est-à-dire surprenante, incongrue, plus elle nécessite un supplément structurel pour pouvoir être interprétée adéquatement par le coénonciateur. Les métaphores dans les occurrences suivantes étant quasiment toutes des métaphores vives, elles sont peu recevables, voire irrecevables sans les divers suppléments de structure réalisés sous forme de prédictions dépliées. On se contentera de souligner les suppléments structurels qui permettent la lecture métaphorique :

- (37) **INT52a** - How little directness there is, when you look around! We have to make things distinct by indirection. What a redundant and fearful dance it all is, as if our feelings are weapons that could kill, and words are their **bullets**.
- (38) **INT82-83** - Susan can be a virulent combatant, with a sharp tongue I've enjoyed. Unfortunately, her bitterness is too urgent to be witty; she lacks detachment.
- (39) **INT95** - I liked taking Nina to restaurants and parties, to openings and exhibitions. [...] Wherever we were, she was my refuge, my pocket of light.
- (40) **EMM145** - Oui, il est méchant. Il a peur qu'on lise les horreurs qu'il écrit en secret, alors il tend des pièges. Elle eut conscience qu'elle parlait trop fort, avec une véhémence proche de l'hystérie.
- Non, coupa le scaphandrier. Le journal, c'est une soupape imaginaire, c'est tout. Un petit théâtre personnel où il se venge de ses humiliations quotidiennes, de sa vie de fils de concierge.
- (41) **EMM138** - Il l'avait espionnée jusque dans ses ablutions, il l'avait épiée des heures durant pendant qu'elle se lavait, s'habillait. C'était une fouine, un petit animal nuisible et puant, de la graine de détraqué, de...

Même si la métaphore « fouine » est lexicalisée, l'auteur la complète par un supplément de structure en dépliant la prédication, car le sens qu'il désire lui donner n'est pas exactement le sens répertorié dans les dictionnaires, comme dans le *Larousse en 22 volumes* : « *Fig. et Fam.* : personne curieuse, indiscrete, maligne, rusée ». Il semble qu'il veuille rapprocher la métaphore lexicalisée de son sens premier (animal).

- (42) **DIA66c** - L'instinct est notre **guide** ; un **guide** qui nous conduit à notre perte.

- (43) **DIA12b** - Ainsi, mon travail de chaque jour était-il peu de chose, mais, comme, travaillant moins de temps que les autres, je travaillais en plus pendant les vacances, ce peu de chose était le bouchon de liège qu'un chat garde toute sa vie au bout de la queue, alors qu'il préférerait sans doute un mois de casserole.

Partant de l'hypothèse qu'il existe un supplément de structure sous la forme d'une prédication dépliée à l'œuvre lors de l'énonciation métaphorique, on a vu que cette prédication dépliée est plus ou moins nécessaire selon les occurrences examinées. Dans certains cas, elle apparaît même comme facultative, et certains exemples semblent ainsi infirmer l'hypothèse initiale :

- (44) **BOD3a** - Though the late 1950s and early 1960s were supposed to be my **heyday**, I don't like the assault of loud music, and I have come to appreciate silence in its many varieties.

Aucune complexification sous la forme d'une prédication dépliée n'est présente ici, car il n'y a ni groupe prépositionnel, ni adjectif, ni relative, etc. Cependant, l'hypothèse n'est en rien infirmée, car dans les cas de métaphores lexicalisées, c'est-à-dire celles dont le locuteur natif n'a plus ou guère souvenir de l'origine métaphorique, rien ne différencie ces prédications d'énoncés métaphoriques des prédications des énoncés dits littéraux.

- (45) **BOD162-163** - Mother had always watched television from the late afternoon until she went to bed. She hadn't wanted Harry or his brother or father to speak. If they opened their mouths, she told them to shut up. She hadn't wanted them in the room at all. She preferred the faces on television to the faces of her family. She was an **addict**.

En (45), on est face à une métaphore lexicalisée, ce que prouve son insertion dans un dictionnaire comme le *Hachette-Oxford*, qui parle de *telly addict*, comme c'est ici le cas, et comme le confirme le cotexte gauche.

- (46) **SIL27a** - L'hiver en France est une douce saison. Chez moi c'est bien dur. Très. Les arbres sont des sapins, des forêts serrées, la neige est lourde là-dessus. Ici les arbres sont fins. La neige dessus c'est une **dentelle**.

On note le même phénomène en (46), où la métaphore est lexicalisée comme le note le *TLFi*⁹ : « *Au Fig.* Tout ce qui rappelle la dentelle par sa finesse, sa délicatesse ».

Conclusion

Cet article a eu pour but de montrer que le repérage des énoncés métaphoriques n'est pas impossible. Contrairement à la vision traditionnelle lexicaliste, qui tient la métaphore pour un phénomène purement sémantique, portant sur le seul mot, la métaphore se développe avant tout dans un énoncé, et si sa raison d'être est conceptuelle et qu'il s'opère un changement sémantique, ce sont certaines structures morpho-syntaxiques privilégiées qui permettent essentiellement au coénonciateur d'interpréter un énoncé comme métaphorique.

En replaçant la métaphore dans le cadre d'une analyse énonciative et cognitive, on se rend compte qu'elle ne saurait être qualifiée de prédication contradictoire, ano(r)male ou déviante. La prédication métaphorique n'est rien d'autre qu'une prédication littérale, qui a besoin d'être dépliée pour assurer son décodage discursif si elle est vive. Alors, si l'on considère que l'opération de prédication fonde l'énoncé, les énoncés métaphoriques seraient plus aptes à correspondre à cette opération en ce qu'ils exhibent une prédication dépliée, et ne peuvent alors plus être qualifiés de « prédictions contradictoires », car comment expliquer alors les cas de syllepse, où le sens littéral et le sens métaphorique cohabitent pour générer un sens encore plus riche, car polyphonique ?

⁹ *Trésor de la Langue Française informatisé* : <http://atilf.atilf.fr/tlf.htm>

Bibliographie

- Beardsley M. C., 1978, « Metaphorical Senses », *Noûs* 12.1 : 3-16.
- Benveniste E., 1966 (1992), *Problèmes de linguistique générale 1*, Paris, 'Tel', Gallimard.
- , 1974 (1992) *Problèmes de linguistique générale 2*, Paris, 'Tel', Gallimard.
- Boisset J.-H., 1979, « In Defense of the Predicative Function of Metaphor », in W. C. McCormack & H. J. Izzo (Eds.), *The Sixth LACUS Forum*, Hornbeam Press, Inc., p. 78-85.
- Cohen J., 1966, *Structure du langage poétique*, Paris, Flammarion.
- Delmas C., 1987, *Structuration abstraite de la chaîne linéaire en anglais contemporain*, Paris, 'linguistique', Cedel.
- Franquart-Declercq C. & Gineste M.-D., 2001, « L'enfant et la métaphore », *L'année psychologique*, 101^{ème} année, 4^{ème} fascicule, Paris, PUF.
- Jamet D., 2002, *Contributions cognitive et énonciative au repérage des lexies métaphoriques. Domaines anglais-français*, thèse de doctorat, Université Jean Moulin – Lyon 3 et Université Sorbonne Nouvelle – Paris 3, direction Malcolm Clay et Claude Delmas.
- (Ed.), 2008a, *Métaphore et perception. Approches linguistiques, littéraires et philosophiques*, Paris, L'Harmattan.
- , 2008b, « La perception d'Internet via ses métaphores », in D. Jamet (Ed.), *Métaphore et perception. Approches linguistiques, littéraires et philosophiques*, Paris, L'Harmattan : 39-56.
- (Ed.), 2008c, *Dérives de la métaphore*, Paris, L'Harmattan.
- « Les dérives potentielles de la métaphore : essai de typologie », in D. Jamet (Ed.), *Dérives de la métaphore*, Paris, L'Harmattan, décembre 2008.
- « Que prédique-t-on dans un énoncé métaphorique ? », Actes du colloque international de linguistique organisé par l'UFR LAG-LEA et le CELA, Université de Provence – Aix-en-Provence, 4, 5 et 6 novembre 2004, *Faits de langue* 30-31, 2009a : 301-311.
- http://www.up.univ-mrs.fr/wclaux/pred/pred_denis_jamet.pdf
- *L'énonciation métaphorique*, Limoges, Editions Lambert-Lucas, 2009b.
- Jongen R., 1980, « La métaphore comme éponyme et comme prédication d'identité », in R. Jongen (Ed.), *La Métaphore ; approche pluridisciplinaire*, Université de St-Louis.

- Kleiber G., 1994, « Métaphore : le problème de la déviance », in R. Landheer (Ed.), « Les figures de rhétorique et leur actualité en linguistique », *Langue française* n°101, Paris, Larousse, p. 35-56.
- Lakoff G. & Johnson M., 1980, *Metaphors We Live By*, The University of Chicago Press, Chicago.
- Loewenberg I., 1973, « Truth and Consequences of Metaphors », *Philosophy and Rhetoric* 6.1, p. 30-46.
- Martin R., (1983) 1992, *Pour une logique du sens*, Paris, 'linguistique nouvelle', Presses Universitaires de France.
- Prandi M., 1992, *Grammaire philosophique des tropes. Mise en forme linguistique et interprétation discursive des conflits conceptuels*, Paris, 'Propositions', Editions de Minuit.
- Prandi M., 1999, « Grammaire philosophique de la métaphore », in N. Charbonnel & G. Kleiber, *La métaphore entre philosophie et rhétorique*, Paris, 'linguistique nouvelle', Presses Universitaires de France, p. 184-206.
- Reddy M. J., 1969, « A Semantic Approach to Metaphor », *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*, p. 240-251.
- Tamba-Mecz I., 1981, *Le sens figuré. Vers une théorie de l'énonciation figurative*, Paris, 'linguistique nouvelle', Presses Universitaires de France.
- Yaguello M., 1998, *Petits faits de langue*, Paris, Seuil.

Ouvrages utilisés pour le corpus :

- ATT** Zola E., (1880) 1997, *L'attaque du moulin*, Paris, Librio.
- BOD** Kureishi H., (2002) 2003, *The Body*, London, Faber & Faber.
- INT** Kureishi H., 1998, *Intimacy*, London – Boston, Faber & Faber.
- DIA** Radiguet R., (1923) 1997, *Le diable au corps*, Paris, Librio.
- EMM** Brussolo S., (1990) 1999, *Les Emmurés*, Paris, Editions de la sentinelle, Le Livre de Poche.
- SIL** Vercors, (1942) 1985, *Le silence de la mer*, *Le silence de la mer*, Paris, Le Livre de Poche.
- SHU** Swift G., *Shuttlecock*, Picador, 1981.

François Maniez
Université Lumière Lyon 2 - Centre de Recherche en
*Terminologie et Traduction*¹

Quelques tendances récentes de la linguistique de corpus

Introduction

En 1991, John Sinclair publiait *Corpus, Concordance, Collocation*. Bien que cet ouvrage ait fortement contribué à définir la linguistique de corpus, l'intérêt des linguistes pour l'utilisation de l'outil informatique afin de mieux décrire la langue était loin d'être nouveau à l'époque. Dès le début des années 70, Labov et d'autres sociolinguistes américains avaient incorporé l'informatique à leurs systèmes de description des phénomènes de variation de l'anglais parlé. De fait, la création des premiers corpus sous forme informatique remonte même aux années 60, période pendant laquelle les linguistes américains Kucera et Francis (1967) ont publié les premiers résultats de l'exploitation du corpus Brown (du nom de l'Université où ils enseignaient), qui totalisait un million de mots d'anglais américain écrit. Au même moment, de l'autre côté de l'Atlantique, les linguistes de l'Université de Lancaster, en collaboration avec des collègues scandinaves, créaient un corpus de taille similaire à partir de textes d'anglais britannique, connu sous le nom de LOB (Lancaster-Oslo-Bergen).

Même si l'emploi du terme « linguistique de corpus » sous-entend de nos jours l'emploi de corpus informatisés, il convient de noter que les linguistes n'ont pas attendu l'avènement de l'informatique pour avoir recours à l'utilisation de grandes quantités de données textuelles. Les linguistes « de terrain » comme Boas le faisaient dès les années

¹ <http://recherche.univ-lyon2.fr/crtt/> & francois.maniez@univ-lyon2.fr

40, et les études portant sur l'acquisition de la langue maternelle et la pédagogie des langues secondes ont traditionnellement été fondées sur une collecte systématique de données qui justifie l'appellation de « corpus ». Tony McEnery et Andrew Wilson² proposent le terme d'«*early corpus linguistics*» pour désigner ce type d'études, tout en précisant qu'il s'agit d'une appellation rétrospective, l'utilisation d'un corpus n'ayant été à l'époque qu'un moyen d'arriver à une description exhaustive d'un phénomène linguistique donné. Même si l'exploitation de tels corpus était nécessairement exempte de toute automatisation, leur taille pouvait parfois être considérable : en 1897, Käding avait ainsi constitué un corpus d'environ 11 millions d'occurrences afin d'étudier la fréquence de distribution des lettres et suites de lettres en allemand.

En Europe, l'intérêt pour les corpus encodés sous forme numérique, qui s'est manifesté en France dès les années 60 avec la création de la base textuelle Frantext,³ s'est généralisé à la fin des années 70 avec l'apparition de l'ICAME (International Computer Archive of Modern and Medieval English), essentiellement sous l'impulsion de linguistes britanniques et scandinaves, et le volume des publications dans ce domaine n'a cessé de croître depuis le début des années 80. La liste électronique CORPORA, créée en 1992, permet à tous les chercheurs qui s'intéressent à la linguistique de corpus d'échanger des informations et de se tenir au courant des manifestations organisées dans leur domaine.

Ce n'est toutefois qu'au début des années 80 que l'on voit émerger la linguistique de corpus en tant que branche distincte de la linguistique générale. L'une des premières mentions du terme « corpus » dans le titre d'une monographie se trouve dans l'ouvrage des grammairiens Svartvik et Quirk intitulé *A Corpus of English Conversation* (1980), mais c'est une étude publiée deux années plus tard par Johansson⁴ qui affirme clairement l'importance de l'utilisation de l'informatique dans cette branche nouvelle. En 1984, Aarts et Meijs publient *Corpus Linguistics*,⁵ qui définit encore plus précisément ce domaine de recherche.

² Mc Enery, T and A. Wilson (1996): *Corpus Linguistics* (Edinburgh: Edinburgh University Press), p. 2.

³ Frantext est un fonds de quelque 160 millions de mots constitué à l'Institut National de la Langue Française (InaLF – CNRS), qui a servi à la rédaction du *Trésor de la Langue Française*.

⁴ Johansson, S. (1982): *Computer Corpora in English Language Research*. Bergen: Norwegian Computing Centre for the Humanities.

⁵ Aarts, J. et W. Meijs (eds) 1984. *Corpus Linguistics*. Amsterdam: Rodopi.

On peut donc considérer que la linguistique de corpus est à l'heure actuelle âgée d'une trentaine d'années, si l'on restreint l'acception de ce terme à l'utilisation de corpus informatisés. Son essor est dû à la conjonction de deux facteurs : la position théorique initiale de linguistes qui ont choisi, afin de décrire la langue, de concentrer leur intérêt sur la performance plutôt que sur la compétence des locuteurs natifs d'une part, et la disponibilité croissante de données sous forme numérique d'autre part. Alors qu'il y a dix ans encore, le linguiste devait utiliser la reconnaissance optique de caractères pour constituer ses corpus, cette pratique est devenue obsolète en raison de l'augmentation du nombre de documents accessibles en ligne grâce à l'Internet, particulièrement dans le cas de l'anglais. Cette évolution, outre le fait qu'elle a considérablement raccourci la phase préparatrice du travail du linguiste de corpus, a eu pour conséquence de lui permettre d'apporter beaucoup plus d'attention à la constitution de ses corpus qu'à leur taille.⁶

Une linguistique de la performance

Si nous avons parlé de position théorique de principe, c'est parce que l'utilité de la linguistique de corpus est encore loin d'être reconnue par l'ensemble de la communauté des linguistes. Aux États-Unis, il semble qu'une guerre oppose encore deux écoles : celle des tenants de la grammaire générative d'une part, et celle des linguistes de corpus de l'autre. Les prises de position de Noam Chomsky contre la linguistique de corpus et contre l'approche empirique en linguistique de manière générale sont connues.

Rappelons brièvement l'un de ses jugements :⁷

Any natural corpus will be skewed. Some sentences won't occur because they are obvious, others because they are false, still others because they are impolite. The corpus, if natural, will be so wildly skewed that the description [based upon it] would be no more than a mere list.

⁶ En comparaison avec le million de mots des corpus mentionnés plus haut, le *British National Corpus*, désormais consultable en ligne, compte plus de 100 millions de mots (dont 10% consistent en transcriptions de production orale), et le *Cobuild's Bank of English Corpus* en dénombre plus de 400 millions.

⁷ Chomsky, N. (1957), *Syntactic Structures*, The Hague, Mouton. (p. 159)

L'approche de Chomsky souligne clairement l'importance que revêt pour lui la description exhaustive de la compétence du locuteur natif, l'utilisation du corpus étant invalidée parce que « certains types de phrases n'y seront pas représentés ». Pour le linguiste de corpus, plus intéressé par la notion de performance, le fait que certains types de phrases soient présents dans certains types de corpus est en fait un phénomène digne d'intérêt. Un autre point de divergence entre ces deux écoles est le désintérêt marqué des linguistes chomskiens pour les mesures quantitatives, dont le pendant est la méfiance proverbiale des linguistes de corpus envers l'introspection. Dans un entretien du 12 janvier 2004⁸, Chomsky se livre à une comparaison visant à démontrer l'inutilité de la linguistique de corpus :

Corpus linguistics doesn't mean anything. It's like saying [...] suppose physics and chemistry decide that instead of relying on experiments, what they're going to do is take videotapes of things happening in the world and they'll collect huge videotapes of everything that's happening and from that maybe they'll come up with some generalizations or insights.

Le linguiste de corpus est décrit comme cantonné dans un rôle d'enregistrement des données, qui selon Chomsky ne le rend pas capable d'arriver à un degré de généralisation suffisant, à une théorie linguistique globale, qu'il présuppose comme étant le seul but digne d'intérêt. Il prolonge ainsi sa comparaison entre physique et linguistique :

If you want to understand how bodies fall, Galileo would not have been interested in videotapes of leaves falling and balls going around and rocks rolling down mountains and so on and so forth. What he was interested in is the highly refined abstract conception of a ball rolling down a frictionless plane, which doesn't even exist in nature. But it turns out that the study of that is what provided insight into the principles of nature and that's the way the sciences have developed ever since.

On voit ici s'exprimer le corollaire de la méfiance de Chomsky vis-à-vis de la performance et de son caractère imparfait : la nécessité d'idéaliser les données linguistiques brutes. Mais c'est en fait tout un

⁸ Andor J. (2004) *The Master and His Performance: An Interview with Noam Chomsky*, in Kecskes, I. (ed.) *Intercultural Pragmatics*, 1-1 (2004), 93-111, Berlin: Mouton de Gruyter. (pp. 35-60)

pan de la recherche scientifique actuelle qui est rejeté comme étant de peu d'intérêt, et ce de manière sans doute un peu péremptoire, car de nombreux protocoles dépendent de l'observation de phénomènes similaires à ceux que décrit Chomsky (*videotapes of leaves falling*) pour aboutir à une modélisation.

Le linguiste de corpus est donc un linguiste de la performance, de l'étude des données brutes, alors que le linguiste « de fauteuil » se fie avant tout à son introspection, qui doit suffire à la description de sa compétence. Les descriptions satiriques de ces deux types de linguistes brossées par Fillmore,⁹ même si elles forcent un peu le trait, sont intéressantes pour qui cherche à comprendre ce qui sépare les deux approches. Voici celle du linguiste de corpus :

He has all the primary facts that he needs, in the form of a corpus of approximately one zillion running words, and he sees his job as that of deriving secondary facts from his primary facts. At the moment, he is busy determining the relative frequencies of the eleven parts of speech as the first word of a sentence versus the second word of a sentence.

Celle du linguiste partisan du recours à l'introspection n'est guère plus flatteuse :

He sits in a deep soft armchair, with his eyes closed and his hands clasped behind his head. Once in a while he opens his eyes, sits up abruptly shouting, 'Wow, what a neat fact!', grabs his pencil, and writes something down... having come still no closer to knowing what language is really like.

Cette amusante caricature n'est pas sans fondement, et les échanges de courrier de listes électroniques telles que CORPORA¹⁰ témoignent de cette divergence d'approche. Il est possible que l'animosité entre ces deux écoles relève quelque peu de la guerre de territoire : si plus personne ne songe à discuter l'utilité des corpus dans le travail du lexicographe ou du lexicologue (en particulier pour la description des collocations et de la phraséologie), certains syntacticiens et sémanticiens persistent à nier son utilité dans le cadre de leurs disciplines. Les linguistes de corpus eux-mêmes ne se privent pas de critiquer le peu d'usage que font les grammairiens des masses de données qui leur sont accessibles, comme le fait John Sinclair¹¹ :

⁹ Fillmore, C. (1992): "Corpus Linguistics" or "Computer-aided armchair linguistics" (p. 35), in Svartvik J. (ed.) (1992) *Directions in Corpus Linguistics*, Berlin: Mouton de Gruyter. (pp.35-60)

¹⁰ <http://gandalf.aksis.uib.no/corpora/>.

¹¹ Sinclair, J. (1991) *Corpus, Concordance, Collocation* (Oxford: Oxford University Press), p.100.

In grammars, the tradition of citing examples has faded away since *A Modern English Grammar* (Jespersen 1909; 1949). Even *A Comprehensive Grammar of the English Language* (CGEL) (Quirk *et al.* 1985) relies heavily on invented examples.

The CGEL had the corpus of the Survey of English Usage available to it: this is a corpus approaching one million words, spanning twenty-five years, and including a substantial proportion of spoken English. Occasional reference is given in the grammar to Survey material, but no attempt is made to confront and account for the evidence. Hardly any of the examples are citations, though citations must have been readily available. One is forced to conclude that the authors were following a methodology which gave low priority to one of the concerns of this book, which is to press for the use of actual language data as a basis for all descriptive statements.

Cette critique à peine voilée de la méthodologie utilisée par les célèbres grammairiens est typique de l'opposition entre les deux camps cités plus haut : ceux qui estiment que les jugements portés par les locuteurs natifs sur la grammaticalité d'énoncés inventés constituent une source suffisante pour décrire un système linguistique, et ceux qui pensent que toute description d'un tel système doit se fonder sur des faits de langue observables. Aux yeux de ces derniers, il est surprenant que les grammairiens aient choisi de négliger les données disponibles dans les corpus, dans la mesure où même des corpus de taille restreinte suffisent amplement à la tâche. Citons de nouveau Sinclair (1991: 100) :

The received wisdom of corpus linguistics is that fairly small corpora, of one million words or even fewer, are adequate for grammatical purposes, since the frequency of occurrence of so-called grammatical or function words is quite high. In the LOB corpus, for example (one million words of English printed in the UK in the year 1961), the commonest 100 words are almost all grammatical, and range in frequency from *the* at 68,315 to *people* at 953 [...].

On pourrait sans doute accuser Sinclair d'un biais lexicologique, puisqu'il semble vouloir réduire la description de phénomènes grammaticaux au seul usage des « mots grammaticaux ». Mais le fait est que nombre de grammairiens continuent d'utiliser dans leurs recherches des exemples qu'ils inventent ou qu'ils relèvent au gré de leurs lectures, négligeant ainsi les indications fournies par l'étude des corpus de plus grande taille sur les fréquences d'emploi de telle ou telle tournure qui, outre le fait qu'elles ont sans doute leur place dans les ouvrages de grammaires conçus pour les apprenants d'une langue

seconde, pourraient sans doute guider la réflexion des chercheurs dont la conception de la linguistique est moins pragmatique.

La question de la taille idéale d'un corpus est souvent posée, et il semble exister un consensus parmi les linguistes du domaine : pour ce qui est de l'étude du lexique, des collocations et de la phraséologie, plus la taille du corpus est élevée, plus la description sera précise (même si la question peut se poser différemment selon que l'on étudie une langue de spécialité ou la langue dite générale). On remarquera au passage que depuis sa première édition (1994) l'un des tout premiers dictionnaires bilingues anglais-français, le Grand Dictionnaire Hachette-Oxford (GDHO), fait précéder le mode d'emploi du dictionnaire d'une section intitulée « Un dictionnaire à partir de corpus » qui explique à l'utilisateur ce qu'est un corpus et l'usage qu'en font les lexicographes. Le fait que les auteurs de ce dictionnaire aient ainsi choisi de détailler pour leurs utilisateurs le processus de création d'une entrée ne doit pas être interprété uniquement comme la mise en avant d'un argument commercial ou la preuve d'un souci de pédagogie. Il prouve leur adhésion à une véritable « doctrine » de l'usage du corpus en vue d'une description exhaustive du lexique. Certains linguistes nous prédisent d'ailleurs un avenir dans lequel les corpus électroniques pourraient tout simplement prendre la place des dictionnaires,¹² et suggèrent que pour la consultation des collocations impliquant un item lexical donné, l'accès direct au corpus sur lequel se fonde le dictionnaire pourrait à l'avenir s'avérer préférable à la version synthétisée et nécessairement incomplète qu'offre le lexicographe dans l'entrée qui lui est consacrée.¹³

Corpus et langues de spécialité

Dans le domaine des langues de spécialité, l'utilisation de corpus électroniques a permis à de nombreux linguistes et informaticiens

¹² Cf. Greffenstette, G (1998): "The Future of Linguistics and Lexicographers: Will there be Lexicographers in the year 3000?" In Fontenelle, Thierry, Philippe Hiligsmann, Archibald Michiels, André Moulin, Siegfried Theissen (eds), *Euralex'98: Proceedings of the Eighth Euralex International Congress* Liège, Université de Liège, pp 25-42)

¹³ Cf. Clear, J. (1996): "Technical Implications of Multilingual Corpus Lexicography" In: *International Journal of Lexicography* Vol 9, N° 3, pp 265-276.

(parmi lesquels Bourigault¹⁴ et Daille¹⁵) de mettre au point des techniques de TAL (Traitement Automatique de la Langue) dont la principale application est l'extraction terminologique. L'une des caractéristiques des langues de spécialité étant la haute fréquence d'occurrence des expressions à base nominale, l'utilisation de corpus étiquetés grammaticalement (*POS-tagged*) permet d'identifier automatiquement les candidats-termes et de soumettre leur validation à des terminologues et à des spécialistes du domaine d'étude. L'étude des collocations et de la phraséologie dans ces langues de spécialité bénéficie également de l'utilisation de ces ressources, même si le repérage automatique est ici plus difficile que celui des termes. Les recherches en langue de spécialité ne se limitent d'ailleurs pas à la seule étude de phénomènes lexicaux : certains chercheurs comme Gledhill¹⁶ ont montré que les discours spécialisés pouvaient également se définir par la récurrence de certaines structures grammaticales.

Les grands corpus de références comme le TLF ou le BNC ont pour but principal de représenter le plus fidèlement possible la langue générale. La linguistique de corpus s'applique toutefois également aux langues de spécialité, parfois appelées langues spécialisées¹⁷ ou sous-langages. Cette dernière dénomination, employée par Z. Harris¹⁸ sous-entend, au delà d'une particularité du lexique, un fonctionnement langagier spécifique, et englobe non seulement les langages des disciplines scientifiques et techniques, mais également les méta-langages comme celui de la linguistique. L'hypothèse de Harris veut que les sous-langages aient un lexique limité et un nombre fini de schémas de phrases, qui seraient des combinaisons de sous-classes de

¹⁴ Bourigault, D. (1992): "Lexter: un logiciel d'extraction de terminologie" in Tama '92, Actes du 2^o Symposium TermNet : Applications terminologiques et microordinateurs, Vienne, Autriche.

¹⁵ Daille, B. (1994) : Approche mixte pour l'extraction automatique de terminologie : statistique lexicale et filtres linguistiques, Thèse de doctorat en informatique, Université Paris 7.

¹⁶ Gledhill, C.J., (1999) "The phraseology of rhetoric, collocations and discourse in cancer research abstracts" in C. Barron and N. Bruce (eds.), *'Knowledge and Discourse' Proceedings of the International Multidisciplinary Conference*. Hong Kong: 18-21 June 1996. University of Hong Kong, Hong Kong. April 1999.

¹⁷ Lerat préfère le terme *langue spécialisée*, car il conçoit la langue de spécialité non pas comme une « sous-langue » mais comme une forme d'usage particulière de la langue. Pour lui, « l'anglais *language for special purposes* (LSP) dit bien cette particularité [...]. Pour dire en français l'unicité de l'idiome et la particularité des univers de connaissances, « langue spécialisée » présente le même avantage : renvoyer au système linguistique pour l'expression et aux professions pour les savoirs. » (Lerat : 1995, p. 12).

¹⁸ Voir notamment : Harris Z., Gottfried M., Ryckman T., Mattick Jr P., Daladier A., Harris T., Harris S. (1989) : *The Form of Information in Science, Analysis of Immunology Sublanguage*, Kluwer Academic Publisher, Dordrecht.

mots propres à ces sous-langages. Pour McEnery et Wilson, c'est cette finitude qui rend l'utilisation des corpus particulièrement propice à l'étude des sous-langages¹⁹ :

It has been hypothesised that certain genres of writing actually represent a constrained subset of a natural language: these are sublanguages – a version of natural language which does not display all the creativity of that natural language. One key feature that has been hypothesised for a sublanguage is that it will show a high degree of closure at various levels of description. [...] In short, a corpus should be an exceptionally good tool for identifying and describing a sublanguage, because they both have an important feature in common – a finite nature. [...] Corpora should give us a very sharp contrast in terms of closure between any supposed sublanguage and any unconstrained language.

La dénomination de sous-langage pose toutefois le problème de l'inclusion de celui-ci dans la langue générale,²⁰ ce que l'appellation, traduite de l'anglais *sublanguage*, semble présupposer, et d'une manière plus générale celui de ce qui relève de la syntaxe ou du lexique dans la définition de sa spécificité.²¹ Ainsi, Habert *et al.* (1997) expliquent dans la section consacrée aux hypothèses de Z. Harris :

[Les] sous-langages ne sont pas forcément des sous-ensembles de la langue générale. Certains traits de la langue générale s'y retrouvent, d'autres leur sont propres. [...] Inversement, les sous-langages peuvent recourir à des patrons syntaxiques particuliers qu'il serait difficile d'intégrer tels quels à une grammaire « de langue ». C'est le cas de certains motifs dénommatifs qui forment de véritables « grammaires locales ».²²

Les auteurs citent en exemple de ces patrons syntaxiques particuliers une structure propre aux manuels informatiques de

¹⁹ Mc Enery et Wilson (1996), p. 148-149.

²⁰ Frandsen (1998) fait remarquer que les modèles d'inclusion des langues spécialisées dans la langue générale sont souvent fondés sur une idée erronée de la généralité. Son propre modèle, fondé sur la distinction saussurienne entre langue et parole, propose une distinction asymétrique entre langue générale et langue spécialisée, la première se situant sur un axe épistémologique, la deuxième sur un axe ontologique.

²¹ On pourrait ajouter que l'appellation de « sous-langue » serait probablement plus pertinente que le calque imposé par l'usage.

²² Habert, B., Nazarenko, A. & Salem A. (1997) : *Les linguistiques de corpus*, Paris, Armand Colin / Masson. (p. 149)

l'anglais mentionnée par Black *et al.* (1993)²³ : *to vary on* [*un dispositif*], signifiant approximativement « le mettre en marche » dans des phrases comme « *The system will be unable to vary on the device* » (dans ce cas précis, il nous semble cependant que les auteurs considèrent comme relevant de la grammaire ce qui représente une simple variation lexicale sur le verbe à particule adverbiale *to turn on*, et non pas un patron syntaxique typique de la langue de spécialité).

Si l'on tente de dissocier ce qui relève de la syntaxe et du lexique dans les langues de spécialité, on observe bien sûr une dissociation partielle des lexiques de la langue générale et des langues de spécialité d'une part, et des langues de spécialité entre elles d'autre part. Pour prendre l'exemple de la langue médicale anglaise, même des corpus de près de 200 millions de mots (comme celui du cédérom *Internal Medicine 93*) ne contiennent aucune occurrence d'adjectifs de la langue générale qui sont d'origine dialectale (*bodacious*), appartiennent au registre familier (*groovy*) ou expriment une opinion dont la coloration affective est incompatible avec l'objectivité requise par les comités de lecture des articles de recherche scientifique (*loathsome*). De la même manière, de nombreux adjectifs composés utilisés en langue médicale (*cervicothoracic-lumbosacral*, *hypothalamic-pituitary-gonadal*) sont absents des corpus de langue générale. S'il s'agit d'adopter une symbolisation qui reflète la relation entre les lexiques des langues de spécialité et de la langue générale, la relation d'intersection des deux ensembles, sous forme d'un diagramme de Venn, semble plus appropriée, les pans ne faisant pas partie de cette intersection représentant les parties du lexique spécifiques aux deux types de langue. Il faut cependant nuancer cette approche purement quantitative, dans la mesure où une grande partie du lexique de la langue générale est investie d'un sens spécifique en langue de spécialité, notamment de manière à réduire les phénomènes de polysémie. Ainsi, le mot « sensibilité », qui désigne dans le vocabulaire des essais cliniques l'équivalent de la notion de rappel en traitement automatique des langues, ne fait jamais référence à une qualité applicable à un humain dans la littérature médicale.²⁴

²³ Black E., Garside R., Leech, G., Eyes E., Mc Enery A., Lafferty J., Magerman D., Roukos S. (1993) : *Statistically-driven computer grammars of English : the IBM/Lancaster approach*, Rodopi, Amsterdam, 1993. (p. 112)

²⁴ De tels blocages contextuels ne s'appliquent pas uniquement au sens des mots polysémiques, mais aussi à la catégorie grammaticale de certains homographes (par exemple, en anglais médical, *novel* est toujours un adjectif et jamais un nom). Nous avons déjà eu l'occasion de souligner que l'intégration de telles restrictions locales aux programmes d'étiquetage morpho-syntaxique

L'utilisation du Web comme Corpus

Alors que le *British National Corpus* (101 millions de mots) est resté figé et décrit donc exclusivement l'anglais de la fin du siècle dernier, d'autres corpus sont en constante extension. Le corpus *Bank of English* compte à l'heure actuelle quelque 650 millions de mots et le *Corpus of Contemporary American English* (COCA, <http://www.americancorpus.org/>) totalise 385 millions de mots. En dépit de leur taille considérable, ces corpus sont parfois insuffisants pour l'étude de structures syntaxiques ou de lexies extrêmement rares. De nombreux chercheurs se tournent donc vers les moteurs de recherche du Web pour en extraire des données linguistiques.

Si l'on examine quelques-unes des définitions fondatrices des corpus, on constate qu'un grand nombre de traits définitoires semblent exclure le Web de la catégorie des corpus. Sinclair (1991, p.171), exprime dans sa définition (*A corpus is a collection of naturally-occurring language text, chosen to characterize a state or variety of a language.*) la notion de choix ou de sélection, ce qui impliquerait une recherche non pas sur l'ensemble du Web, mais sur une partie de celui-ci seulement.²⁵ Biber (1998, p. 4) le rejoint sur ce principe de sélection en mentionnant « *a large and principled collection of natural texts* ». La définition donnée par le groupe d'experts EAGLES (*Expert Advisory Group on Language Engineering Standards*) est encore plus restrictive :

A corpus is a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language.

L'utilisateur des moteurs de recherche du Web est très loin de cet échantillonnage attentif censé prétendre à la représentativité.

Si l'on considère ce qui rapproche le Web d'un corpus, on peut énumérer les traits suivants :

pourrait en améliorer l'efficacité (Maniez, F: (2001): « Désambiguïisation syntaxique des groupes nominaux en langue spécialisée : le cas des adjectifs en anglais. » Actes du colloque TALN de Tours, 2-5 juillet 2001, Tome 1, pp. 273-282.).

²⁵ Cette possibilité est donnée par des logiciels tels que Kwicfinder (<http://www.kwicfinder.com/KWiCFinder.html>), développé par William Fletcher.

- le Web s'apparente à un corpus de très grande taille, constamment mis à jour, mais qui échappe par nature à tout type d'annotation, et dont la composition est constamment changeante.

- les données linguistiques qui y sont présentes peuvent être considérées comme représentatives de la manière dont les personnes ayant les moyens financiers et les connaissances techniques nécessaires pour mettre des documents en ligne utilisent leur langue maternelle. Il ne s'agit donc pas d'un échantillon représentatif de la population dans son ensemble.

- les langues y sont encore très inégalement représentées (à titre indicatif, on dénombre 146 000 résultats pour le mot *oreillons* et 3 270 000, soit 22 fois plus, pour son équivalent anglais *mumps*).²⁶

Un autre problème souvent évoqué lorsque l'on envisage l'utilisation du Web à des fins linguistiques est le pourcentage important de formes non standard²⁷ (notamment sur ses forums) et le fait qu'il est souvent impossible d'affirmer avec certitude que les écrits qui y sont postés proviennent de locuteurs natifs (l'anglais étant *de facto* la *lingua franca* de la communauté scientifique et du monde des affaires, ce dernier problème le concerne tout particulièrement). Parfois, ce n'est pas l'acceptabilité grammaticale qui est contestable, mais la manière dont certaines expressions sont utilisées par les locuteurs natifs eux-mêmes. Ainsi, l'expression *steep learning curve* (courbe d'apprentissage à croissance rapide) fait régulièrement l'objet d'un contresens par les locuteurs anglophones du fait que la prosodie sémantique de l'adjectif *steep* (raide, pentu) est essentiellement négative, suggérant la difficulté. L'immense majorité des occurrences de la séquence *learning curve is steep but* concerne ainsi des emplois erronés de cette expression par des natifs.

Le Web permet cependant des recherches concernant certains patrons syntaxiques rares pour lesquelles des corpus tels que le BNC ou le COCA (*Corpus of Contemporary American English*) sont de taille insuffisante. Ces corpus se sont avérés insuffisants, par exemple, lors d'une étude sur la lexicalité des suites Adjectif-Nom-Nom en

²⁶ Ces données et les suivantes concernent des recherches effectuées sur la version française du moteur de recherche Google (<http://www.google.fr/>) le 15 février 2009.

²⁷ On relève par exemple 655 000 résultats pour la forme standard de l'anglais *I couldn't care less* et plus de deux fois ce chiffre pour son équivalent non standard *I could care less*. En français, un passé simple inconnu de Bescherelle (*conclua*) compte plus de trois millions d'occurrences.

anglais,²⁸ lorsque les auteurs ont recherché dans les corpus existants des énoncés contenant des emplois attributifs de certains adjectifs composés anglais (par exemple *long-term*, *full-time*, *large-scale*, *last-minute* ou *high-speed*). Le tableau ci-dessous reproduit les chiffres obtenus pour les vingt adjectifs les plus fréquents de ce type. Dans la mesure où la séquence formée d'une forme du verbe *be* et d'un adjectif est majoritairement suivie d'un nom en anglais (utilisation épithétique), la conjonction *and* a été ajoutée à cette séquence, les formes de ce type rencontrées sur le Web étant majoritairement attributives (par exemple, l'utilisation de l'adjectif *long term* – parfois épelé *long-term* – dans l'énoncé : *Brown said Britain's commitment to Afghanistan was long term and he urged other countries to do more in sharing the burden*).

Le Tableau 1 (page suivante) compare donc l'ensemble des formes de verbe *be* suivies des adjectifs formés des mots des deux premières colonnes et de *and* dans le COCA (385 millions de mots) à l'équivalent de ce qui n'est sur le Web qu'un sous-ensemble de ces mêmes formes (celui contenant l'une des formes de *be*, c'est-à-dire *was*). Les chiffres parlent d'eux-mêmes, même si certains auteurs supposent que les résultats de certains moteurs de recherche sont surestimés.²⁹

²⁸ Pierre J.L. Arnaud, Emmanuel Ferragne, Diana M. Lewis et François Maniez (2008), "[Adj+N] sequences in attributive or NP-final positions: observations on lexicality" in : *Phraseology : an interdisciplinary perspective* (Granger, S. & Meunier, F. eds), pp. 111-125. Benjamins, Amsterdam.

²⁹ Voir à ce sujet l'une des pages du blog de Jean Véronis : <http://aixtal.blogspot.com/2005/02/web-le-mystre-des-pages-manquantes-de.html>.

Adjectif	Nom	“[TO BE] ADJ-N and” (COCA 2009)	“was ADJ N and” (Google 2009)
long	term	20	972
short	term	7	16 100
full	time	7	13 000
high	tech	6	446
high	quality	4	9 050
small	scale	3	1 940
large	scale	2	557
open	air	1	4 910
high	speed	1	1 930
last	minute	0	6 950
full	length	0	5 850
low	cost	0	4 280
long	distance	0	3 110
high	level	0	1 620
high	end	0	1 250
low	level	0	1 220
late	night	0	503
long	range	0	48
full	scale	0	43
white	collar	0	36

Tableau 1 : Estimations des emplois attributifs de certains adjectifs composés anglais dans le corpus COCA et sur le Web.

Quel est le rapport de taille entre le Web actuel et les corpus de référence que nous avons cités? Les chiffres obtenus pour le corpus COCA dans le Tableau 1 ne sont pas suffisamment élevés pour qu’une comparaison soit statistiquement significative. Toutefois, l’étude des fréquences d’emploi sur le Web de certains adjectifs de fréquence comprise entre 10 et 100 dans le BNC (Tableau 2) peut nous donner une vague idée du facteur multiplicatif. La moyenne des coefficients multiplicateurs était supérieure à 5 000 en 2005 et elle est supérieure à 50 000 en 2009. Ces chiffres suggèrent donc une multiplication par dix de la taille du Web anglais au cours des quatre dernières années. La taille du BNC étant légèrement supérieure à 100 millions de mots, la partie anglaise du Web serait ainsi passée de 500 milliards de mots à 5 000 milliards de mots. Au vu de l’ampleur du « corpus » ainsi disponible, il n’est pas surprenant qu’un grand nombre de linguistes et de spécialistes du TALN aient recours aux données linguistiques présentes sur le Web, qu’il s’agisse d’étudier la fréquence de diverses

séquences polylexicales, de créer automatiquement des corpus multilingues parallèles, ou de tenter de résoudre automatiquement les problèmes d'ambiguïté sémantique, ces deux dernières activités étant par ailleurs la clé de l'amélioration qualitative des programmes de traduction assistée par ordinateur (TAO).

Adverbe	Occurrences du BNC	Occurrences du Web (2005)	Coeff. 2005	Occurrences du Web (2009)	Coeff. 2009
abominably	22	71 900	3 268	176 000	8 000
abysmally	20	116 000	5 800	269 000	13 450
acceptably	33	457 000	13 848	826 000	25 030
admiringly	89	130 000	1 460	401 000	4 505
adoringly	18	56 900	3 161	619 000	34 388
advantageously	29	269 000	9 275	2 860 000	98 620
adventurously	10	33 900	3 390	2 950 000	295 000
aerobically	31	145 000	4 677	395 000	12 741
agonizingly	26	123 000	4 730	767 000	29 500
agreeably	69	246 000	3 565	946 000	13 710
Moyenne	34,7	164 870	5 317,4	1 020 900	53 494,4

Tableau 2 : Estimations des emplois de certains adverbes anglais dans le « British National Corpus » et sur le Web.

Conclusion

Il est probable que l'ignorance réciproque des « linguistes de fauteuil » et des linguistes de corpus est appelée à se prolonger, leurs méthodes et leurs buts étant fondamentalement opposés. Il est pourtant bien clair que les deux approches peuvent et doivent coexister, les intuitions des uns pouvant parfois être vérifiées quantitativement par les autres. Mais les corpus n'ont pas forcément vocation à servir à la découverte des universaux ou de modèles cognitifs ayant pour but d'expliquer la capacité des humains à s'exprimer par le langage. Leur usage est résolument tourné vers l'étude de la performance, qu'il s'agisse de faire progresser la didactique de l'apprentissage des langues étrangères par l'exploitation de corpus d'apprenants ou de décrire les problèmes de traduction relatifs à divers couples de langues

par l'utilisation des corpus multilingues. Certains linguistes, à l'instar d'Elena Tognini-Bonelli (2001) vont même jusqu'à considérer que l'utilisation des corpus dans ces deux domaines ne doit plus se contenter d'être confirmatoire de certaines intuitions du linguiste, mais devenir purement exploratoire (*corpus-driven linguistics*) afin de dépasser les limites des grammaires et des dictionnaires destinés aux apprenants d'une langue seconde :

[...] we have to abandon the fiction that each word is some kind of independent selection, and accept that the choice patterns of words in text can create new, large and complex units of meaning...

C'est cette possibilité d'accès à l'ensemble des contextes d'utilisation d'un mot donné qui permet effectivement à l'utilisateur des corpus de s'imprégner du sens (de tous les sens) d'un lexème donné. C'est d'ailleurs le paradoxe de l'utilisation du concordancier, qui extrait les exemples d'usage de leur contexte, mais qui en les rapprochant en crée une autre vision. La linguistique d'avant les corpus a parfois été comparée à l'astronomie d'avant le télescope. C'est peut-être le principal mérite de la linguistique de corpus que de nous donner une nouvelle vision de ce que peut être l'accès au sens des mots, qui comme le rappelle Cruse (1986) passe fondamentalement par l'étude de leur contexte :

[The] meaning of a word is fully reflected in its contextual relations; in fact, we can go further, and say that, for present purposes, the meaning of a word is constituted by its contextual relations.

Qu'il s'agisse de sémantique ou de syntaxe, d'apprentissage des langues, de traductologie ou d'étude des langues de spécialité, la linguistique de corpus continuera d'apporter sa pierre à l'édifice d'une linguistique de la performance.

Références

- Aarts, J. Meijs W. (eds) (1984), *Corpus Linguistics*. Amsterdam, Rodopi.
- Arnaud, P. J.L., E. Ferragne, D. M. Lewis et F. Maniez (2008), “[Adj+N] sequences in attributive or NP-final positions: observations on lexicality”, in *Phraseology: an interdisciplinary perspective* (Granger, S. & Meunier, F. eds), pp. 111-125. Benjamins, Amsterdam.
- Biber, D. (1998), *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press.
- Black, E., Garside, R., Leech, G., Eyes, E., McEnery, A., Lafferty, J., Magerman, D., Roukos, S. (1993), *Statistically-driven computer grammars of English : the IBM/Lancaster approach*. Amsterdam, Rodopi.
- Bourigault, D. (1992), “Lexter : un logiciel d'extraction de terminologie.”, in *TAMA '92, Actes du 2^o Symposium TermNet : Applications terminologiques et microordinateurs*. Vienne, Autriche.
- Bourigault, D., Jacquemin, C., L'Homme, M.-C. (eds) (2001), *Recent Advances in Computational Terminology*. John Benjamins Publishing Company, Amsterdam.
- Chomsky, N. (1957), *Syntactic Structures*. The Hague, Mouton
- Chomsky, N (1962), *paper* donné à l'University of Texas 1958, 3rd Texas Conference on Problems of Linguistic Analysis in English. Austin: University of Texas.
- Church, K.-W. et Hanks, P. (1990), “Word Association Norms, Mutual Information and Lexicography”. *Computational Linguistics*, vol. 16, n^o 1, pp. 22-29.
- Clear, J. (1996), “Technical Implications of Multilingual Corpus Lexicography”, in *International Journal of Lexicography* Vol 9, N^o 3, pp 265-276.
- Cruse, D. (1986), *Lexical Semantics*. Cambridge University Press, 328 pages.
- Daille, B. (1994), *Approche mixte pour l'extraction automatique de terminologie : statistique lexicale et filtres linguistiques*, Thèse de doctorat en informatique, Université Paris 7.
- Daille, B. (2001), “Qualitative terminology extraction: Identifying relational adjectives”, in Bourigault, D., Jacquemin, C., L'Homme,

- M.-C., (eds) *Recent Advances in Computational Terminology*. John Benjamins Publishing Company, Amsterdam, pp 149-166.
- Fellbaum, C. (ed) (1998), *WordNet: An Electronic Lexical Database*. MIT Press.
- Fillmore, C. (1992), "Corpus Linguistics" or "Computer-aided armchair linguistics", in SVARTVIK J. (ed.) *Directions in Corpus Linguistics*. Berlin: Mouton de Gruyter. (pp.35-60)
- Frandsen, F. (1998), "Langue générale et langue de spécialité – une distinction asymétrique?", in GAMBIER Y. (éd.) *Discours professionnels en français*. Peter Lang, Frankfurt am Main, pp. 15-34.
- Gledhill, C.-J. (2000), "The Discourse Function of Collocation in Research Article Introductions", in *English for Specific Purposes*. Volume 19/2:115-135.
- Granger, S. (ed.) (1998), *Learner English on Computer*. Longman.
- Greffenstette, G. (1998), "The Future of Linguistics and Lexicographers: Will there be Lexicographers in the year 3000? ", in Fontenelle, Thierry, Philippe Hiligsmann, Archibald Michiels, André Moulin, Siegfried Theissen (eds), *Euralex'98: Proceedings of the Eighth Euralex International Congress* Liège, Université de Liège, pp 25-42)
- Habert, B., Nazarenko, A., Salem, A. (1997), *Les linguistiques de corpus*. Paris, Armand Colin / Masson.
- Johansson, S. (1982), *Computer Corpora in English Language Research*. Bergen: Norwegian Computing Centre for the Humanities.
- Kucera, H. & Francis, W. (1967), *Computational Analysis of Present-Day American English*. Providence, Rhode Island: Brown University Press.
- Maniez, F. (2001), "Désambiguïstation syntaxique des groupes nominaux en langue spécialisée : le cas des adjectifs en anglais." Actes du colloque TALN de Tours, 2-5 juillet 2001, Tome 1, pp. 273-282.
- McEnery, T. et Wilson, A. (1996), *Corpus Linguistics*. Édinburgh, Edinburgh University Press.
- Sinclair, J. (1991), *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Svartvik, J. et Quirk, R. (eds.) (1980), *A Corpus of English Conversation*. Lund Studies in English 63.

- Teubert, W. (1996), "Comparable or Parallel Corpora?", in *International Journal of Lexicography* Vol 9, N° 3, pp 238-264.
- Tognini-Bonelli, E. (2001), *Corpus Linguistics at Work*. Benjamins, 224 pages.
- Véronis, J. et Khouri, L. (1995), "Étiquetage grammatical multilingue : le projet MULTEXT", in *T.A.L., Volume 36, 1995, numéros 1-2*, pp. 233-248.

Dan McIntyre & Brian Walker
University of Huddersfield (United Kingdom)

Annotating a corpus of Early Modern English writing for categories of discourse presentation

1. Introduction

This article discusses the process of annotating a small corpus of Early Modern English writing that we have constructed in order to investigate the diachronic development of speech, writing and thought presentation. The work we have done so far is a pilot investigation for a planned larger project. We have constructed a corpus of approximately 40,000 words of Early Modern English (EModE) fiction and news journalism and annotated it for categories of discourse presentation (DP) drawn from a model originally proposed by Leech and Short (1981). This has allowed us to quantify the types of discourse presentation within the corpus and to compare our findings against those from a similarly annotated corpus of Present Day English (PDE) writing (reported in Semino and Short 2004). Our results so far appear to indicate developing stylistic tendencies in fiction and news texts in the Early Modern period, and suggest that it would be profitable to extend the project through the construction of a larger corpus incorporating a greater number of text-types in order to test our hypotheses more rigorously. In this article we concentrate specifically on describing the annotation phase of the project. We discuss the criteria by which we defined the various discourse presentation categories in order to make clear our analytical methodology, as well as the issues we were confronted with in trying to annotate in a systematic and retrievable way. We conclude with some preliminary results to illustrate the value of this kind of annotation and suggest some hypotheses resulting from this pilot investigation.

2. Discourse presentation

Prototypically, discourse presentation (also known as speech, writing and thought presentation – or SW&TP) refers to the presentation in a posterior discourse context of speech, writing or thought from an anterior discourse context. A person may report the speech, writing and/or thoughts of a third party using a variety of different forms. Hence, the original utterance ‘I love corpus linguistics!’ may be reported by a third party using any of the following structures (see Table 1 for a description of the categories).

- (i) ‘I love corpus linguistics!’
- (ii) ‘I love corpus linguistics!’ he said.
- (iii) He said that he loved corpus linguistics.
- (iv) He loved corpus linguistics!
- (v) He expressed his enjoyment.
- (vi) He spoke loudly.

Each of the above forms expresses varying degrees of what Leech and Short (1981, 2007) have termed narrator interference, as well as decreasing claims to faithfulness with regard to the reporting of the original utterance. Example (i) expresses the exact words of the original utterance; (ii) includes the exact words plus a reporting clause indicating the presence of a narrator; (iii) presents the original utterance in an indirect form, with the original speaker’s words contained within a subordinate clause, having been subjected to a backshift in tense; (iv) is a free indirect rendering that blends aspects of a narratorial report with a flavour of the original speaker’s utterance (in this case, the exclamation mark); (v) reports only the speech act of the original speaker (none of the propositional content of the original utterance can be reconstructed); (vi) reports only the fact that speech occurred.

Examples (i) to (vi) constitute speech presentation, though the same principles apply to the presentation of a third party’s writing or thoughts. Discourse presentation can also refer to the presentation of speech, writing or thought in some future discourse context. For example, *He’s about to say how much he loves corpus linguistics*. Table 1 shows the categories of discourse presentation that we used in our project:

Speech presentation		Writing presentation		Thought presentation	
(F)DS	(Free) Direct Speech	(F)DW	Free Direct Writing	(F)DT	Free Direct Thought
FIS	Free Indirect Speech	FIW	Free Indirect Writing	FIT	Free Indirect Thought
IS	Indirect Speech	IW	Indirect Writing	IT	Indirect Thought
NRSA	Narrator's (Re)presentation of a Speech Act	NRWA	Narrator's (Re)presentation of a Writing Act	NRTA	Narrator's (Re)presentation of a Thought Act
NV	Narrator's Presentation of Voice	NW	Narrator's Presentation of Writing	NT	Narrator's Presentation of Thought
				NI	Internal Narration
NRS	Narrator's Report of Speech	NRW	Narrator's Report of Writing	NRT	Narrator's Report of Thought
N	Narration	N	Narration	N	Narration

Table 1 Speech, writing and thought presentation model.
Based on the description in Short (2007)

The categories in Table 1 are those described in Short (2007), itself a development of those presented originally in Leech and Short (1981) and later expanded by Short and a project team at Lancaster University. The model proposed by Leech and Short (1981) suggests that with each move along the cline of discourse categories comes an increased claim to faithfulness with regard to the reporting of the original discourse. (The only deviation from this in the model concerns the Free Direct (FDS/W/T) and Direct (DS/W/T) categories, which are conflated because they represent the same degree of faithfulness to the original). In later conceptions of the model, Short *et al.* dispensed with the notion of discourse report, preferring instead to describe the phenomenon as discourse presentation, as a result of the fact that hypothetical and forward-facing discourse presentation does not involve the report or representation of something already said, written or thought. Nonetheless, the 'R' element (for 'report' or 'representation') has been retained in favour of 'P' (for 'presentation') in some of the acronyms in Table 1 to avoid confusion with earlier publications on the subject. While we occasionally refer to discourse report or representation, this is only to avoid the confusion that might arise from changing the acronyms in Table 1, and all such references should be taken as referring to the *presentation* of discourse.

Discourse presentation as a linguistic phenomenon has been studied from a wide range of academic perspectives, including philosophy

(Clark and Gerrig 1990), applied linguistics (Baynham and Slembrouck 1999, Myers 1999), sociology (Holt 1999, Holt and Clift 2006) and psychology (Ravotas and Berkenkotter 1998). Our interest in the phenomenon relates to its stylistic import, hence our use of a discourse presentation model developed from research in linguistic stylistics. Our interest in the diachronic development of the phenomenon is what prompted our study of Early Modern English discourse presentation. Our choice of this period was determined by the fact that this phase of the development of English saw the rise of a standard form of the language as well as an increase in printed texts and literacy. Since the Early Modern period was such a rich era for linguistic development, we reasoned that discourse presentation as a stylistic technique might be used differently from how it is in PDE. There has been some work on the phenomenon from a historical linguistic perspective, though none has used the methodological framework we employ here. Moore (2002), for instance, explores the phenomenon from a qualitative angle, while Jucker (2006), although taking a corpus-based approach, analyses only one text-type (news discourse) and uses an un-annotated corpus. One consequence of this is that Jucker's findings are limited by the structural forms of discourse presentation that it is possible to search for. For example, Jucker does not analyse free indirect discourse, since this is impossible to retrieve through concordancing; free indirect discourse is only apparent by its context, not by its linguistic form. Our technique of first annotating our data means that we are able to retrieve all instances of all the categories of discourse presentation outlined in Short's (2007) model (see Table 1).

The nature of our project locates it within the growing field of corpus stylistics (see, for example, Semino and Short 2004, Mahlberg 2007, O'Halloran 2007a, 2007b), and particularly historical corpus stylistics (Studer 2008, Culpeper and Kytö 2002, 2006). Our two principle aims are to investigate the forms and functions of discourse presentation in Early Modern English writing and to provide quantitative evidence of the relative frequencies of presentational forms. Long term, a subsidiary aim is to provide a perspective on the history of English that is focused on stylistic developments and goes beyond formal levels of language, thereby contributing to the 'alternative histories of English' advocated in Watts and Trudgill (2002).

3. Corpus construction

Since our aim was to compare the forms and functions of discourse presentation in EModE with those of PDE, the sampling frame for our corpus follows the principles of the Lancaster SW&TP Written Corpus, a 260,000-word corpus of contemporary English writing annotated for the categories of speech, writing and thought presentation outlined in Table 1. The Lancaster corpus comprises equal numbers of 2000-word samples of serious and popular fiction (broadly akin to ‘high’ and ‘low’ literature), tabloid and broadsheet news journalism and biography and autobiography. The labour intensive nature of speech, writing and thought presentation annotation meant that we were unable to construct a corpus of equivalent size, and so we restricted our text-types to just fiction and news reports (needless to say, our quantitative comparisons in section 5 are with the fiction and news sections of the Lancaster corpus only, and we carried out log-likelihood calculations to determine whether differences in tag frequencies between the two corpora were statistically significant). We have around 20,000 words of each text-type, divided equally across fifty-year segments of the Early Modern English period. In defining this time-frame we took the common consensus of historical linguists who date the period from approximately 1500 to approximately 1750, these dates being, respectively, roughly synonymous with Caxton’s printing press revolution taking effect and the publication of Johnson’s dictionary. (This is not to suggest that these two events had an equal impact on all varieties of English; we are primarily interested in the developing Standard English of the period, on which Caxton and Johnson clearly did have an effect). Tables 2 and 3 outline the content of the fiction and news sections of our corpus, as well as the time periods they are representative of. It is worth noting that our earliest examples of news journalism are somewhat different from PDE newspapers, since the newspaper as a text-type did not evolve until mid-way through the Early Modern period. The newspaper (as we might recognise it) did not exist at the earlier end of our time frame, and news was often in the form of letters or personal accounts, which were printed and distributed on a fairly limited basis. News pamphlets (also called *Corantos*, or *News books*) first started appearing towards the end of the 16th Century and became established in the early 17th Century.

What is often regarded as the first proper newspaper, *The London Gazette*, did not appear until 1666. Our earliest samples of news journalism are therefore of the contents of letters describing newsworthy events (for example, J1.2 ‘Hevy news of an earthquake’), and while the data is not absolutely equivalent to the Lancaster news data, it does afford an opportunity to gain an insight into how the genre develops across the period. A further point to note is that, unlike the Lancaster team, we did not distinguish between serious and popular fiction, since the distinction did not exist in the Early Modern period in quite the same way. Furthermore, sub-dividing our data in this way would not have been a good methodological strategy, since this would have generated raw figures too small to draw reliable conclusions from.

EModE Corpus – Prose Fiction sub-section					
Extract No.	Period	Title	Word count	Author	Pubn Date
PF1.1	1500-1549	The Noble History of King Ponthus	2072	Henry Watson	1511
PF1.2		The Mad Men of Gotham	2002	William Tyndale	1547
PF2.1	1550-1599	The Carde of Fancie	2154	Robert Greene	1584
PF2.2		Arcadia	2022	Philip Sydney	1590
PF3.1	1600-1649	The Blacke Booke	2057	attr. Thomas Middleton	1604
PF3.2		Cloria and Narcissus	2047	attr. Percy Herbert	1653
PF4.1	1650-1699	The blazing-world	2097	Margaret Cavendish	1668
PF4.2		Oroonoko	2073	Aphra Behn	1688
PF5.1	1700-1750	Moll Flanders	1993	Daniel Defoe	1722
PF5.2		Tom Jones	2079	Henry Fielding	1751
Total Words			20596		

Table 2 The composition of the fiction section of the EModE corpus

EModE Corpus – News Report sub-section					
Extrect No.	Time Period	Title	Word count	Author	Pubn Date
J1.1	1500-1549	An account of the Battle of Flodden	2200	Not Known	1513
J1.2		Hevy newes of an earthquake	825	Not Known	1542
J1.3		A copy of a letter containing certayne newes	1017	Not Known	1549
J2.1	1550-1599	The Spoyle of Antwerpe	2122	George Gascoigne	1576
J2.2		The English Mercurie	1391	Not Known	1588
J3.1	1600-1649	The weeklely Newes	1079	Not Known	1606
J3.2		The courant of newes	1386	Not Known	1620
J3.3		The marchings of Two Regiments	2101	Henry Foster	1643
J4.1	1650-1699	Every Day's Intelligence 1	1019	Heneage Finch	1653
J4.2		Every Day's Intelligence 2	1013		1653
J4.3		A true designe of the Late Eruption of Mt Etna	2170		1669
J5.1	1700-1750	The Flying Post	1107	Not Known	1700
J5.2		London Post	1184	Not Known	1700
J5.3		Country Journal	1876	Not Known	1736
Total Words			20490		

Table 3 The composition of the news section of the EModE corpus

We collected our data from a variety of sources, using texts that were already in electronic format, as this represented a great time saving and, where possible, checked the electronic version against facsimiles of original publications of the texts. This was to make sure that the later edited version of early texts, which the electronic forms were often drawn from, did not contain, for example, extra or altered punctuation. Our sources included: Early English Books Online (EEBO); the Oxford Text Archive (OTA); Renaissance Editions (University of Oregon); Project Gutenberg; the Lampeter Corpus; the Lancaster Newsbook Corpus; the Corpus of English Dialogues 1560-1760; and a fairly new resource called The Burney Collection, which is a collection of facsimiles of newspapers available from the British Library.

4. The annotation process

The annotation scheme we used was a development of that outlined in McIntyre et al. (2004), which describes the results of a similar project to annotate a corpus of contemporary spoken English for discourse presentation. This involves the application of TEI-conformant XML mark-up that comprises an element *dptag* (*discourse presentation tag*) and an attribute *cat* (*category*). These are enclosed within angle brackets forming what, in shorthand reference, is called a *tag*. The *cat* attribute consists of fifteen fields into which pre-designated alphanumeric codes are entered detailing the SW&TP categories outlined in Table 1. Each field has a limited number of possible constituents and the combination of constituents from different fields allows for a detailed description of the discourse presentation being marked (these are set out in Table 4).

Field	Possible constituents	Definition of constituent
1	x N F	Narrator's; Narration; Free
2	x R I D	Representation; Indirect; Direct
3	x S T W V I M	Speech; Thought; Writing; Voice; Internal state; Use
4	x A	Act
5	x p	Topic
6	x #	# = odd/interesting cases
7	x y	discourse summary
8	x g a	grammatical negative; absence of speech, thought and/or writing
9	x h	hypothetical
10	x i	inferred
11	x q	quote
12	x r	iterative
13	x v p	interrogative; imperative
14	x m	nominalisation
15	x 1 2 3 4	no.s = DP split into sections

Table 4 Constituents of the fields of the *cat* attribute

The possible constituents designated to the first four fields relate to the major DP categories (outlined in Table 1) and are always capital letters. The constituents designated to the remaining eleven fields relate to DP sub-categories and provide further details about the DP. These are generally lower-case letters, but the hash symbol (#) and numbers are also possible in certain fields. We use *x* as a placeholder and do not mark empty positions following the final attribute value.

This means that “cat” attribute constituents always occur in the same field position, making searches of the annotated corpus for particular DP categories using computer tools more straightforward.

Below is an example of three annotated sentences from the fiction section of the corpus:

```
<dptag cat="N"> Here the book dropt from her hand, and a shower of tears ran  
down into her bosom. In this situation she had continued a minute, when the  
door opened, and in came Lord Fellamar. Sophia started from her chair at his  
entrance; </dptag> <dptag cat="NRS"> and his lordship advancing forwards,  
and making a low bow, said, </dptag> <dptag cat="xDS"> "I am afraid, Miss  
Western, I break in upon you abruptly." </dptag>
```

The example shows that the code to mark-up direct speech is ‘xDS’: the “cat” fields that contain the constituents for direct speech are 2 and 3; field 1 is not required, so is filled with an ‘x’; fields 4 to 15 are left blank. Notice that in the example reporting clause (NRS) is tagged. Our annotation policy for this study was to also tag narration and narration phenomena (such as reporting clauses) as well as DP, since this often impacts on the stylistic effect of the DP, as we will show in section 5. It is also the case that our example shows that for every tag that marks the start of a new section of DP or narration phenomena, there is also an end tag (</dptag> in our case) which marks the end of that stretch of DP or narration.

We also indicate instances of embedded discourse presentation using an *e[n]dptag*, where *e* stands for ‘embedded’ and *n* indicates the level of discourse embedding. An example of this can be seen below:

```
<dptag cat="xDS"> So yelde I me to you & in to your pryson as your knyght &  
ye to haue power to doo as of your owne  
  <e1dptag cat="NRS">& yet he bad me</e1dptag>  
  <e1dptag cat="xIS">yt I sholde salewe you from hym.</e1dptag>  
</dptag>
```

The example immediately above shows an instance of one level of discourse embedding. More are possible though it is rare to find instances beyond three levels.

It should be recognised that ambiguity is a large part of discourse presentation and we marked this by using ambiguous tags. Consider the following example from the fiction section of the corpus:

But in the other Chappel lined with the Star- stone, she preached Sermons of Comfort to those that repented of their sins [...]

Focusing just on the underlined section of the extract, it is unclear whether repenting, in this case, involved a speech act ('I repent of my sins') or some sort of thought process or act, or both. The case, therefore, is genuinely ambiguous, and our annotation scheme reflects this by using a *cat2* attribute to mark an alternative analysis. The *cat2* attribute follows exactly the same format as the *cat* attribute. Thus, the resulting tagging format for the above example is:

```
<dptag cat="N"> But in the other Chappel lined with the Star- stone, </dptag>  
<dptag cat="NRSAP"> she preached Sermons of Comfort </dptag>  
<dptag cat="NRSA" cat2="NRTA"> to those that repented of their sins [...]  
</dptag>
```

While we do not discuss ambiguous examples in this article, we will return to this issue in future research.

Annotating in this way has a number of methodological advantages. For instance, it forces the analyst to be clear about what constitutes a particular category of discourse presentation. As far as possible, we tagged on the basis of linguistic form (e.g. indirect discourse presentation always involves two clauses, while the NR{S/W/T}A category was used for one clause structures; we discuss this in greater detail below), though we recognised that context often plays a role in determining a particular structure (e.g. free indirect forms). All the texts in the corpus were tagged initially by one of us and then checked by the other and revised in the light of our discussions. A further advantage of this approach is that as our tagging progressed, we were able to revise decisions made earlier in the project on the basis of our increasing experience of identifying the various discourse presentation structures. Annotating also enables the retrieval of problematic structures, for discussion at a later date – for example, ambiguous cases.

In Table 5 we give an example from our corpus, where one exists, of each of the categories set out in Table 1. In each example the DP is underlined. Following this, we discuss the criteria we employed in assigning stretches of text to particular discourse presentation categories.

D. McINTYRE & B. WALKER – ANNOTATING A CORPUS

Speech presentation		Writing presentation		Thought presentation	
(F)DS	Whan he did com home to his house his wife sayd, <u>where is my Brandiron or trefete.</u>	(F)DW	he began and wrote again -- <u>'Be mine, with all your poverty.'</u>	(F)DT	<u>'Very well,'</u> thought I;
FIS	the rogues presented each a pistol to them, and bid them deliver, <u>or they would blow the brains out of their heads;</u>	FIW	No occurrence in the corpus	FIT	but Dedalus finding, he could not build his determinations upon these uncertainties, <u>wherein both the safety of the Towne and his own honour, might probably suffer, by reason of the protraction as also the person of the princesse Cloria be endangered by his slownesse and neglect.</u>
IS	the Princesse told her, <u>that she had beene lately troubled with a most untoward and fearfull dreame,</u>	IW	Middleton also writes to them out of Holland, <u>that Colonel Dezmond was shipped away ...</u>	IT	and he shou'd have been entirely comforted, but for the Thought <u>that she was possess'd by his Grand-father.</u>
NRSA	the wynde also began to blow agayne: wherfore we were glad and lauded and <u>thanked god</u>	NRWA	and comytted unto hym the same by Instruccyon <u>sygned and subscrybed with his owne hande</u>	NRTA	All this, you may be sure, <u>was as I wished,</u>
				NI	his arguments and divisions being so many, <u>that they caused a great confusion in his brain</u>
NV	<u>My lord then made another and a longer speech of the same sort</u>	NW	The late Parliament having upon their dissolution delivered up the Power which they received from his Excellency at their first sitting, <u>by a Writing under their Hands and Seal</u>	NT	<u>filled her imagination with some unprofitable thoughts</u>

Table 5 EModE corpus examples of SW&TP categories

Wherever possible, we used linguistic form to guide our tagging. Having clear criteria for deciding between one DP category and another was particularly important for direct and indirect discourse forms. The example for (F)DS in Table 5 shows that quotation marks were not always used to mark direct speech. However, the example is clearly one of direct speech because there is (i) a reporting clause that introduces the speech; (ii) a shift to present tense; and (iii) a shift in deixis that is appropriate to the original speaker, marked by the pronoun. Indirect discourse consists of a reporting clause and a subordinate reported clause, along with a corresponding back-shift in tense. The important criteria here is that the reported discourse must be in a separate clause, which can be finite or non-finite. Table 5 shows a prototypical example involving a reporting clause and a subordinate reported clause signaled by the subordinating conjunction *that*. Non-prototypical but fairly common forms are those where the subordinating conjunction is elided, for example:

<dptag cat="NRT"> for now she fear'd </dptag>
<dptag cat="xIT"> the Storm wou'd fall on the Prince; </dptag>

Additionally, the reported clause can also be marked by an infinitive verb (underlined below), for example:

<dptag cat="NRS"> I chargde them </dptag>
<dptag cat="xIS"> to stay and watch the house belowe, </dptag>

Without these criteria it can sometimes be difficult to distinguish between indirect discourse and Narrator's Report of Speech, Writing or Thought act (NR{S/W/T}A), particularly when there is a topic specified, as the following two examples demonstrate:

<dptag cat="N"> and hearing some one sighing in the other Room, she pass'd on, and found the Prince in that deplorable Condition, </dptag>
<dptag cat="NRTAp"> which she thought needed her Aid: </dptag>

<dptag cat="NRSAp"> when presently I demaunded of this Leiuetenant the place of his abode, and when hee last heard of him </dptag>

The above examples demonstrate the use of a DP sub-category *p* to indicate topic. Using the formal criteria described above helps to

distinguish between propositional content and topic, which can sometimes be problematic.

While we endeavoured to tag on form, some DP categories, particularly free indirect examples, also require consideration of the wider context. The examples of FIS and FIT in Table 5 (no examples of FIW occur in our corpus) demonstrate this. The FIS has tense and pronouns appropriate to an indirect form, but there is no reporting clause introducing the discourse; hence this was distinguished on formal grounds. The FIT, however, is a more difficult case. We tagged this as free indirect thought because the preceding clauses introduce Dedalus' internal state and indicate his thought process. Consequently, we decided that the underlined section of the example is not simply narration but relates to the propositional content of his thoughts, containing some flavour of the original discourse.

5. Results

Analysis of the corpus is ongoing and here we present some of our initial quantitative findings, along with some qualitative analysis, in order to demonstrate the usefulness of the kind of annotation we have undertaken.

Figure 1 shows the overall distribution of speech, writing and thought presentation in both the EModE and PDE data:

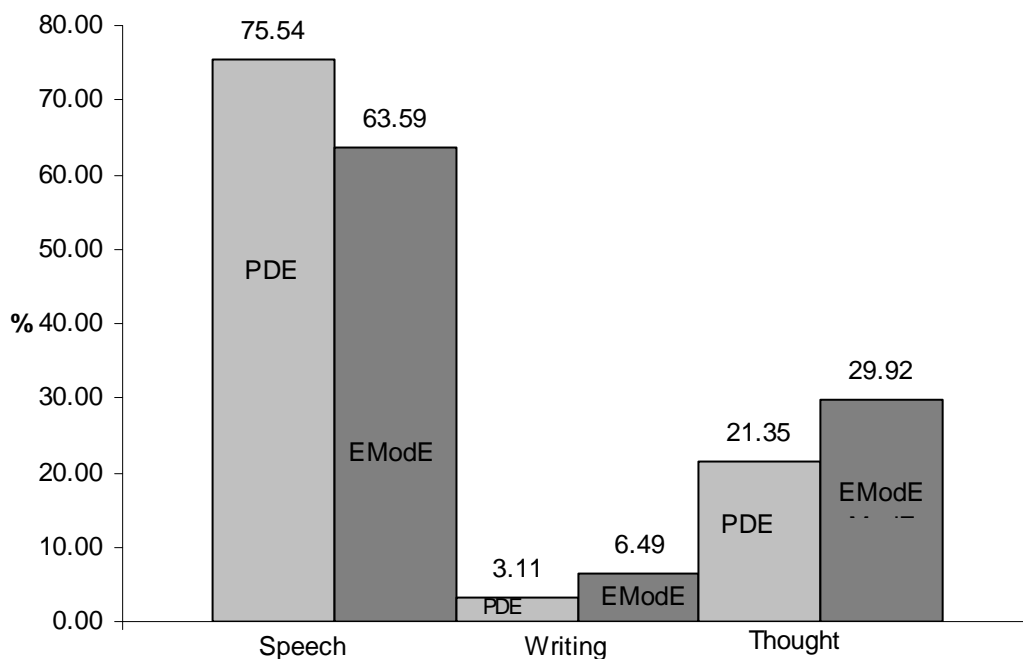


Figure 1 Overall distribution of speech, writing and thought presentation in the EModE and PDE data

What we can observe from this graph is that the overall distribution of discourse presentation in the EModE data follows that of the PDE corpus; that is, speech presentation dominates, followed by thought and writing presentation. While the histogram suggest that there is more thought and writing presentation in the EModE data than in PDE, log-likelihood tests show that these are not significant differences. Initially then, the distribution of SW&TP in EModE is the same as for PDE. However, we begin to see differences when we consider the distribution of individual categories on the speech, writing and thought presentation clines. Tables 6, 7 and 8 show, respectively, the frequency of categories of speech, writing and thought presentation in the corpus compared against the PDE data. The percentages are based on number of tags as opposed to number of words within each category. We provide a fuller discussion of these results in McIntyre and Walker (forthcoming) and here concentrate particularly on what the results reveal about writing presentation. Log-likelihood (LL) figures in bold indicate statistically significant differences between the two corpora.

Category	PDE				EModE				LL
	No. of tags	% of total	% of cline	Rank	Tag Freq	% of all DP	% of cline	Rank	
(F)DS	2339	40.38	53.45	1	275	27.89	43.87	1	37.12
FIS	90	1.55	2.06	5	11	1.12	1.75	4	1.17
IS	784	13.53	17.92	3	120	12.17	19.14	2	1.20
NRSA	918	15.85	20.98	2	109	11.05	17.38	4	13.89
NV	245	4.23	5.60	4	112	11.36	17.86	3	64.73
Totals	4376	75.54	100.00		627	63.59	100.00		5.79

Table 6 Frequencies of instances of speech presentation (number of speech tags) and rank orderings

Category	PDE				EModE Corpus				LL
	No. of tags	% of all DP	% of cline	Rank	Tag Freq	% of all DP	% of cline	Rank	
(F)DW	43	0.74	23.89	2	17	1.72	26.56	2	6.96
FIW	12	0.21	6.67	5	0	0.00	0.00	5	6.96
IW	30	0.52	16.67	3	36	2.64	40.63	1	19.68
NRWA	82	1.42	45.56	1	10	1.01	15.63	4	10.38
NW	13	0.22	7.22	4	11	1.12	17.19	3	5.26
Totals	180	3.11	100.00		74	6.49	100.00		0.59

Table 7 Frequencies of instances of writing presentation (number of writing tags) and rank orderings

Category	PDE				EModE Corpus				LL
	No. of tags	% of all DP	% of cline	Rank	Tag Freq	% of all DP	% of cline	Rank	
(F)DT	84	1.45	6.79	4	5	0.51	1.70	5	15.00
FIT	230	3.97	18.59	2	1	0.10	0.34	6	51.08
IT	119	2.05	9.62	3	66	6.69	22.37	2	45.13
NRTA	71	1.23	5.74	5	39	3.96	13.22	3	28.39
NT					11	1.12	3.73	4	4.38
NI	733	12.65	59.26	1	173	17.55	58.64	1	1.69
Totals	1237	21.35	100.00		295	29.92	100.00		12.34

Table 8 Frequencies of instances of thought presentation (number of thought tags) and rank orderings

If we focus on writing presentation in EModE, it is clear that the foregrounded category is Indirect Writing. This is statistically over-used in the EModE data when compared to the PDE data. The beginnings of one potential explanation for this can be found in the fact that the majority of the Indirect Writing presentation in the EModE corpus (34 out of 36 examples) occurs in the news journalism data. Below is a concordance of all the instances of indirect writing presentation in this sub-section of the corpus:

itch; the substance of which was, <IW> that the charge of filling up, the fixing of posts remarkable, they write from thence, <IW> that his majesty one day took three wild boars, e for the said county, threatening, <IW> that in case he proceeded any farther in taxing 'd on the spot. They write from Lynn, <IW> that on Sunday se'nnight they had such a viole Marquis de Monti has lately wrote to the magistrates of Dantzick, <IW> that they may soon the captain of which vessel reports, <IW> that two Maltese men of war have taken the Adm We have like wise advice from Genoa, <IW> that a ship belonging to Majorca is arrived in he has, as they write from Vienna, <IW> settled the succession own for good of the publick, and the honour of that mighty empire, <IW> he has, as they wound. We have an account, <IW> that one Mons. Munier, who has lived in England the pril next ensuing: They also tell us, <IW> that the Reform of the troops, which was actua Our letters from Tournay, of the 30th past, say, <IW> that an arrest of the council state Our letters this day from Brussels say, <IW> that the burgers, who have fled from their h Our Advices from Copenhagen say <IW> they were busy there fitting out a squadron of men We have advice from Moscow <IW> that his Czarish Majesty had disbanded a great many Our letters from Paris make mention, <IW> as if the Pope, who had been relapsed, were re We have advice from Lubeck, <IW> that 5 ships were lately cast away on the coast of They tell us from Stetin, <IW> that the Governour General Mellin had, by Placaet, We have an account from Lysland, <IW> that they are busy levying a tax there, which is to We have advice from Warsaw, <IW> that, pursuant to the accommodation made with the E

Last Sunday Publication was made throughout the kingdom, <IW> that the Month of February k Dec 24 Our letters from Poland say, <IW> that Prince Alexander Sobietzki, designed to g her her collar. The port-letters say, <IW> that the Mary of London was put into Plymouth, Our accounts from most of the provinces of this kingdom say, <IW> that there's nothing but ordering the landgrave of Hess d'Armstadt <IW> to forbear his hostilities against the Middleton also writes to them out of Holland, <IW> that Colonel Dezmond was shipped away t a Letter to Glencarn , assuring him <IW> that the K. of France , and Denmark , the Duke tters from the Hague , it is written, <IW> that the French Ambassador there, ther in men nor money, desiring him <IW> to be with in what he formerly Promised unto the of their Garrisons, and they wonder <IW> that he is able to send them no aid , neither in unto the sayd Generall, advising him <IW> that the sayd Ile of Lantore did belong unto t brought Advice into Plymouth, <IW> that he had descried the Spanish Armado near the Li Capt. Fleming, who had beene ordered <IW> to cruize in the Chops of the Channell, for Di certen requestes, as he termed them) <IW> to remedye the grieffes of the Devonshirmen,

What is particularly interesting about these examples is the reporting clause that precedes the discourse presentation. In each case, considerable emphasis is placed on identifying the source of the report that follows. Thus, we have clauses such as 'it is written', 'the port-letters say', 'they tell us from Stetin', 'we have advice from Lubeck' and 'they write from Lynn'. It appears, then, that there is a concern among EModE writers of news reports to make clear that the report of news is taken from a identifiable source, rather than being, say, conjecture on the part of the writer. That these reporting clauses should be followed by indirect writing presentation is perhaps explained by the fact that an indirect report allows for the reconstruction of the exact words of the original writer. This seems appropriate when so much emphasis is placed on accounting for the source of the story. In effect, the indirect category makes it clear that the news report is a *representation* of an original source, as opposed to, say, a summary report. It does this by presenting the original writer's words in an alternative format but one which also allows the reader recourse to the words and structures of the original discourse. Conboy (2007) makes the point that EModE journalism relied heavily on written reports, though he makes no suggestion as to why these reports were presented in the forms that they were:

Throughout the eighteenth century, news often comprised the contents of letters received, conveying both opinion and information, and the language reflected the letter-writing style of the time.

We have a report here, but we hope without foundation, that his Majesty's frigate Minerva was not lost on the back of the Isle of Wight on Friday last night last, when it really blew a hurricane. (*London Evening Post*, 31 December to 3 January 1764)

Newspapers depended on such reports for their own content, together with letters from readers to fill their pages. Communication and distribution technologies available at the time meant that maintaining a regular flow of news was a problem. It meant that the language of the reports which were in regular supply could be more elaborate.

(Conboy 2007: 6-7)

Conboy's chosen example is similar in structural terms to those in our corpus in that the reporting clause identifies the source of the report which is then presented in an indirect form. However, when it comes to explaining this kind of structure, Conboy seems to be suggesting that the style of newspaper reports was in part due to a need to fill up space. This seems counter-intuitive, since an easier way to fill up space than establishing a more long-winded written style would have been to print in larger type. We suggest another possibility; that indirect presentational forms are more dominant in the EModE data because of a desire to be seen to *represent* the news. This would accord with the relative absence of freedom of the press in this period, and the necessity of avoiding overtly critical comment in news writing (taking care to indicate that a report is based on information in another source, and presenting that source in such a way that the original discourse is recoverable, is one way of implicitly claiming no responsibility for the content of the report; it is also noteworthy that most of the news report is anonymous). The systematic annotation and analysis of both reported and reporting clauses allows us to note this as a pattern in the news reporting of the time. We can also note that the quantitative norm of NRWA (Narrator's Representation of Writing) for Present Day English has the function of summarising more than reporting. Indirect Writing, on the other hand, is the closest we can get to the original discourse while still allowing the reporting of this from a different viewpoint. This, we suggest, may be indicative of the developing nature of the news report genre from report to summary, in effect, a move towards the narrator end of the discourse presentation cline. We might further speculate that the use of IW in EModE news continues in a written form the word of mouth tradition from which the transmission of news grew. More research would be needed to validate these hypotheses.

6. Conclusion

This brief article is intended as a record of the decisions we made during the tagging of our pilot corpus, and as an example of what can be gained through stylistic annotation. Our pilot investigation has already generated a number of hypotheses which might be tested further in a larger project. Stylistic annotation thus offers the possibility of moving us closer towards being able to make generalisations about stylistic development over time.

References

- Baynham, M. and Slembrouck, S. (1999) 'Speech representation and institutional discourse', *Text* 19(4): 439-57.
- Clark, H. H. and Gerrig, R. J. (1990) 'Quotations as demonstrations', *Language* 66: 764-805.
- Conboy, M. (2007) *The Language of the News*. London: Routledge.
- Culpeper, J. and Kytö, M. (2006) "'Good, good indeed, the best that ere I heard": exploring lexical repetitions in the Corpus of English Dialogues, 1560-1760', in Taavitsainen, I., Härmä, J. and Korhonen, J. (eds) *Dialogic Language Use / Dimensions du Dialogisme / Dialogischer Sprachgebrauch* (Mémoires de la Société Néophilologique de Helsinki, Vol. 66), pp. 69-85. Helsinki: Société Néophilologique.
- Culpeper, J. and Kytö, M. (2002) 'Lexical bundles in Early Modern English: a window into the speech-related language of the past', in Fanego, T., Méndez-Naya, B. and Seoane, E. (eds) *Sounds, Words, Texts, Change. Selected Papers from the Eleventh International Conference on English Historical Linguistics (11 ICEHL)*, pp. 45-63. Amsterdam: John Benjamins.
- Holt, E. (1999) 'Just gassing: an analysis of direct reported speech in a conversation between employees of a gas supply company', *Text* 19(4): 505-37.
- Holt, L. and Clift, R. (eds) (2006) *Reporting Talk: Reported Speech in Interaction*. Cambridge: Cambridge University Press.
- Jucker, A. H. (2006) "'but 'tis believed that...": speech and thought presentation in Early English newspapers', in Brownlees, N. (ed.) *News Discourse in Early Modern Britain. Selected Papers of CHINED 2004*, pp. 105-25. Bern: Peter Lang.
- Leech, G. and Short, M. (1981) *Style in Fiction*. London: Longman.
- Leech, G. and Short, M. (2007) *Style in Fiction*. 2nd edn. London: Pearson Education.
- Mahlberg, M. (2007) 'Clusters, key clusters and local textual functions in Dickens', *Corpora* 2(1): 1-31.
- McIntyre, D., Bellard-Thomson, C., Heywood, J., McEnery, A., Semino, E. and Short, M. (2004) 'Investigating the presentation of speech, writing and thought in spoken British English: a corpus-based approach' *ICAME Journal* 28: 49-76.

- McIntyre, D. and Walker, B. (forthcoming) 'Discourse presentation in Early Modern English writing: a preliminary corpus-based analysis.'
- Moore, C. (2002) 'Reporting direct speech in Early Modern slander depositions', in Minkova, D. and Stockwell, R. (eds) *Studies in the History of the English Language: A Millennial Perspective*, pp. 399-416. Berlin: Mouton de Gruyter.
- Myers, G. (1999) 'Unspoken speech: hypothetical reported discourse and the rhetoric of everyday talk', *Text* 19(4): 571-90.
- O'Halloran, K. A. (2007a) 'The subconscious in James Joyce's 'Eveline': a corpus stylistic analysis which chews on the 'Fish hook'', *Language and Literature* 16(3): 227-44.
- O'Halloran, K. A. (2007b) 'Corpus-assisted literary evaluation', *Corpora* 2(1): 33-63.
- Ravotas, D. and Berkenkotter, C. (1998) 'Voices in the text: the uses of reported speech in a psychotherapist's notes and initial assessments', *Text* 18(2): 211-39.
- Semino, E. and Short, M. (2004) *Corpus Stylistics: Speech, Writing and Thought Presentation in a Corpus of English Writing*. London: Routledge.
- Short, M. (2007) 'Thought Presentation 25 Years on', *Style* 41(2): 227-57.
- Studer, P. (2008) *Historical Corpus Stylistics: Media, Technology and Change*. London: Continuum.
- Watts R. and Trudgill P. eds. (2002), *Alternative Histories of English*. London: Routledge.

Laure Gardelle
Université Jean Moulin Lyon 3 – CEL, EA 1663

Que faut-il entendre par « genre grammatical » en anglais moderne ?

Le genre en anglais présente certaines spécificités majeures par rapport aux autres langues qui disposent de cette catégorie grammaticale ; or ces spécificités posent un problème théorique important en linguistique générale, et il s'agit ici de voir comment il pourrait être résolu.

Certes, il existe de grandes disparités dans les systèmes du genre des différentes langues, qu'il s'agisse de l'anglais ou d'autres langues. Les plus évidentes sont d'ordre syntaxique : les éléments susceptibles de porter une marque de genre ne sont pas toujours les mêmes. Même le nom, qui joue pourtant un rôle essentiel, ne porte pas toujours une telle marque ; Whorf oppose ainsi les systèmes patents du genre, dans lesquels il est possible de deviner le genre d'un nom d'après sa morphologie, aux systèmes latents, dans lesquels ces marques distinctives sont absentes (Corbett 1991 : 202). Le plus souvent, il s'agit en réalité de systèmes *plutôt* patents, comme en allemand – où la voyelle finale indique souvent le genre du nom –, ou *plutôt* latents, comme en anglais, où les noms porteurs de marques de genre, tels *woman* ou *actress*, sont extrêmement peu nombreux. Les différences entre les langues concernent également d'autres parties du discours. En français, par exemple, le genre se manifeste surtout dans les déterminants, les adjectifs, les pronoms – personnels, possessifs ou démonstratifs – et parfois les participes passés, tandis qu'en anglais, seuls certains pronoms personnels, relatifs et interrogatifs ont cette propriété. En russe, le verbe peut lui aussi indiquer le genre ; ainsi *ležal* (« était posé ») suppose un sujet masculin, tandis que *ležala* implique un nom féminin dans cette position (ibid. 110).

Outre ces variations syntaxiques, il existe des divergences afférant au système même du genre. D'une part, le nombre de sous-classes qu'il établit varie d'une langue à l'autre ; surtout, ces genres ne peuvent pas toujours être désignés au moyen des étiquettes *masculin*, *féminin* ou *neutre*. C'est là le signe de différences sur le plan même de la catégorisation. Pour certaines langues, les trois étiquettes issues de la tradition gréco-latine peuvent être conservées ; c'est le cas par exemple pour le zande, langue de la branche Niger-Congo – s'y ajoute un quatrième genre, qualifié d'animal. Mais ces étiquettes ne sont possibles que parce que les critères de classification sont assez proches de ceux que nous connaissons : le zande réserve le masculin aux noms d'hommes, le féminin aux noms de femmes, et applique le neutre au reste du monde – la quatrième classe regroupe les noms d'animaux (ibid. 14). Parfois, les critères de classification sont trop différents pour que les appellations *masculin*, *féminin* et *neutre* puissent être appliquées. Ainsi, en dyirbal traditionnel – langue aborigène d'Australie –, une même classe regroupe les noms d'hommes et ceux d'animés non humains ; c'est pourquoi Dixon la qualifie seulement de genre I. La classe dite II regroupe les noms qui font référence aux femmes, mais aussi à l'eau, au feu et aux combats ; elle ne peut donc porter simplement l'étiquette de féminin. Le genre III s'applique aux noms qui désignent tout aliment qui n'est pas fait de chair ; seule la classe IV regroupe tous les autres noms et pourrait éventuellement être qualifiée de neutre, puisque, étymologiquement, le terme signifie *ni l'un ni l'autre* (ibid. 16). Par ailleurs, même deux systèmes décrits au moyen d'étiquettes identiques, ainsi *masculin*, *féminin* et *neutre*, peuvent s'avérer extrêmement différents ; il existe peu de points communs au français et à l'anglais moderne par exemple, hormis pour les noms qui désignent des personnes.

En effet, anglais et français sont emblématiques des deux grands types de classification des noms et référents nominaux. Dans le premier type, auquel s'apparente le système de l'anglais moderne, la sélection du genre repose sur des critères essentiellement sémantiques – ou du moins à dominante sémantique. En d'autres termes, c'est le contenu notionnel du nom ou l'image mentale du référent qui vient déterminer le genre utilisé. Il y a alors une forte correspondance entre les plans linguistique et extralinguistique. La perception du monde ayant une certaine influence sur le choix du genre, les croyances et autres composantes culturelles peuvent jouer un rôle important. En

telegu par exemple, langue parlée dans le sud de l'Inde, les locuteurs établissent une distinction entre les êtres humains (associés respectivement au masculin et au féminin) et le reste du monde (associé au neutre), mais les noms d'êtres divins sont assignés au même genre que les noms d'humains ; ainsi, le nom du Gange, *ganga*, est-il féminin (ibid. 10). A ces systèmes sémantiques viennent s'opposer les systèmes dits formels, bien plus fréquents, parmi lesquels se trouve le français. Dans ce second type, l'assignation à un genre n'est pas fonction du sens du nom, mais de son signifiant. Elle dépend notamment de critères morphologiques ; l'étude de Tucker et al. (1977) sur le français, par exemple, a permis de conclure que les composés dont le premier élément était un verbe étaient pour la plupart masculins (cité par Corbett 1991 : 58). Des critères phonologiques peuvent également entrer en considération : toujours en français, la même étude a montré que 90% des noms se terminant par le son /z/ étaient féminins (type épouse, chemise, etc.). Au premier abord, les deux grands types de systèmes ne semblent guère présenter de points communs : le premier présente des catégories motivées sémantiquement, tandis que dans le second, l'assignation est avant tout arbitraire ; l'accord dans les systèmes sémantiques fait largement appel à des critères extralinguistiques, alors que dans les systèmes formels, c'est un phénomène avant tout linguistique, grammatical. En réalité, ils ne sont pas totalement étrangers l'un à l'autre : comme le souligne Corbett, aucun système n'est totalement formel, c'est-à-dire n'ignore totalement le rapport à l'extralinguistique ; le critère sémantique intervient souvent pour les noms d'êtres humains, voire plus généralement pour les êtres sexués¹. Ainsi en français, *roi* est masculin, *reine* est féminin. Mais dans les langues à systèmes formels, la correspondance entre plans linguistique et extralinguistique ne concerne qu'une fraction des noms.

¹Dans certaines langues dont le système du genre est essentiellement formel, le rapport à l'extralinguistique l'emporte parfois même sur le critère morphologique pour les êtres humains ; la langue compte alors quelques noms dits hybrides. Un exemple typique est le nom allemand de la jeune fille, *Mädchen*, assigné au neutre par son suffixe en *-chen* alors qu'il désigne un être de sexe féminin. L'accord se fait avec le neutre au sein du groupe nominal – l'article défini, par exemple, est obligatoirement *das* –, mais au-delà, le pronom personnel privilégié est *sie*, pronom féminin. L'accord est donc formel dans le premier cas, mais sémantique dans le second. En russe, c'est directement en langue que le nom de l'oncle, *djadja*, qui aurait dû relever par sa forme de la deuxième déclinaison, et donc être féminin, est assigné au masculin ; il en va de même en latin avec le nom du paysan par exemple, *agricola* (Corbett 37). Il ne s'agit cependant que de cas isolés.

Si la plupart de ces disparités entre les systèmes du genre des différentes langues ne sont que diverses manifestations d'une même catégorie grammaticale, certaines spécificités de l'anglais moderne remettent en cause la définition même qu'en donne la linguistique générale – à tel point que certains linguistes ont rejeté l'idée d'un système grammatical du genre dans cette langue. Examinons d'abord leurs arguments, afin de voir ensuite dans quelle mesure ces spécificités doivent réellement faire de l'anglais une langue à part dans le paysage linguistique.

I. Définition générale du genre et problèmes posés par l'anglais

Le genre est défini dans les encyclopédies de linguistique (voir notamment Trask 1993 : 12) comme un système de classification des noms, par lequel les noms d'une langue sont regroupés en un petit nombre de classes – deux dans le cas du français, trois en anglais par exemple. L'étymologie même du mot *genre* fait écho à cette définition, puisqu'il vient du latin *genus*, qui signifie *sorte*. Pour la majorité des linguistes, le genre se caractérise par le phénomène de l'accord, c'est-à-dire qu'il y a covariance, sur un plan syntaxique, soit entre les divers éléments d'un constituant qui portent une marque de genre, soit entre plusieurs constituants grammaticalement reliés. Ainsi en français, au sein du syntagme nominal *un banquet somptueux*, l'article indéfini et l'adjectif s'accordent au masculin parce que le substantif qui les régit est masculin ; au contraire, avec le nom *fête*, l'accord aurait lieu au féminin : *une fête somptueuse*. Comme le montre cet exemple, c'est toujours le nom qui est présenté comme la source des marques de genre présentes dans le reste de l'énoncé.

Cette définition traditionnelle, qui s'applique bien au latin ou au français, ainsi qu'à de nombreuses autres langues non apparentées², pose problème pour l'anglais moderne sur ses deux caractéristiques

² Le genre n'est pas seulement présent dans les langues issues de l'indo-européen. Bien que de nombreuses langues ne disposent pas du genre – notamment en Asie, où ce sont les systèmes de classificateurs qui sont les plus représentés, ou encore en basque ou dans les langues turques –, la catégorie se retrouve dans toutes les régions du globe, qu'il s'agisse des langues du nord du Caucase, de celles de la famille dravidienne en Inde du sud, de la famille afro-asiatique, des langues de Nouvelle-Guinée et d'Australie, ou encore, sur le continent américain, des langues de la famille algonquin (Corbett 1991 : 2).

principales : le principe de l'accord et celui d'une classification des noms en langue.

La première objection soulevée par Nesfield (1924 : 12), Erades (1956)³ ou encore Creissels (1995 : 87) est d'ordre morphologique : Erades (1956) par exemple note que le genre, même s'il n'est pas marqué morphologiquement dans le nom, doit apparaître au moins au sein du syntagme nominal – déterminant ou adjectif –, car c'est là qu'il prend sa source⁴. Or en anglais moderne (l'anglais du XVI^e siècle à nos jours), ni le déterminant, ni l'adjectif ne marquent le genre, et ce alors qu'en vieil-anglais (l'anglais employé jusqu'au milieu du XI^e siècle environ), ce principe de l'accord au sein du syntagme nominal était respecté : le fonctionnement du genre était plus ou moins analogue à celui du français aujourd'hui. Erades conclut donc qu'il n'y a plus de système du genre en anglais moderne – il n'y aurait plus que les résidus d'une catégorie en déroute.

Un second point apparaît encore plus problématique : il est fréquent qu'en anglais moderne, un même référent, chez un même locuteur, soit tour à tour désigné par des pronoms de genres différents. C'est le cas notamment pour les animaux. Dans un compte rendu de pêche, Jon Callan (2003) – britannique – écrit par exemple d'un même référent :

1. (barbeau) We carefully checked her over and there wasn't mar on her at all, it looked like it had just come out of mint. / We measured her at 31 ½ inches with an 18 inch Girth, so the potential of her growing some more yet is a very real possibility.

Il en va de même dans les autres variétés d'anglais ; citons un extrait d'un ouvrage de Vance Packard (US [1950]) :

2. (47. énoncé spécifique) The journal ran a series of photographs showing how this ingenious rat solved a problem that bothered him. The rat's private shelter was separated from the hot, exposed area where it had to eat its meals by a foot-high plank. Its meal was always served in a small bowl.

³ Cité par A. Joly (1975 : 205).

⁴ De même, Nesfield conclut : « In the proper sense of the word, this is not gender at all, since it is not based on the form of the word. 'Gender' appears only in *he, she, it*. » (12). Creissels conclut également qu'il est impossible de définir des classes d'accord en anglais car le genre consiste avant tout en un classement des formes substantivales et des formes qui peuvent leur servir de déterminant (87).

Ce constat vaut également en énoncé générique – ici un extrait d'une histoire naturelle des vertébrés (Willson [US, 1984]) :

3a. (207) Figure 5-19. A male *Pterocles* sand grouse. Its modified belly feathers enable it to transport water to its young

3b. (ibid.) Sand grouse (Fig. 5-19) fly many kilometres from the nesting area to reach water holes. The male rumples his breast feathers against the ground and water then is easily taken up when he wades into the shallows to drink.

Ces variations de genre, si elles sont anecdotiques pour les humains et les inanimés, sont relativement fréquentes dans les références aux animaux : dans un corpus test de 530 occurrences, outre les cas de variations d'un passage à un autre d'un même ouvrage, on compte 11 cas de variations d'un paragraphe à un autre, 11 cas de variations d'une phrase à la suivante, et 6 de variations au sein d'une même phrase (pour plus de détails, voir Gardelle [2006]). Ces variations de genre étaient d'ailleurs déjà fréquentes aux débuts de l'anglais moderne (XVI^e-XVII^e siècles), en particulier pour les animaux ; on lit par exemple dans l'histoire naturelle de Thomas Mouffet (1658 : 1070) :

4. (araignée) Amongst the Net-workers I saw one of the greatest of all, I have set down the picture of it here. In Autumn amongst small Rose-boughs it extendeth an artificiall Net, and it catcheth either another Spider running over it, or Gnats or Flies that come to it, when she pulls her cord with wonderful dexterity, and when she hath hanged them thus up, she leaves them till she growes hungry again.

L'existence de telles variations de genre, ainsi que la fréquence de ces variations pour un certain nombre de référents, pose le problème de la définition du genre en anglais, catégorie qui, selon la linguistique générale, établit des classes de noms. Erades résume bien cette objection : « Can we speak of gender in a language where the same noun may at one moment be masculine, at another feminine or neuter, and, let us mark it well, in the language of the same speaker and sometimes in one and the same sentence ? » (1956 : 9). Là encore, l'objection a d'autant plus de poids qu'en vieil-anglais, les classes de noms étaient stables. Il semble par ailleurs que cette caractéristique de l'anglais moderne le différencie non seulement des langues à systèmes formels, mais aussi des autres langues à système sémantique du

genre : Corbett ne mentionne pas de telles variations dans d'autres langues.

Pour tenter de résoudre le problème de classification du genre de l'anglais moderne, intéressons-nous à présent plus spécifiquement à son fonctionnement. La question à laquelle il s'agit de répondre est la suivante : si le nom ne régit pas un accord grammatical strict comme en français ou en vieil-anglais, selon quel mécanisme le genre est-il choisi ?

II. Le mécanisme de choix du genre en anglais moderne

Pour répondre, intéressons-nous plus précisément aux cas les plus problématiques : les références aux animaux, et plus particulièrement au choix entre neutre et genres dits animés (masculin et féminin) pour ces référents. Les grammaires proposent différentes hypothèses explicatives, parmi lesquelles le critère du sexe, celui du sexe pertinent, l'animation, le degré de familiarité avec l'animal, la personnalité, la caractérisation ou encore la personnification. Au-delà du manque de consensus dans la description du système du genre, ces critères suggèrent trois pistes possibles : le critère du sexe et celui de l'animation impliquent que la source de choix du genre se trouve dans une propriété objective du référent ; le genre opèrerait donc une classification non pas des noms, mais des référents eux-mêmes. Au contraire, les notions de « sexe pertinent » et de « personnalité » orientent vers la subjectivité de la perception du locuteur en contexte ; dans cette hypothèse, le genre opèrerait donc une classification des référents, mais avec la médiation du locuteur. Enfin, évoquer le degré de familiarité avec l'animal implique de s'intéresser au locuteur uniquement, plus particulièrement à son vécu.

Examinons chacune de ces trois pistes, afin de déterminer laquelle peut vraiment s'appliquer.

1. Les propriétés objectives du référent

Pour explorer cette piste, prenons le cas du critère du sexe, le seul à avoir été retenu dans les grammaires de l'anglais entre le XVII^e siècle,

date des premières descriptions, et le début du XX^e. Le principe d'une corrélation entre genre et sexe apparaît particulièrement séduisant car il correspond à l'intuition des anglophones : les informateurs interrogés décrivent ainsi le système : « we use *he* for males, *she* for females, *it* for things ». D'ailleurs, dans le langage courant, le verbe *neuter* signifie castrer, c'est-à-dire supprimer le sexe.

Si le sexe du référent est le seul critère discriminant dans le choix du genre, on doit pouvoir édifier deux règles : *he* et *she* sont employés pour tout animal dont le sexe est connu, et *it* désigne tout animal dont le sexe est inconnu. C'est d'ailleurs ainsi que le système du genre est expliqué dans les grammaires de l'anglais jusqu'au début du XX^e siècle. En réalité, aucune de ces deux règles ne peut être appliquée à l'anglais. D'une part, *he* et *she* sont loin d'être employés seulement pour des animaux dont le sexe est connu ; une analyse statistique du corpus test montre que ce cas représente à peine 70% des occurrences des deux pronoms – plus précisément 64% des occurrences de *he* et 80% de celles de *she*. Le critère du sexe ne permet pas non plus de rendre compte des variations de genre rencontrées dans les énoncés cités précédemment, ainsi dans l'énoncé 1 cité ci-dessus :

1. (barbeau) We carefully checked her over and there wasn't mar on her at all, it looked like it had just come out of mint. / We measured her at 31 ½ inches with an 18 inch Girth, so the potential of her growing some more yet is a very real possibility.

Il est difficile de postuler que le locuteur reconnaît le sexe du barbeau dans le premier segment, et encore plus qu'il l'ignore dans le second pour s'en souvenir à nouveau dans la phrase suivante ; il faut donc en conclure que *he* et *she* ne sont pas employés qu'en fonction du critère du sexe.

La seconde « règle » – selon laquelle *it* serait employé pour un animal de sexe inconnu – ne fonctionne pas non plus : dans le corpus test, non moins de 19% des occurrences de *it* désignent un référent de sexe connu. L'énoncé 3a cité précédemment le montre :

3a. (207) Figure 5-19. A male *Pterocles* sand grouse. Its modified belly feathers enable it to transport water to its young.

L'auteur choisit le neutre alors que le syntagme nominal source spécifie le sexe (*male*), et alors que plus loin (en 3b) elle emploie le masculin. Le critère du sexe n'est donc pas pertinent dans le choix du

neutre par rapport au masculin ici. L'existence même de variations pour un même référent montre également, plus largement, qu'aucune autre propriété objective des référents ne peut être seule à l'origine du choix du genre ; il faut chercher ailleurs le mécanisme qui le régit. Puisqu'il y a variation en discours, intéressons-nous à l'axe du contexte – pris au sens de co-texte et de situation.

2. L'axe du contexte

Lorsqu'un même locuteur emploie deux genres différents pour un même référent, ni les propriétés intrinsèques de ce référent, ni le vécu du locuteur, ne peuvent avoir influé sur le choix du genre ; seul le contexte évolue.

Mais peut-on pour autant conclure que les facteurs contextuels régissent à eux seuls le choix du genre ? Non, car parfois le contexte n'influe pas sur le genre. Citons un énoncé dans lequel un fauconnier évoque le mode de vie du faucon pèlerin :

5. « He lives in South America... He also eats worms... » (entendu, Royal Armouries [Leeds], 2003).

Le locuteur emploie ici le masculin alors que les données sont totalement factuelles ; dans les mêmes conditions, un ouvrage scientifique emploierait *it*, et les informateurs interrogés voient dans cette utilisation du masculin l'indice que le locuteur travaille dans le monde des oiseaux de proie ; il est possible d'en déduire que dans ce cas, c'est le vécu du locuteur qui a favorisé le choix de *he*.

Il reste donc à présent à déterminer dans quelle mesure ce vécu peut influencer sur le genre.

3. Le vécu du locuteur

Les grammaires qui s'intéressent au vécu du locuteur proposent deux orientations. La première consiste à classer les référents en fonction de leur statut dans la société ; il y a alors influence du vécu du locuteur par le biais des représentations culturelles qu'il a intériorisées. L'autre postule un degré de familiarité du seul locuteur avec le référent.

La première voie est illustrée principalement par les structuralistes, notamment Quirk et al. (1985, p. 68) :

	<u>Classe</u>	<u>Exemple</u>	<u>Pronoms</u>
personnel he/she he/she/it animé who – they	mâle	<i>brother</i>	who - he
	femelle	<i>sister</i>	who - she
	duel	<i>doctor</i>	who -
	commun	<i>baby</i>	- who -
	collectif	<i>family</i>	- which - it which - it /
non-personnel - it / (who) - he - it / (who) - she - it (/he/she)	animal mâle	<i>bull</i>	which
	supérieur anl. femelle sup.	<i>cow</i>	which
	anl. inférieur	<i>ant</i>	which
inanimé	inanimé	<i>box</i>	which - it

Les auteurs établissent ici trois classes au sein des animaux : animal mâle supérieur, animal femelle supérieur et animal inférieur. Ces classes se distinguent par des potentialités d'emploi des genres différentes : pour les animaux inférieurs, masculin et féminin sont annoncés comme très rares, tandis que pour les êtres supérieurs, ils apparaissent plus fréquents. Mais une fois encore, ils ne peuvent indiquer que des tendances d'utilisation ; pour aucune espèce il n'est possible d'exclure l'un des trois genres. Ce constat vaut même pour les animaux de compagnie. Il est vrai, comme le rappellent la majorité des grammaires, qu'ils sont rarement associés au neutre ; le corpus test, par exemple, ne compte aucune occurrence de neutre chez les propriétaires qui évoquent leur propre animal, ou encore chez les vétérinaires qui répondent en ligne au courrier des lecteurs. Mais le seul statut d'animal familier ne suffit pas à imposer le masculin ou le féminin, comme le montrent les énoncés 6a et 6b, extraits d'un même ouvrage de vulgarisation scientifique, *The Human Side of Animals : Animal IQ* (Vance Packard, [US, 1950]) :

6a. (chat. [preface p.Viii-x]) Even Man's best friends, the dog, the cat and the horse, have distinctive personality traits... The cat is far more independent than the dog even though she accepts the support of Man. She is stand-offish, no hand-licker.

6b. (chat. 104) the cat is moderately well up the scale of animal intelligence. It has a fairly adaptable, alert mind.

Il s'agit dans les deux cas du même animal, en énoncé générique. Pour expliquer l'emploi de genres différents dans les deux énoncés, force est de conclure une nouvelle fois à des facteurs contextuels, qui influent sur la perception du référent au moment de l'énonciation.

Ces extraits montrent par ailleurs qu'une simple familiarité d'un locuteur individuel avec un animal ou une espèce (en entendant *familier* comme synonyme de *bien connu*) ne suffit pas non plus à exclure le neutre : le degré de familiarité de Packard avec les chats est nécessairement le même tout au long de l'ouvrage.

Il est possible à présent de conclure sur les facteurs d'influence à l'œuvre dans le choix du genre. L'analyse a montré que les propriétés objectives du référent, les éléments contextuels et le vécu du locuteur pouvaient tous trois orienter le choix du genre, mais à chaque fois *dans certains énoncés* seulement. Cela signifie par conséquent qu'ils ne constituent pas des types de critères fondamentaux pour le choix du genre, mais seulement trois composantes susceptibles d'être au service d'un autre élément. Où se trouve alors cet élément déterminant pour le genre ? Les trois composantes analysées ont pour point commun d'entrer dans la représentation mentale du référent en contexte, c'est-à-dire dans la représentation qu'a le locuteur du référent au moment de l'énonciation. Puisque dans cette représentation, aucune des trois composantes ne peut être retenue pour le genre, il ne reste plus que le point de vue du locuteur ; c'est donc lui qui doit constituer le domaine fondamental pour le choix du genre – les propriétés objectives, le vécu et les éléments contextuels n'influent sur le genre que s'ils motivent ce point de vue. Que faut-il entendre par « point de vue » ? Puisque le vécu du locuteur influe parfois, il semble que ce point de vue consiste plus précisément en un positionnement du locuteur lui-même par rapport à ce référent ; et puisque le contexte ne fait pas nécessairement varier le genre, ce positionnement ne correspond pas à une simple localisation, comme dans le cas des déictiques par exemple ; il s'agit plutôt d'un ressenti vis-à-vis du référent.

Il reste à présent à voir si ces spécificités de l'anglais peuvent être intégrées à la définition du genre à travers les langues. En d'autres termes, peut-on, malgré les objections avancées par certains linguistes, parler de système grammatical du genre en anglais moderne ?

III. Vers une redéfinition du genre

Il semble que l'anglais moderne puisse être ainsi qualifié, à condition que la définition traditionnelle du genre soit légèrement revue.

D'une part, l'étude a montré que l'anglais opérait une classification des référents. Au lieu d'être défini comme un système de classification des noms, le genre pourrait être identifié comme « une catégorisation grammaticale des noms *ou des référents nominaux* », comme le propose Cotte, linguiste angliciste (1999 : 65). Notons que cette nuance ne modifierait pas le rapport entre la catégorie du genre et celles auxquelles elle est toujours opposée – cas, nombre et classificateurs ; elle peut donc être retenue comme trait définitoire.

Par ailleurs, contrairement à des langues telles que le latin, le vieil-anglais ou le français, le genre de l'anglais moderne n'opère pas une classification en langue, mais en discours ; c'est ce qu'ont montré les nombreuses variations de genre pour un même référent – là encore, la nuance ne modifie pas le rapport entre le genre et d'autres catégories, donc elle peut être retenue. Dire que la classification des référents est opérée en discours ne signifie pas, bien sûr, que le genre est absent des catégories nominales en langue. On fait l'hypothèse (Gardelle [2006 : 647]) qu'en langue, une catégorie signifiée par un substantif compte des *potentialités* de genre – non pas nécessairement une association à un seul genre comme en français, mais, selon les noms, à un ou deux genres, voire parfois à trois⁵. Ces potentialités d'association seraient largement liées au vécu du locuteur – qu'il s'agisse de la culture partagée par toute une communauté ou d'expériences individuelles –, au sens où des pronoms et plus généralement le discours fréquemment entendu à propos de référents d'une catégorie, associés au propre

⁵Certains substantifs ne paraissent associés qu'à un genre ; ainsi *Rover* employé pour désigner un chien impose le masculin, le prénom librement donné excluant le neutre (Gardelle 2006, Huddleston & Pullum 2000). Mais contrairement aux langues à système formel telles que le français, un très grand nombre de substantifs, notamment parmi ceux qui désignent des animaux, comptent plusieurs potentialités d'association.

ressenti du locuteur, contribueraient à véhiculer certaines représentations de ces référents. Ces représentations seraient progressivement intériorisées, permettant parallèlement l'acquisition plus globale des invariants du masculin, du féminin et du neutre⁶. Le genre finalement choisi en discours par le locuteur adulte serait celui qui, parmi ces potentialités d'association intériorisées en langue, correspondrait le mieux à la représentation du référent au moment de l'énonciation, en fonction de l'invariant de chacun des genres. Par exemple, un pays n'est jamais associé au masculin ; selon cette hypothèse, au moment de l'énonciation, un locuteur ne sélectionnerait pas le genre le plus approprié parmi les trois genres, mais seulement parmi deux : féminin et neutre.

Dans cette hypothèse, il reste possible de retenir, même pour l'anglais moderne, le principe d'un accord entre genre du pronom et nom : c'est bien le nom qui déclenche des potentialités d'association restreintes. A la différence du français, cependant, cet accord n'est pas formel, syntaxique, mais passe par le biais de la perception du référent en contexte par un locuteur donné.

En conclusion, il semble que la position de linguistes tels qu'Erades soit trop extrême, car la définition de la catégorie du genre à travers les langues est peut-être trop centrée sur le modèle des langues à système formel. L'anglais dispose bien d'un système grammatical du genre, même s'il présente certaines spécificités ; celles-ci pourraient s'expliquer en partie par un degré de grammaticalisation différent, plus élevé pour les langues comme le français. Par ailleurs, les deux types de système ne sont peut-être pas aussi éloignés : même dans les langues à système formel, on note certains noms dits conflits, de genre masculin mais désignant des référents de sexe féminin ; le genre peut alors changer en discours – ainsi *l'écrivain... elle...* Une étude plus précise du fonctionnement de l'anaphore pronominale permettrait de préciser ces mécanismes.

⁶ On sait par exemple qu'un adulte, une fois ses propres représentations intériorisées, est capable néanmoins de noter l'emploi d'un genre inhabituel chez d'autres locuteurs. Des informateurs ont ainsi pu se souvenir avoir entendu tel locuteur employer le masculin pour une plante en pot ou un moustique. On propose donc qu'au stade de l'acquisition du langage, ce processus a lieu, à une plus grande échelle, et participe à la structuration des catégories. Les vécus différents expliqueraient alors certains phénomènes, ainsi le fait que seuls des locuteurs familiers du monde de la pêche, par exemple, envisagent d'employer le féminin pour un poisson.

Bibliographie

- Adamczewski, Henri et Gabilan, Jean-Pierre (1995), *Déchiffrer la grammaire anglaise*. Paris : Didier.
- Attal, Jean-Pierre (1987), *Grammaire et usage de l'anglais*, avec la collaboration de G.S. Tothill et H. Swan. Paris : Duculot.
- Baron, Dennis E. (1986), *Grammar and Gender*. New Haven : Yale UP.
- Callan, Jon (2003) '14 Days in Barbel Heaven'. *Barbel Fishing World*, <<http://www.barbel.co.uk>>, (12/12/2003)
- Christophersen, Paul et Sandved, Arthur O. (1969), *An Advanced English Grammar*. Londres : Macmillan.
- Corbett, Greville G. (1991), *Gender*. Cambridge : Cambridge UP.
- Corblin, Francis (1995), *Les formes de reprise dans le discours : anaphores et chaînes de référence*. Rennes : PU de Rennes.
- Cornish, Francis (1999), *Anaphora, Discourse and Understanding: Evidence from English and French*. Oxford : Clarendon.
- Cotte, Pierre (1998), *L'explication grammaticale de textes anglais*, 2e édition. Paris : PUF.
- , 'Le genre est une métalangue' (1999), *Féminin / Masculin : Littératures et cultures anglo-saxonnes*. Rennes : PU Rennes.
- Creissels, Denis (1995), *Eléments de syntaxe générale*. Paris : PUF, coll. Linguistique Nouvelle.
- Curme, George O. (1931), *A Grammar of the English Language in Three Volumes*, vol. 3 : *Syntax*. Boston : D.C. Heath & Co.
- Gardelle, Laure (2006), *Le genre en anglais moderne (XVI^e siècle à nos jours) : le système des pronoms*, thèse de doctorat (non publiée).
- Gildon, Charles et Brightland, John (1730), *A Grammar of the English Tongue*. Londres: F. Clay, ca.
- Greenbaum, Sydney (1996), *The Oxford English Grammar*. Oxford: Oxford UP.
- Guillaume, Gustave (1992), *Leçons de linguistique*, vol. 12 : 1938-1939, dir. Roch Valin, Walter Hirtle et André Joly. Lille : Université de Laval, p. 58ff.
- Hjelmslev, Louis (1971), *Essais linguistiques*. Paris : Editions de Minuit.
- Huddleston, Rodney et Pullum, Geoffrey K. (2002), *The Cambridge Grammar of the English Language*. Cambridge : Cambridge UP.

- Jespersen, Otto (1953), *A Modern English Grammar on Historical Principles*, vol.7 : *Syntax*. Londres : Allen & Unwin, 1953 [1942].
- Joly, André (1987), « Toward a Theory of Gender in Modern English », *Essais de systématique énonciative*. Lille : PU de Lille, 1987 [1975].
- Kleiber, Georges (2001), *L'anaphore associative*. Paris : PUF.
- Kruisinga, E. (1932), *A Handbook of Present-Day English*, 2e partie : *English Accidence and Syntax*, 5e édition. Groningen : Noordhoff.
- et Paul A. Erades (1960), *An English Grammar*, vol. 1, *Accidence and Syntax*, 8^e édition. Groningen : Noordhoff.
- Lapaire, Jean-Rémi et Rotgé, Wilfrid (1998), *Linguistique et grammaire de l'anglais*. Toulouse : PU du Mirail.
- Larrea, Paul et Rivière, Claude (2005), *Grammaire explicative de l'anglais*, 3^e édition. Paris : Pearson éducation - Longman.
- Leech, Geoffrey et Svartvick, Jan (1994), *A Communicative Grammar of English*, 2^e édition. Londres : Longman.
- Milner, Jean-Claude (1982), *Ordres et raisons de langue*. Paris : Seuil.
- Fernand, Mossé (1959), *Manuel de l'anglais du Moyen-Âge, des origines au XVI^e siècle*, 2^e édition. Paris : Montaigne.
- Mouffet, Thomas (1663), *The Theater of Insects ; or, Lesser Living Creatures. The Whole Revised, Corrected, and Inlarged with the Addition of Two Useful Physical Tables, by John Rowland, MD*. Londres : E. Cote, 1663 [1658].
- Nesfield, J.C. (1957), *Modern English Grammar*. Edition revue et corrigée. Londres : Macmillan, 1957 [1924].
- Packard, Vance (1950), *The Human Side of Animals: Animal IQ*. New York : Pocket.
- Poutsma, Henry (1904-1929), *A Grammar of Late Modern English*. Groningue : Noordhoff.
- Quirk, Randolph et al. (1985), *A Comprehensive Grammar of the English Language*. Londres: Longman.
- Roggero, Jacques (1988), *Grammaire anglaise*. Paris : Nathan-université.
- Swan, Michael (1997), *Practical English Usage*, 2^e édition. Oxford : Oxford UP.
- Sweet, Henry (1892), *A New English Grammar Logical and Historical*. Oxford : Clarendon.
- Tellier, André-R. (1967), *Cours de grammaire anglaise*. Paris : Sedes.

- Trask, Robert Lawrence (1993), *A Dictionary of Grammatical Terms in Linguistics*. Londres : Routledge.
- E.S.C. Weiner, compil. (1983), *The Oxford Guide to English Usage*. Oxford : Clarendon.
- Willson, Mary (1984), *Vertebrate Natural History*. Philadelphie : Saunders College.
- Zandvoort, R.H. (1965), *A Handbook of English Grammar*, 3^e édition. Londres : Longman.

Veronika Koller
Lancaster University (United Kingdom)

Analyzing metaphor and gender in discourse

1. Introduction

Discourse, defined as the use of language as a social practice based on cognition, is recognized as a means to represent reality from a particular viewpoint and to negotiate relationships between text producers and recipients. The present article brings together these two functions of discourse by looking at the role that metaphor plays in constructing reality, a question that is further narrowed down by addressing how such constructions transport notions of gender identity that position discourse participants in particular relations to each other. The links between metaphor and gender are multi-faceted, involving how speakers of different genders use metaphor, how metaphor is employed to describe men and women, and how metaphors can serve to define a social domain in masculine or feminine terms. Sections 2 and 3 address some theoretical and methodological issues that arise when one investigates metaphor and gender in spoken or written discourse, while section 4 briefly presents three case studies that illustrate the different ways in which metaphor use in discourse can be gendered, and gendering.

These latter notions are explained in the following section.

2. Research into language, gender and metaphor

The first part of this section will challenge essentialist approaches to gender and language, which are contrasted with a social constructivist approach that centres on discourse analysis. After this, a brief outline of the history of metaphor research will be given, with a focus on conceptual metaphor theory. The final part of this theory

section provides definitions of discourse and links the concept to gender and metaphor.

2.1. Language and gender research

Overviews on gender and language research (Jule 2008, Litosseliti 2006) usually distinguish three historical approaches, which, although roughly following one another, overlap to some extent and enjoy more or less currency in different contexts. What has come to be known as the deficit approach can be roughly located in the first half of the 20th century and attracted its label because of a view of women's language as an inferior variety. With the advent of second-wave feminism – the first wave being represented by the suffragette movement at the turn of the 19th and 20th century –, much 1970s research into language and gender¹ subscribed to the idea of language and conversation as sites of women's oppression. While Lakoff's (1975/2004) seminal work still relied on intuition and anecdotal examples, conversation analysts such as Zimmermann and West (1975) and Fishman (1983) focused on examples of mixed talk in which men frequently interrupted women and exerted control over topics while female speakers took a supportive role and hedged their statements to be less assertive. Roughly in the 1980s, the so-called dominance approach gradually gave way to what came to be known as the difference approach, which posited that women and men are socialised into different sub-cultures, including language use, with the result that mixed talk amounts to quasi-intercultural miscommunication (Maltz and Borker 1982). Tannen's (1990) work was pivotal in this regard, starting a string of popular books on allegedly irrevocable social, if not innate, gender differences, a deceptively simple and obvious line of thinking that still enjoys widespread allegiance twenty years later. The aim of much academic research in the difference framework was to reevaluate perceived women's speech as cooperative, often taking a relativist approach to real power differences, and their reinforcement in language use, between women and men. Key works are Coates' studies of all-female (1986) and all-male (2003) talk.

¹ Of course, gender has always been, and continues to be, seen as a fixed demographic variable in language and variation studies.

The publication of Judith Butler's landmark book *Gender Trouble* in 1990 ushered in a new era in language and gender research that signalled a radical break with the approaches outlined above. The deficit, dominance and difference views have in common an essentialist belief in gender as an inherent property that is determined by sex. Simply put, males will be masculine in their behaviour and females feminine, and they will show such behaviour irrespective of context or other identity facets. By contrast, what has come to be known as the social constructivist or discourse approach posits that biological sex and gender are not in a cause-effect relationship. Gender identity is seen as an effect of behaviour, including language behaviour, that is culturally associated with masculinity or femininity. Rather than being stable and context-independent, therefore, gender can be enacted differently across situations and times. Such fluid gender identities are co-constructed in discourse, i.e. negotiated through social interaction, including language use. Further, gender identities and their enactment are modelled on a hegemonic ideal (Connell 1995, Connell and Messerschmidt 2005) which speakers attempt to approximate or, less often, avoid. In short, while the three earlier approaches all assumed an interpretative stance of 'You are a man (woman) and therefore you talk like this', the contemporary approach interprets language use in terms of 'You talk like this and that makes you come across as masculine (feminine)'.

Of course, there are limits to how flexibly gendered language can be used, and it is indeed often the case that female or male speakers try to project, through their language use, ideals of femininity or masculinity, respectively. Rather than explaining this purely as an effect of their sex and concomitant socialisation, however, the discourse approach to language and gender allows for the possibility that speakers react to perceived expectations of 'ideal' gendered behaviour and that they may at times even do so consciously and strategically. It needs to be remembered that there are strong social pressures in place to patrol 'appropriate' language and other behaviour, ranging from ridicule and challenges to threats and physical violence. Yet, one corollary of the discourse approach is that masculinity and femininity are linguistic resources that speakers of any sex can, albeit within different limits, use to project gender identities that they deem appropriate in the context, or to capitalize on positive cultural associations of particular genders. For example, the

works of Baxter (2003, 2010), Holmes (2006) and Mullany (2007) have shown how, depending on context and institutional norms, female professionals use language in stereotypically masculine ways, e.g. in order to be perceived as assertive and competent. Likewise, male professionals deploy what is widely regarded as feminine language features and conversational styles in order to present themselves as caring leaders and good ‘people persons’. (This evidence sits alongside that of stereotypical and hybrid language behaviour.) In section 4, examples of gendered speech behaviour in relation to metaphor use will be provided, which further call into question any deterministic link between sex and gender.

After this brief outline of language and gender research, the next sub-section will turn to metaphor.

2.2. Metaphor research

The word ‘metaphor’ derives from the Greek μεταφέρειν (‘to carry across’) and it is therefore no surprise to find one of the earliest treatments in Western philosophy in the works of Aristotle (ca. 350BC). Defining it as ‘an application of an alien name by transference’ in his *Poetics*, he further elaborated in his *Rhetoric* that a metaphor had to ‘correspond to the thing signified’. For Aristotle, then, metaphor was less the interplay between different domains than what was later called the source domain, in terms of which something else is defined. While he commented, on the use of metaphor in literature, that ‘a command of metaphor ... is the mark of genius, for to make good [new] metaphors implies an eye for resemblances’ (*Poetics*), he also, in his *Rhetoric*, acknowledged the ubiquity of metaphor, noting that ‘the proper or regular and the metaphorical ... are used by everybody in conversation’, and also pointed out its pedagogical value (see also Ortony 1975, Cameron 2003). Aristotle thus prefigured the modern view of metaphor as ubiquitous, but differed in that he regarded it as a purely linguistic, rather than a cognitive, phenomenon.

In modern times, I.A. Richards’ *Philosophy of Rhetoric* (1936) is widely cited as an important contribution to metaphor theory. Writing in the Romantic tradition, he advanced that imagination and metaphor shape our perception of the world and clearly took a cognitive view of

metaphor when saying that ‘thought is metaphoric ... and the metaphors of language derive therefrom’. Richards further distinguished between the tenor (target domain), vehicle (source domain) and ground of a metaphor; in the clichéd gendered metaphor WOMEN ARE FLOWERS, for example, women are the vehicle, i.e. that which is conceptualized in terms of something else, and flowers are the tenor, i.e. the terms in which the vehicle is thought of, while the ground for comparison is semantic features such as fragile or beautiful. Importantly, Richards prefigured the theory of conceptual blending (Fauconnier and Turner 2002, Grady et al. 1999) and its notion of emergent structure (see below) by claiming that tenor and vehicle show similarity but can be of varying importance, together forming more than the sum of their parts. Moving on to the latter half of the 20th century, Max Black (1962, 1977) not only took up the idea, developed in antiquity by Cicero (55BC), that metaphor helps to fill lexical gaps and express the otherwise inexpressible (see Ortony 1975). More importantly, he advanced the interaction theory of metaphor, which says that metaphor producers not only draw on knowledge they share with recipients, but also help to shape the recipients’ understanding of both secondary and primary subject, the latter being Black’s terms for source and target domain. Because metaphor creates rather than reflects similarity, it is based on knowledge but also produces new knowledge, an idea which anticipates conceptual metaphor theory and blending theory.

Lakoff and Johnson’s seminal work *Metaphors We Live By* (1980) first spelled out what has since become the standard in linguistic metaphor research, namely the conceptual theory of metaphor. The central idea is that metaphor is the cognitive process of ‘understanding and experiencing one kind of thing in terms of another’ (p.5). Thus, metaphor is primarily a cognitive phenomenon which is realised in language and other semiotic modes, e.g. images, as metaphoric expressions. As such, it is ubiquitous and a universal feature of human cognition. Metaphor is formed, according to this theory, by mapping semantic features from a source to a target domain, thereby conceptualizing one in terms of the other. Source and target domain are each a structured set of elements, their properties and relations, thus forming an idealised cognitive model (Lakoff 1987). Moreover, mapping from source to target domain is selective, highlighting some features and hiding others, often for ideological purposes (see Ungerer

2000 on ‘muted metaphors’). Thus, the WAR metaphor so prevalent in business discourse (Koller 2004b) is typically used to convey notions of mutually exclusive interests, battle and victory, but features such as death, displacement and horror are usually disregarded, so as not to invalidate the positive value that is ascribed to aggression in much of business discourse.

According to conceptual metaphor theory, selective mappings can be accounted for by the invariance hypothesis (Lakoff 1993), which holds that the conceptual structure of the target overrides that of the source domain, i.e. precedence is given to what the speaker wants to express about the area that is being conceptualized. To elaborate the previous example, expressions of war may be used to talk about business in order to focus on the aspect of competition rather than cooperation. Mappings can be more or less exhaustive, with single-feature-mappings at one extreme and systematic mappings at the other. Systematicity works both at the level of language and that of cognition; thus, metaphoric expressions are related if they express the same underlying metaphor. For instance the expressions ‘she boosts the morale of the troops’, ‘I am ruthless in using every weapon I have’ and ‘She barked out new marching orders’ as used by and about businesswomen all realize a conceptual metaphor BUSINESSWOMEN ARE MILITARY LEADERS (Koller 2004a; see sub-section 4.2). At the same time, conceptual metaphors themselves are related if they elaborate on a higher-level metaphor. For example, MARKETING IS ATTACKING (realized as ‘Oracle has launched a two-pronged assault’), COMPETITORS ARE ENEMIES (‘the company has no greater rival’) and ADVERTISING IS A WEAPON (‘customers are bombarded with commercials’) can all be traced back to BUSINESS IS WAR. This latter kind of systematicity shows that conceptual metaphor operates at different levels of complexity. Especially more recent versions of conceptual metaphor theory (Lakoff and Johnson 1999) posit that cognition in general, and metaphor in particular, are grounded not only in socio-cultural but also in physical experience. Such primary metaphors, deriving either from the experience of moving one’s body through space or from perceiving natural phenomena such as rising levels when liquid is poured into a container, draw on the source domain of spatial orientation, with more abstract metaphors deriving from them in subsequent cognitive processes. Primary metaphors based on physical experience are embodied, e.g. PURPOSES ARE

DESTINATIONS, which relies on a source-path goal schema (Johnson 1987) and entails PROBLEMS ARE OBSTACLES, PURSUING A PURPOSE IS MOVING ON A PATH (e.g. ‘we’re running around in circles’) etc. The metaphor thus structures a whole event and yields the more abstract conceptual metaphor LIFE IS A JOURNEY (see sub-section 4.10).

The theory part of this article will now close by linking gender and metaphor to the notion of discourse.

2.3. Metaphor and gender in discourse

Combining notions advanced by Fairclough (2003) and van Dijk (2003), we can define discourse as a use of language as a social practice that is based on cognition. In this context, we can, again elaborating on Fairclough (1995a), distinguish between discourse as a mass noun (which can be further characterized by adjectival pre-modification to indicate the social area in which the discourse is produced, distributed and received, e.g. political discourse), and discourse as a count noun (defined as different ways of representing aspects of the world).

Again, discourses can be further differentiated through adjectival pre-modification and/or prepositional post-modification to indicate

- the topic of the discourse, e.g. discourse on immigration
- the discourse producers, e.g. the Conservative Party’s discourse on immigration
- the geographical origin, e.g. British discourse on immigration
- the historical period of the discourse’s production, distribution and reception, e.g. 1980s discourse on immigration
- the stance or aim that the text producer takes on the topic, e.g. the Conservative Party’s restrictive discourse on immigration. (Obviously, there are close links between producer and stance.)

Discourse is instantiated in texts that originate in a social domain where it functions as a means to negotiate identities and relations. Related to that, *a* discourse is the totality of texts that are defined by a topic etc., and is produced, distributed and received in certain ways collectively known as discourse practice. From a language and gender point of view, a discourse can deal explicitly with gender issues as its

topic, e.g. medical discourse on men's health. Beyond that, however, we can also identify *gendered* discourses in which an originally neutral topic is implicitly linked to gender, e.g. when political or media discourse describes childcare policies as 'woman-friendly'. *Gendered* discourses are by the same token *gendering* (Sunderland 2004), i.e. they communicate ideas about gender identity and appropriate gendered behaviour. The above example, for instance, not only assumes childcare to be of higher relevance to women, but defines womanhood itself in terms of childcare responsibilities and, ultimately, motherhood.

Obviously, gendered and gendering discourses have an ideological function in that they impact on power relationships between different social groups. Metaphor aids in this process in that, as Lakoff and Johnson (1980: 156-157) noted, 'metaphors ... highlight and make coherent certain aspects of our experience ... metaphors may create realities for us, especially social realities ... people in power get to impose their metaphors'. This is echoed by Fairclough (1995a: 74), who observes that 'metaphor is a feature of language and discourse [which] may be ideologically vested'. He further posits that 'dominant metaphors construct domains in a way which helps marginalize other constructions from the perspective of oppositional groups' (1995b: 71-2).

An example is the following self-description by a female executive, who depicts herself as a combatant in a war or disaster scenario: 'My single greatest strength is seeing through the smoke into chaos, and operating where everything else is exploding'. The similarity thus established between the source domain of fighting and the target domain of managing is that both are conceptualized as a chaotic, violent environment. In gender terms, the example is an instance of how hegemonic metaphors are often taken up by non-dominant groups, here a female executive using a culturally masculine WAR metaphor (see sub-section 4.2).

After this outline of the main theories on gender, discourse and metaphor, and a discussion of the links between them, the next section will present some methods that have been fruitfully employed to analyze metaphor and gender in discourse.

3. Research methods: qualitative and quantitative, manual and automated

This section will first follow on from the previous points made on discourse by introducing a model for its analysis. After that, it will look at questions guiding the qualitative, manual analysis of metaphor in discourse and present methods that have been suggested to identify both metaphoric expressions and conceptual metaphors in language use. This will be followed by an outline of how corpus linguistic parameters, in particular keywords, concordances and collocations, can support qualitative manual analysis. The section will finish by flagging up a few ways of combining qualitative and quantitative, manual and (semi-)automated analysis. The issue of gender is backgrounded throughout this section, but will be taken up again in section 4.

3.1. A model for discourse analysis

Operating with the definition of discourse as the use of language as a social practice that is based on cognition, it follows that discourse analysis involves looking at the way writers and speakers selectively realize mental models, or socio-cognitive representations (SCRs), including metaphors, in language in order to position themselves vis-à-vis others. A critical stance in discourse analysis also includes investigating ‘the role of discourse in the (re)production and challenge of dominance’ (van Dijk 1993: 249). Given this broad remit, it is essential that texts, as instantiations of discourse, be analyzed in the context of interaction as well as in the wider social context. Indeed, text, interaction and social context can be seen as embedded within, and mutually constitutive of, each other. At each level, cognitive aspects are at work, as demonstrated in Figure 1.

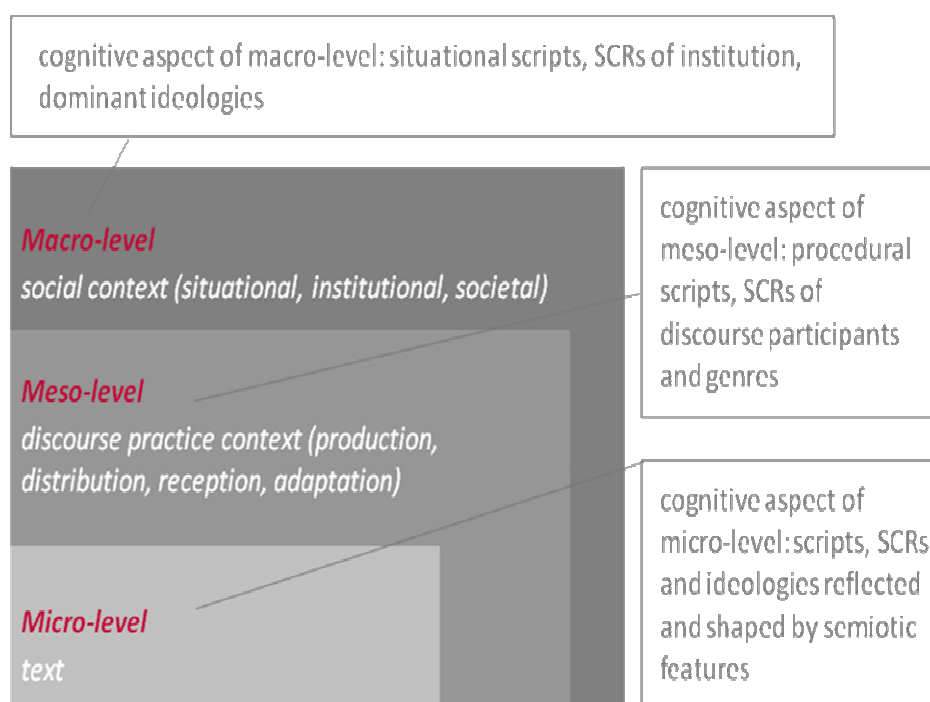


Fig.1: A tripartite model of discourse and its cognitive underpinnings (adapted from Fairclough 1995a: 98)

The starting point for a critical analysis of discourse is often the researcher's interest in a social problem or phenomenon, which can be studied through the 'semiotic entry point' (Fairclough) of textual data. The actual methods to analyse texts in contexts will be chosen to fit the research questions and data, but in general terms, a critical analysis of discourse would be structured roughly as follows: the textual analysis is descriptive and can address questions relating to both content and form. Thus, the researcher will seek to address *what* actors, entities and events are represented in the text, looking e.g. at presences and absences of social actors, propositional analysis and argumentation strategies. Still at the level of text, the description will also be of the text's semiotic features, including metaphoric expressions, asking *how* actors etc. are represented. This stage of the analysis will entail a detailed description of linguistic and other parameters to see how text producers represent themselves and position the recipient and any third parties that are referred to in the text. A semiotic analysis points out the links between different

linguistic parameters, and between linguistic and other parameters, such as visuals.

Moving from text to context, we need to distinguish between discourse practice, a mid-range level context that encompasses the production, distribution, reception and possible adaptation of texts, on the one hand, and the social context on the other. Relevant questions about discourse practice are related to the notion of genre and address the relationship between text producer and recipient, audience design and communicative purpose. Another issue is access to means of production and distribution, as it ensures the impact a text has. Cognitive aspects at the meso-level include SCRs of various text types — e.g., what does a prototypical interview look like? —, producers and audiences, all of which influence production and reception. Further, there are procedural scripts, or structured knowledge about the conditions and temporal sequences of discourse practices, e.g. the question of how press releases are transformed into news articles. At the macro-level, discourse analysis addresses the wider social context, looking at the values, attitudes and ideologies as well as any institutional contexts that impact on discourse practice and, ultimately, on texts. In short, while textual analysis investigates, in its content part, *what* actors etc. are represented and, in its semiotic part, *how* they are represented, context analysis seeks answers to the question as to *why* certain actors etc. and not others are represented, and why they are represented in a particular way. As such, context analysis explains the results of textual analysis with recourse to social and discourse practice contexts, and their cognitive underpinnings. It therefore represents the interpretive stage of analysis, and the one where cognitive aspects are most important. The next section will focus on how to analyze a particular form that the cognitive underpinnings of discourse can take, i.e. metaphor.

3.2 . Identifying metaphor in discourse

In section 2.2, it was pointed out that in the most prevalent current approach to metaphor, conceptual metaphor theory, metaphor is seen as primarily a cognitive phenomenon that is reflected at the level of text, or any other semiotic mode. However, for the analysis of metaphor in discourse it is necessary to work backwards, first

ascertaining metaphoric expressions in language and inferring from them what conceptual metaphors may have given rise to them.

A recent but already very influential method for identifying metaphoric expressions in discourse has been advanced by the Pragglejaz Group (2007). Taking text, i.e. a semantically coherent and grammatically cohesive instance of language use, as an analytical unit, their metaphor identification procedure is outlined as follows:

1. Read the entire text to establish a general understanding of the meaning.
2. Determine the lexical units in the text (words and phrasal verbs).
 - 3a. For each lexical unit in the text, establish its meaning in context, that is, how it applies to an entity, relation, or attribute in the situation evoked by the text. Take into account what comes before and after the lexical unit.
 - 3b. For each lexical unit, determine if it has a more basic (\neq frequent) contemporary meaning in other contexts. Basic meanings tend to be more concrete or historically older.
 - 3c. If the lexical unit has a more basic contemporary meaning in another context, decide whether the contextual meaning contrasts with the basic meaning but can be understood in comparison with it.
4. If yes, mark the lexical unit as metaphorical.
5. Check difficult cases against at least two recent corpus-based dictionaries.

(based on Pragglejaz Group 2007: 3)

For example, a sentence such as ‘When I am verbally assaulted I become very furious’² can be analyzed by first paraphrasing the meaning as ‘I become angry when people speak words I consider harmful to me’, and then identifying the lexical units of the utterance

² The example is taken from the ongoing PhD work by Gladys Ansah at the Department of Linguistics at Lancaster University (UK), which investigates metaphor use as a window on bilingual cognition. At the same department, Ayako Tominari’s PhD-in-progress (on the representation of male heroes in sports reporting) adapts the Pragglejaz method to analyze data in Japanese.

as when/I/am/verbally/assaulted/I/become/very/furious. After that, the analysis proceeds word-by-word. For instance, the contextual meaning of ‘assaulted’ refers to the use of words in a way the speaker finds unacceptable or face-threatening, while the basic meaning of the word is a violent attack, an act that threatens a person physically. Considering the fact that the context of the word ‘assaulted’ does not refer to actual physical harm to the speaker, the lexical unit has been used metaphorically.

While the Pragglejaz approach is an important step towards validating metaphor identification at the linguistic level, it is not intended to deal with conceptual metaphor. In order to start tackling the issue of conceptual metaphor identification,³ Steen (2002) has developed a five-step method that involves identifying, in this order, metaphorical focus, metaphorical idea, metaphorical comparison, metaphorical analogy and metaphorical mapping. It should be noted that the last step to identifying metaphoric expressions according to the Pragglejaz method is the beginning of identifying conceptual metaphors, namely comparing the basic and contextual meanings of a metaphorically used word. As such, this step coincides with the third step in Steen’s method. For example, in the word ‘assaulted’, discussed above, the contextual meaning (speak harmful words) can be contrasted with its more basic meaning (physical attack) to understand the contextual meaning through a metaphorical comparison between the target domain (speech acts) and the source domain (physical attack). Going beyond the Pragglejaz method, however, Steen’s five-step procedure involves breaking up a text into its propositions, or ‘conceptual representations of the basic idea units or thoughts in a text’ (p.22). These minimal units are then identified as metaphoric or literal by drawing an analogy between their meaning in the text and any more concrete meaning. If there are two meanings that can be located in different conceptual domains, the unit counts as metaphoric. The final step of the procedure is then to make explicit the mapping that has taken place. (for further detail, see Steen 1999).

A final caveat concerns the labelling of conceptual metaphors: The researcher needs to be aware that the way they name conceptual metaphors that they identify in early stretches of a text, and the level

³ Although there is no space to discuss it here, Cameron and Deignan (2006) have developed the notion of ‘metaphores’, which they theorize as being positioned between linguistic and conceptual metaphor and emerging in local stretches of discourses.

of generality at which they identify them, can bias subsequent analysis. For instance, the phrase ‘steer the business through turbulent times’ can be traced back to conceptual metaphors SHIP IS COMPANY, MANAGING IS STEERING or BUSINESS IS A SEA, but the alternatives obviously capture different semantic features.⁴ As a rule of thumb, the label given to reconstructed conceptual metaphors should be closely oriented towards the textual evidence, which for this example would favour the second alternative.

In sum, procedures such as those outlined above limit the role of ‘informed intuition’ in metaphor identification (Deignan 1999: 180), even if traces of it may remain. However, qualitative methods, while offering valid ways to identify metaphor, cannot deal with large numbers of texts. Yet if we accept Eubanks’s (2005) point that metaphor is at least in parts constituted and reinforced intertextually, we need to look at substantial amounts of authentic data. The next section will outline how corpus linguistics can help with that.

3.3. Corpus linguistics

Corpus linguistics involves analyzing large electronic collections of natural language with the help of special computer software. As such, it has to be distinguished from computational linguistics, which is a branch of Artificial Intelligence that investigates the computational aspects of the human language faculty, including automated natural language processing (e.g. speech recognition or automated translation) and language generation. Despite ongoing attempts (Barnden 2009, Birke and Sakar 2006, Berber Sardinha 2007, Fass 1997, Mason 2004), some of them rather promising, it is not (yet) possible to reliably identify metaphor in natural language in a fully automated way. We therefore have to rely on corpus-based research, which helps identify metaphor *candidates* and facilitates both quantitative and qualitative analysis. As far as metaphor analysis is concerned, corpus work is semi-automated in that data may require manual annotation and automatically generated results are in need of manual editing.

⁴ Note that at the conceptual level, metaphors are unlikely to be structured in parts of speech. It seems more plausible that conceptualizations are either image schematic or embodied (personal communication with Elena Semino, 29 April 2010).

Seen within these limits, however, corpus-based metaphor research allows the analyst

- to study metaphor use in a discourse community;
- to check intuitions about the nature, frequency/importance and systematicity of conceptual metaphor;
- to ascertain textual properties of conceptual metaphors;
- to identify different functions of metaphoric expressions in different parts of text;
- to link metaphoric expressions to grammatical patterns in the co-text;
- to investigate diachronic and cross-linguistic differences.

(see Stefanowitsch 2006)

Data can either be provided by large off-the-shelf corpora combining several genres, which can be continuously added to (so-called ‘open’ corpora, e.g. the Bank of English, which in 2008 stood at 650 million words) or not (closed corpora, e.g. the 100-million-word British National Corpus or the 20-million-word PAROLE French Corpus). These usually come with inbuilt analysis software. Alternatively, corpora can be purpose-built to help answer particular research questions (on pitfalls and good practice in corpus-building, see Baker 2006: 25-46). Such corpora will then be analyzed with external software packages such as Wordsmith Tools or AntConc. Among the typical parameters to be analyzed in corpus linguistics are keywords, i.e. words that are over- or underused with statistical significance compared to a reference corpus. It should be noted that keywords are not usually metaphoric expressions — few metaphoric expressions are frequent enough to be key —, but they can feature in or reinforce them (see sub-section 4.1 for examples). Further, concordances establish the frequencies of certain words, and the lexical and grammatical patterns around them, while collocations show which words co-occur with statistically significant frequency with a particular search word. Taken together, concordances and collocations can demonstrate how metaphors structure texts. Figure 2a shows a concordance, generated by Wordsmith Tools 3, for the search word ‘women’ in a corpus of business magazine texts, while Figure 2b represents its collocates.

N	Concordance	Set	Tag	Word No.	File	%
1	nsume (table, page 53). Women are moving into			613	29~1.doc	13
2	<ms> enough. After all, women who suffer from			432	28~1.doc	16
3	usual all-women panel judges, in			1,070	ft25~1.txt	87
4	households. These are women who			168	sft9~1.txt	38
5	emy. It will be aimed at women who want a soph			1,172	32~1.doc	25
6	oned its planned career-women line.			370	o2~1.doc	18
7	status as a boys' club; women now			2,241	24~1.doc	37
8	tels & Resorts features women executives in ma			1,734	er~1.wbk	31
9	ass--bras for full-figured women were her career f			857	29~1.doc	15
10	literary award for women writers. The origi			174	ft25~1.txt	14
11	ver wanted to dress just women who lunch." Klei			1,202	29~1.doc	20
12	can drive with their kids. Women want small cars			1,778	25~1.doc	37
13	gan to change as more women went			281	ft88~1.txt	34
14	er PMDD is real or not, women beware. Those			542	28~1.doc	19
15	ith an estimated 35% of women now working out			1,726	29~1.doc	30
16	y rate of women is much lower th			233	ft92~1.txt	51
17	Lilly says that 3-5% of women of child-bearing			198	28~1.doc	10
18	xisting core audience of women over 45.			344	ft20~1.txt	60
19	w> is 20- to 30-year-old women in urban centers			2,021	w6~1.doc	40
20	anchor desk with pretty women a generation or t			314	c6~1.doc	12
21	Single women - particularly sin			24	sft9~1.txt	7
22	lds consisting of single women is growing. Spen			921	25~1.doc	21
23	ficant numbers of single women have above avera			66	sft9~1.txt	17
24	men - particularly single women with children - ar			27	sft9~1.txt	8
25	Taiwanese women are very much in			72	ft92~1.txt	17
26	only the small print tells women that Sarafem an			488	28~1.doc	18
27	the name change is that women who do not cons			412	28~1.doc	16
28	. Few of the women in the admiring c			264	w6~1.doc	8
29	r. When the women said they had th			884	w6~1.doc	19
30	of them women working part-tim			404	ft88~1.txt	49
31	gnore these women at their peril.'			227	sft9~1.txt	53
32	stopped two women on the street an			867	w6~1.doc	19
33	ed by urban women on the go.			1,231	w6~1.doc	26
34	self and the two young women the g			577	18~1.doc	18
35	h firm, found that young women attracted by an i			717	15~1.doc	20

Fig. 2a: Concordance for search word 'women' in a business magazine corpus

N	WORD	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	WOMEN	67	4	5	1	1	2	0	0	58	0	1	1	2	1
2	THE	33	17	16	6	4	3	0	4	0	0	8	0	5	3
3	OF	26	24	2	1	2	2	8	11	0	1	0	0	0	1
4	WOMEN'S	23	1	3	0	0	0	1	0	19	2	0	1	0	0
5	TO	20	17	3	6	6	3	0	2	0	0	0	1	0	2
6	AND	18	8	10	1	3	2	1	1	0	1	1	3	4	1
7	S	17	0	17	0	0	0	0	0	0	17	0	0	0	0
8	A	12	6	6	2	1	2	0	1	0	1	0	1	1	3
9	ARE	12	2	10	0	0	1	0	1	0	3	1	2	2	2
10	IN	11	1	10	0	0	0	0	1	0	2	2	0	2	4
11	THAT	10	6	4	2	2	0	1	1	0	1	0	1	0	2
12	FOR	9	8	1	1	1	1	2	3	0	0	0	0	0	1
13	WITH	8	4	4	2	0	1	1	0	0	1	2	0	0	1
14	AN	7	4	3	0	3	0	1	0	0	0	0	3	0	0
15	SINGLE	7	5	2	0	1	0	0	4	0	0	1	1	0	0
16	WHO	7	0	7	0	0	0	0	0	0	5	2	0	0	0
17	AS	6	5	1	1	1	1	2	0	0	0	0	0	0	1
18	IS	6	2	4	0	1	0	1	0	0	2	0	1	1	0
19	PRODUCTS	6	2	4	1	0	0	0	1	0	0	3	0	0	1
20	APPAREL	5	0	5	0	0	0	0	0	0	0	0	3	2	0
21	ITS	5	3	2	0	1	1	0	1	0	0	0	0	1	1
22	NOW	5	0	5	0	0	0	0	0	0	2	2	0	1	0

Fig. 2b: Collocates for search word 'women' in a business magazine corpus

To date, most corpus-based metaphor research has taken a lexical approach of some sort.⁵ Thus, work has been done that involves searching for source domain vocabulary in the form of individual lexical items or sets of words, as based on pre-defined lists or key word analysis (Partington 1997, Koller 2005). Other studies have started at the other end, identifying target domain vocabulary and ascertaining if and how words are embedded in metaphoric expressions (Stefanowitsch 2004). Yet other research has opted for a combined approach, looking for sentences containing lexical items from both source and target domain (e.g. Martin 2006). The last two methodologies, however, are ailed by the fact that source and target domain vocabulary rarely co-occur in language use. To circumvent this problem, Cameron and Deignan (2003) have searched for markers of metaphor in corpora, e.g. ‘as it were’, ‘so to speak’.

From the above, it is clear that while corpus-based analysis provides a broad view of large amounts of data, qualitative analysis of sample texts helps gain an in-depth view of selected texts. Ideally, the two should therefore be combined. For the analysis of (gendered) metaphor in discourse, this can be done in essentially two different ways. On the one hand, corpus analysis software can be used to identify possible sample texts with high density of relevant metaphoric expressions. Those sample texts can then be analyzed qualitatively for other parameters that reinforce metaphoric concepts. Alternatively, metaphors can be manually identified in a small sample of a larger corpus and then extended to the whole corpus with the help of corpus analysis software (see sub-section 4.1). In both cases, results will need manual checking.

According to the discourse analysis model presented in sub-section 3.1, the identification of metaphoric expressions, whether done manually or in a semi-automated fashion, is part of descriptive textual analysis, while the reconstruction of underlying conceptual metaphors shades over into interpretation. Since there are no intrinsically gendered language features — at least not according to the social constructivist approach, which is followed here —, any claims about how metaphor contributes to discursively constructing gender identities arise from the interpretative analysis of discourse practice

⁵ A different approach is taken in work that uses semi-automated semantic annotation to identify metaphor candidates in corpora (Hardie et al. 2007, Koller et al. 2008).

and social context. The next section will present three case studies highlighting the links between metaphor, gender and discourse.

4. Case studies

In sub-section 2.1, a distinction was made between discourses that explicitly deal with gender as a topic on the one hand, and gendered discourses on the other. Gendered discourses, it was suggested, are simultaneously gendering. It should also be noted that text producers not only have the option to draw links between an originally neutral topic and notions of gender, a process which can be uncovered through content analysis. In addition, they can utilize language and conversation features to perform certain gendered styles in order to fulfill or defy expectations. With regard to metaphor, this leaves us with three ways in which it can be related to gender in discourse: Firstly, we can investigate if and how speakers and writers use metaphor in different ways according to their gender, and why they may do so. Secondly, we can look at examples of discourse on gender and see how metaphors are used to talk about men and women. And finally, we can abstract from the gender of speakers and those they talk about and analyze how metaphor is employed to construct discourses and the social domains in which they originate as culturally feminine or masculine. The three case studies below will address these different links between metaphor and gender.

4.1. Men and women using metaphor: a case study from political discourse

This section will summarise a study of the use of metaphor by two German politicians, former chancellor Gerhard Schröder and current chancellor Angela Merkel (see Koller and Semino 2009). The assumption undergirding the research is that speakers can use gendered metaphor to be perceived as masculine or feminine and to benefit from the respective cultural associations. However, it is not assumed that the metaphors used by speakers of different genders can necessarily be explained as an effect of their gender, i.e. the research does not subscribe to the difference approach. Rather, it seems plausible that metaphor use can vary according to a range of factors,

including context, topic and individual preferences. Given the fact that politics and its discourse is a historically male and culturally masculine domain, it is interesting to look at how male and female speakers position themselves in it.

The two politicians under investigation have at different times (1998-2005 and since 2005, respectively) held the same office, i.e. head of government. However, while Gerhard Schröder's political home is the Social Democratic Party, Angel Merkel doubles as the federal chairperson of the conservative Christian Democratic Union. Further differences are that the former chancellor comes from a working-class background and first trained as a salesman before completing a law degree and being admitted to the bar. As chancellor, he projected an anti-intellectual working-class image, but was perceived as a social climber, earning him the epithet 'comrade of bosses'. In personal terms, he very much presented himself as a 'red-blooded male': he was married for the fourth time, and cultivated an ebullient image, even appearing drunk in public after his re-election and bullying Angela Merkel after his electoral defeat by bluntly refusing, in a television debate, to step down as chancellor. Ten years his junior, Angela Merkel grew up in the former German Democratic Republic, completed a PhD in physics and worked as a scientist until the Wende. First known as '[then chancellor Helmut] Kohl's little girl', her political career was meteoric, which is all the more remarkable as her demographic profile as a childless, Protestant woman from eastern Germany is very atypical for a senior Conservative in Germany. Her public image is that of a well-organized, diplomatic and cool person. Married for the second time, she largely refuses to discuss her private life and accordingly, her husband keeps a very low public profile.

The data for the study comprise 116,421 words in total, divided into two sub-corpora of just under 60,000 words for each politician. Each sub-corpus includes approximately 30,000 words of speeches and interviews, with the Schröder sub-corpus spanning the years 2000-2005 and the Merkel sub-corpus covering the period 2005-2007. In terms of genre, the Schröder sub-corpus contains ten speeches and ten interviews, compared to seven speeches and eight interviews in the Merkel sub-corpus. All speeches were delivered in institutional context, e.g. in parliament or at conferences, and address employment, European politics and the federal budget. Interviews are more diverse

in their subject matter, which for Angela Merkel also includes journalists asking about private matters such as her marriage. Finally, the interview data cover a range of media outlets such as newspapers (tabloids, broadsheets and financial papers), news magazines, public radio stations as well as public and private TV channels.

The sub-corpora were first analyzed for keywords with the help of *Wordsmith Tools*, in order to obtain an overview of speech styles, as manifested in lexical contrasts. The results of this initial stage show that references to political issues and entities are very much dictated by the politics of the day. There are also differences with regard to abstract entities such as values; the reference to ‘participation’ (‘Teilnahme’) is as predictable coming from a Social Democrat as is the Christian Democrat’s focus on ‘belief’ (‘Glaube’). The most notable differences concern Merkel’s use of the intensifiers ‘very’ (‘sehr’) and ‘many’ (‘viele’), which contrast with Schröder’s low-affinity ‘occasionally’ (‘gelegentlich’). Moreover, Merkel refers to herself and others as well as to emotional states and verbal processes, which are absent in Schröder’s language use. This makes for a gender-stereotypical impression where the female politician addresses not only political issues but also emotions, refers to the ‘people’ and ‘we,’ and shows an intensified speech style.

Moving on to metaphor, metaphoric expressions were manually analyzed in a core section of the corpus containing approximately 30,000 words, equally divided between the two sub-corpora. Following the Pragglejazz method (2007), this resulted in a list of metaphoric expressions used by each politician. The analysis then returned to corpus methods by employing the Wordlist facility in *Wordsmith Tools* to broaden the analysis to the whole of the data. This generated lists for all the different words in each sub-corpus, which were scanned to identify further metaphor candidates. The resulting extended lists of potential metaphoric expressions were grouped according to the source domain that they realize, the most notable ones being JOURNEY and WAR. These two display interesting contrasts across the two sub-corpora, and the related words from the word list were concordanced to find all instances of the expressions relating to these source domains, and to identify all metaphoric uses of those expressions in the corpus.

The JOURNEY metaphor is extremely productive in language use and has been confirmed for political discourse in countries as far apart as

Britain, the US and China (Charteris-Black 2005, Teng and Sun 2001). Its ideological function is to positively evaluate certain policies, since the metaphorical journey's destinations are seen as desirable (Charteris-Black 2004: 93). In the corpus as a whole, 22 different words realize the JOURNEY metaphor, with Merkel using a higher number of different expressions and also using them more densely, particularly in her speeches. Merkel further reinforces the JOURNEY metaphor through other linguistic parameters, such as adverbs and adjectives ('...not to stand still on this path of reform but to continue on it with determination')⁶, and high-affinity deontic modality ('What we have to do here is to clearly define the goal').⁷ This emphatic language use is consistent with her keywords 'very' ('sehr') and 'many' ('viele'), and is combined with empathy, as conveyed in the keyword 'people' and reference to the 'pain' felt by others. (Note that Schröder only refers to the pain he himself felt at professional disappointments.) However, Merkel's metaphor use is repetitive across the sub-corpus, which suggests a strategic employment of (scripted) language, including metaphor, to appear engaged and passionate and counter her cool image. Empathy and emphatic language can be interpreted as stereotypically feminine traits, but also as the (gendered) characteristics of a 'compassionate Conservative', a type of politician embodied e.g. by David Cameron in the UK, and Merkel may strive to project that image. Schröder's treatment of the JOURNEY metaphor differs in that it foregrounds the notion of problems as obstacles on a journey, which need to be 'approached with courage', 'overcome', 'tackled' and 'attacked'. The last expression relates the JOURNEY to the WAR metaphor, and indeed the Schröder sub-corpus shows a higher density of these, namely 0.72 per 1,000 words compared to Merkel's 0.43. However, this is largely an effect of context, as occurrences of the 'fight' cluster in interviews conducted with Schröder in the run-up to the general election in September 2002.

This brief outline shows that differences in language and metaphor use between male and female politicians can be due to a range of factors, including party membership, worldview and impression management as well as personality and professional background: the

⁶ German original: '...auf diesem Reformweg nicht stehen zu bleiben, sondern ihn ganz entschieden weiter zu gehen.'

⁷ German original: 'Hier müssen wir nur das Ziel klar definieren.'

former physicist Merkel shows a much more formulaic and unvaried metaphor use than does the lawyer Schröder. Other factors are current politics, e.g. the more pronounced use of the WAR metaphor before a general election, and genre: it stands to reason that the WAR metaphor will be used more in election rallies than in institutional speeches. In sum, performance of gendered identities should not be ignored when considering variation in metaphor use in individuals, but such variation can only be properly explained by taking into account a range of factors, of which the performance of masculinity and femininity is one.

The second case study will shift the focus from metaphor and speakers' gender to metaphor used in the representation of gendered social actors.

4.2. Metaphor describing men and women: a case study from business magazines

This second piece of research investigates what metaphors journalists use to talk about businessmen and women (Koller 2004b). As such, it looks at how metaphor can be employed to position people in gender roles, reinforce or challenge stereotypes, and negotiate power relations between groups. To this end, the study not only discusses metaphoric expressions and conceptual metaphors, but also coherent metaphor complexes (Lakoff and Johnson 1980: 97-105). These combine related conceptual metaphors; for example, the metaphoric expression 'she is a street fighter' can be traced back to a conceptual metaphor BUSINESSWOMEN ARE FIGHTERS, which in turn is part of the metaphor complex aggression and competition. The data that were analyzed are shown in Table 1:

Corpus A	Corpus B
40 articles	40 articles
Five business magazines/papers	Five business magazines/papers
8 <i>BusinessWeek</i>	10 <i>BusinessWeek</i>
12 <i>Economist</i>	12 <i>Economist</i>
4 <i>FinancialTimes</i>	13 <i>FinancialTimes</i>
7 <i>Forbes</i>	1 <i>Forbes</i>
9 <i>Fortune</i>	4 <i>Fortune</i>
Published between 1996 and 2001	Published between 1995 and 2002
81,378 words	56,692 words
Average article length 2,034 words	Average article length 1,417 words
Eight compilations and 32 portraits of 20 different women	Four compilations and 36 portraits of 35 different men

Table 1: Data for study of metaphoric description of businessmen and women

The above data were analyzed manually by searching for relevant metaphoric expressions and similes (e.g. ‘she is like a modern nomad’) in the descriptions of businessmen and women by journalists and in quotes. The linguistic evidence was traced back to conceptual metaphors, broadly following Steen’s (2002) five-step method. The quantitative analysis identified the absolute number of relevant metaphoric expressions for each corpus, the number of different conceptual metaphors in each case and the respective percentage of the WAR metaphor. The last aspect was included not only because the WAR metaphor is so prevalent in business discourse as to be a defining feature of it (Koller 2004a), but also because it draws on a historically male and culturally masculine source domain. The subsequent qualitative analysis in turn ascertained conceptual relations and oppositions between metaphors, and divided the conceptual metaphors into higher-level metaphor complexes.

Results show an absolute number of relevant metaphoric expressions of 186 for businesswomen (Corpus A) compared to 190 for businessmen (Corpus B). While these numbers are very similar, there is a wider range of metaphors for businessmen, namely 33 compared to the 23 used to describe businesswomen. This suggests that the masculine domain of business is based on more socio-cognitive representations for masculinity, making it the more differentiated model. Interestingly, the proportional relation of metaphoric expressions of war to all metaphoric expressions is higher

in the businesswomen corpus; 47.85 per cent compared to 31.05 per cent for businessmen. This means that businesswomen are relatively more often described in terms of the WAR metaphor than are businessmen. This rather surprising finding may be most plausibly explained with the WAR metaphor here serving to co-opt women into the ideal of corporate masculinity, which positions globalized (male) managers as representatives of hegemonic masculinity and *ersatz* soldiers (Connell 1998).

In qualitative terms, we can identify three metaphor complexes:

- aggression and competition, including BUSINESS(WO)MEN ARE HUNTERS and BUSINESS(WO)MEN ARE WARRIORS (‘her plan is to hunt bigger prey’, ‘he has lost ground to rivals on international fronts’);
- care and attention, including BUSINESS(WO)MEN ARE NURTURERS (‘[she is] one of the most hands-on nurturers’, ‘[he will] nurture the euro towards maturity’);
- control and leadership, including BUSINESS(WO)MEN ARE MONARCHS and BUSINESS(WO)MEN ARE DRIVERS (‘Europe’s queen of the web’, ‘he is at the wheel’).

Table 2 shows the distribution of these metaphor complexes across the two corpora.

	Corpus A	Corpus B
aggression and competition	40.42%	32.61%
care and attention	66.66%	27.27%
control and leadership	37.5%	35.29%

Table 2: Metaphor complexes in the description of businesswomen (A) and men (B)

The higher proportion of the aggression and competition complex for businesswomen, which is due to the disproportionate use of the WAR metaphor, is balanced by the much higher proportion of the care and attention complex in that corpus, and the fact that the same complex is by far the most frequently employed in the description of businesswomen. In addition, the complexes comprise many gender-specific metaphors, e.g. cheerleaders, flowers, and diamonds for women and cowboys, firefighters and religious leaders for men.

In sum, we find a hybrid metaphoric description of businesswomen, which indicates their changing and unstable image in a masculine social domain. The last case study will briefly outline how at an abstract level, particular metaphors serve to define social domains in gendered terms.

4.3. Metaphors gendering social domains: a case study from marketing

So far, we have seen that text producers use metaphor both to perform masculinity and femininity in particular contexts and to position others in particular gender roles. This final case study illustrates how text producers further draw on the gendered nature of particular metaphor to construct social domains. The data used for this are selected advertisements from British, US and Austrian business magazines, published between 1996 and 2001 (see Koller 2008). The examples were collected to see how particular metaphors are realized, so only advertisements featuring the metaphors in question were included in the sample. Because advertisements form a multimodal genre, the analysis will not only identify and categorize verbal metaphoric expressions, but also draw on several frameworks for visual analysis (Barthes 1977, Forceville 1996, Kress and van Leeuwen 2006) to address how linguistic and visual elements interact to express metaphors.

Within the data, the focus is on RELATIONSHIP and WAR metaphors. This is motivated by shifts in marketing theory and practice, i.e. the paradigm change from transactional to relationship marketing (RM), which began in the early 1990s.

Transaction marketing	Relationship marketing
focus on product	focus on customer
goals: customer acquisition, profit maximization	goal: customer retention, mutual benefit
short-term tactics	long-term strategy
externally oriented (customers)	Externally and internally oriented (customers, employees, investors)
competitors as rivals	competitors as partners
product at centre	brand at centre

Table 3: Transaction vs. relationship marketing (Parvatiyar & Sheth 2001)

While some of RM's tenets may with hindsight seem overly idealistic, the paradigm has nevertheless had a lasting impact on marketing strategies and practices. This is not least reflected in advertisements as one form of marketing communication.

The first finding of the analysis is that the WAR metaphor is absent in advertising. This is unsurprising, as the negative connotations of war are not conducive to the genre's persuasive purpose. On the other hand, there is an overarching metaphor MARKETER-CUSTOMER RELATIONSHIPS ARE PRIVATE RELATIONSHIPS, which is further differentiated into the source domains FAMILY and ROMANCE.⁸ In the FAMILY metaphor, marketers are positioned as parents and customers as children. For instance, an advertisement for oneworld airlines shows a picture of two penguins bending their heads to a baby penguin that faces the viewer and is positioned between the larger penguins. The copy reads:

It's natural to want to be cared for. Oneworld is an alliance of eight airlines, with one goal: to care for our passengers. Whether that's rewarding your loyalty with frequent flyer miles, giving you access to over 340 lounges or simply being there to look after you. Why? Because oneworld revolves around you.

The text here anchors the image (Barthes 1977), which would otherwise seem out of place in a business magazine. The metaphoric expression is verbo-pictorial (Forceville 1996), with the copy referring to the target domain of service and the image representing the source domain of family, here transferred to animals. The layout centres the metaphorical customer horizontally and vertically, and provides additional focus through the parents' bent heads. Although certainly not intended, it is worth noting that the FAMILY metaphor is double-edged, positioning as it does the customer as protected but powerless.

The ROMANCE metaphor is frequently realized in the selected advertisements, typically through images of couples, as in the advertisement for hotel chain Accor that is shown in Figure 3.

⁸ There is also a FRIENDSHIP metaphor, which is not discussed here.



Fig.3: Accor advertisement

The copy reads:

Throughout the world, it's our pleasure to earn your smile. Accor is known for its innovation in the hotel industry. With more than 3,500 hotels worldwide, from budget to luxury establishments, you will always find one that's just right for you. Accor is equally innovative in services to corporate clients and public institutions designed to the well-being and productivity of their employees. It's all part of Accor's ambition — to earn the smiles of our customers.

The text spells out the target domain and also repeatedly verbalizes the smiles depicted in the visuals, thus anchoring the image. The image itself combines both source and target domain in that the people in it may be read as customers or as expressing the metaphorical relationship between marketer and customer. In the latter interpretation, the woman can be interpreted as the wooed customer; the focus is clearly on her as her face is shown twice and her back is reflected in a mirror, while the man is looking at her. The advertisement is also typical in showing a heterosexual couple; indeed, romance is represented as exclusively heterosexual and monogamous, very likely because other forms are not entrenched enough as socio-cognitive representations to function as source domain material. Incidentally, the Accor advertisement also depicts fairly stereotypical gender roles, with a woman in evening dress snuggling up to a man in business attire.

Gender stereotyping is even more pronounced in the final example of the ROMANCE metaphor, an advertisement for online software shop Tendi that is reproduced in Figure 4.



Fig.4: Tendi advertisement

The German copy translates as follows:

Which is the right one? Software Service Shopping for the undecided. Many mothers have pretty daughters. But which is the right one for you? Tendi.com is software shopping on the Internet. We offer more service, more information, more clarity — and save you from being spoiled for choice fast. Whether you finally consent to marry Adobe, Corel, Lotus, Macromedia, Microsoft, Networks Associates, Quark or Symantec, will then be your own decision. And that can be for life.

The example is another verbo-pictorial metaphor, although here the image represents the target while the text incorporates both source and target. In line with the typical software customer, the reader is clearly constructed as male, and beyond that is elevated and flattered by the angle and central perspective, which make him look down at a vast number of women to have his pick from. These examples show that the ROMANCE metaphor is ambiguously gendered, conceptualizing customers either as women to be courted by the marketer or as men choosing metaphorically female brands. Clearly, the RELATIONSHIP, especially the ROMANCE, metaphor positions people in gendered roles

and also constructs a social domain such as marketing as a gendered space. Most remarkably, even alternative, potentially subversive metaphors can be employed to reproduce gender stereotypes.

5. Conclusion

This article has presented a wide range of topics. In terms of theory, it outlined language and gender research, conceptual metaphor theory and discourse analysis, while also covering issues in metaphor identification, including corpus linguistic methods. The guiding idea that links all these is that the function of metaphor in discourse is to construct reality from a particular vantage point, which in turn influences relationships between discourse participants. The particular focus was on how metaphor is linked to gendered realities, i.e. representations of the world in discourse that centre on notions and ideologies of masculinity and femininity. The case studies showed that links between metaphor and gender play out in different ways, including the gender of text producers and its possible relation to metaphor use, the way metaphor positions men and women, and, on an abstract level, how metaphor can evoke socio-cognitive representations of masculinity and femininity and thereby gender social domains. Much more could have been said about each of these links and each opens up new questions that will have to be addressed in future research. Within the space of this article, I hope to have shown that metaphor and gender are closely related in discourse and that such a complex topic needs a firm grounding in theory just as much as a valid methodology.

References

- Aristotle (ca. 350 BC/1932): *Poetics*. (Trans. W.H. Fyfe). Cambridge, MA: Harvard University Press.
- Aristotle (ca. 350 BC/n.d.): *Rhetoric*. (Trans. W. Rhys Roberts). Available: <http://classics.mit.edu/Aristotle/rhetoric.html>. Accessed 1 May 2010.
- Baker, P. (2006): *Using Corpora in Discourse Analysis*. London: Continuum.
- Barnden, J.A. (2009): 'Metaphor and context: A perspective from artificial intelligence.' In A. Musolff and J. Zinken (eds) *Metaphor and Discourse*. Basingstoke: Palgrave, pp. 79-94.
- Barthes, R. (1977): 'Rhetoric of the image'. In *Image, Music, Text* (Trans. S. Heath). London: Fontana, pp. 269-286.
- Baxter, J. (2003): *Positioning Gender in Discourse: a Feminist Methodology*. Basingstoke: Palgrave.
- Baxter, J. (2010): *The Language of Female Leadership*. Basingstoke: Palgrave.
- Berber Sardinha, T. (2007): 'A program for identifying metaphor candidates in corpora'. Paper presented at the Corpus Linguistics 2007 conference, University of Birmingham/UK. Available: http://www2.lael.pucsp.br/~tony/temp/publications/2007_birmingham.pdf. Accessed 1 May 2010.
- Birke, J. and Sakar, A. (2006): 'A clustering approach for the nearly unsupervised recognition of nonliteral language'. In: Association for Computational Linguistics (ed.) *Proceedings of 11th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 329-336. Available: <http://acl.ldc.upenn.edu/E/E06/E06-1042.pdf>. Accessed 1 May 2010.
- Black, M. (1962): *Models and Metaphors: Studies in language and philosophy*. Ithaca, NY: Cornell University Press.
- Black, M. (1977): 'More about metaphor'. *Dialectica* 31(3/4): 43-57.
- Butler, J. (1990): *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge. (2nd edn 1999)
- Cameron, L. (2003): *Metaphor in Educational Discourse*. London: Continuum.

- Cameron, L. and Deignan, A. (2003): 'Combining large and small corpora to investigate tuning devices around metaphor in spoken discourse'. *Metaphor and Symbol* 18(3): 149-160.
- Cameron, L. and Deignan, A. (2006): 'The emergence of metaphor in discourse'. *Applied Linguistics* 27(4): 671-690.
- Charteris-Black, J. (2004): *Coliticians and Rhetoric*. Basingstoke: Palgrave.
- Charteris-Black, J. (2005): *Corpus Approaches to Critical Metaphor Analysis*. Basingstoke: Palgrave.
- Cicero, M.T. (55 BC/1969): *De Oratore*. Leipzig: B.G. Teubner.
- Coates, J. (1986): *Women, Men and Language*. London: Longman. (2nd edn 1993)
- Coates, J. (2003): *Men Talk: Stories in the making of masculinities*. Oxford: Blackwell.
- Connell, R.W. (1995): *Masculinities*. Cambridge: Polity.
- Connell, R.W. (1998): 'Masculinities and globalization'. *Men and Masculinities* 1(1): 3-23.
- Connell, R. and Messerschmidt, J.W. (2005): 'Hegemonic masculinity: Rethinking the concept'. *Gender and Society* 19(6): 829-859.
- Deignan, A. (1999): 'Corpus-based research into metaphor'. In L. Cameron and G. Low (eds) *Researching and Applying Metaphor*. Cambridge: Cambridge University Press, pp. 177-199.
- Eubanks, P. (2005): 'Globalization, "corporate rule", and blended worlds: A conceptual-rhetorical analysis of metaphor, metonymy, and conceptual blending'. *Metaphor and Symbol* 2(3): 173-197.
- Fairclough, N. (1995a): *Critical Discourse Analysis*. London: Longman.
- Fairclough, N. (1995b): *Media Discourse*. London: Edward Arnold.
- Fairclough, N. (2003): *Analysing Discourse: Textual analysis for social research*. London: Routledge.
- Fass, D. (1997): *Processing Metonymy and Metaphor*. London: Ablex.
- Fauconnier, G. and Turner, M. (2002): *The Way We Think: Conceptual blending and the mind's hidden complexities*. New York: Basic Books.
- Fishman, P. (1983): 'Interaction: the work women do'. In B. Thorne, C. Kramarae and N. Henley (eds) *Language, Gender and Society*. Rowley, MA: Newbury House, pp. 89-101.

- Forceville, C. (1996): *Pictorial Metaphor in Advertising*. London Routledge.
- Grady, J., Oakley, T. and Coulson, S. (1999): 'Blending and metaphor'. In: R. Gibbs and G. Steen (eds) *Metaphor in Cognitive Linguistics*. Amsterdam: Benjamins, pp. 101-124.
- Hardie, A., Koller, V., Rayson, P. and Semino, E. (2007): 'Exploiting a semantic annotation tool for metaphor analysis'. In: M. Davies, P. Rayson, S. Hunston and P. Danielsson (eds) *Proceedings of the Corpus Linguistics 2007 conference*. Available: http://corpus.bham.ac.uk/corplingproceedings07/paper/49_Paper.pdf. Accessed 1 May 2010.
- Holmes, J. (2006): *Gendered Talk at Work*. Oxford: Blackwell.
- Johnson, M. (1987): *The Body in the Mind: The bodily basis of meaning, imagination, and reason*. Chicago: University of Chicago Press.
- Jule, A. (2008): *A Beginner's Guide to Language and Gender*. Clevedon: Multilingual Matters.
- Koller, V. (2004a): 'Businesswomen and war metaphors: "Possessive, jealous and pugnacious"?''. *Journal of Sociolinguistics* 8(1): 3-22.
- Koller, V. (2004b): *Metaphor and Gender in Business Media Discourse: A critical cognitive study*. Basingstoke: Palgrave.
- Koller, V. (2005): 'Critical discourse analysis and social cognition: Evidence from business media discourse'. *Discourse & Society* 16(2): 199-224.
- Koller, V. (2008): 'Brothers in arms: Contradictory metaphors in contemporary marketing discourse'. In: M.S. Zanotto, L. Cameron and M.C. Cavalcanti (eds) *Confronting Metaphor in Use*. Amsterdam: Benjamins, pp. 103-125.
- Koller, V. and Semino, E. (2009): 'Metaphor, politics and gender: A case study from Germany'. In: K. Ahrens (ed.) *Politics, Gender and Conceptual Metaphors*. Basingstoke: Palgrave, pp. 9-35.
- Koller, V., Hardie, A., Rayson, P. and Semino, E. (2008): 'Using a semantic annotation tool for the analysis of metaphor in discourse'. *Metaphorik.de* 5: 141-160.. <http://www.metaphorik.de/15/>. Accessed 1 May 2010.
- Kress, G. and van Leeuwen, T. (2006): *Reading Images: The grammar of visual design*. 2nd edn. London: Routledge.

- Lakoff, G. (1987): *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.
- Lakoff, G. (1993): 'The contemporary theory of metaphor'. In A. Ortony (ed.) *Metaphor and Thought*. 2nd ed. Cambridge: Cambridge University Press, pp. 202-251.
- Lakoff, G. and M. Johnson (1980): *Metaphors We Live By*. Chicago: University of Chicago Press.
- Lakoff, G. and M. Johnson (1999): *Philosophy in the Flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.
- Lakoff, R. (1975): *Language and Woman's Place*. New York: Harper Colophon Books. 2nd revised and expanded edn 2004 (ed. M. Bucholtz). Oxford: Oxford University Press.
- Litosseliti, L. (2006): *Gender and Language: Theory and practice*. London: Hodder Arnold.
- Maltz, D. and R. Borker (1982): 'A cultural approach to male-female miscommunication'. In J. Gumperz (ed.) *Language and Social Identity*. Cambridge: Cambridge University Press, pp. 196-216.
- Martin, J.H. (2006): 'A corpus-based analysis of context effects on metaphor comprehension'. In: A. Stefanowitsch and S.Th. Griess (eds) *Corpus-Based Approaches to Metaphor and Metonymy*. Berlin: de Gruyter, pp. 214-236.
- Mason, Z.J. (2004): 'CorMet: A computational, corpus-based conventional metaphor extraction system'. *Computational Linguistics* 30(1): 23-44. Available: <http://wing.comp.nus.edu.sg/acl/J/J04/J04-1002.pdf>. Accessed 1 May 2010.
- Mullany, L. (2007): *Gendered Discourse in the Professional Workplace*. Basingstoke: Palgrave.
- Ortony, A. (1975): 'Why metaphors are necessary and not just nice'. *Educational Theory* 25: 45-53.
- Partington, A. (1997): *Patterns and Meaning*. Amsterdam: Benjamins.
- Parvatiyar, A. and Sheth, J.N. (2001): 'Conceptual framework of customer relationship management'. In J.N. Sheth, A. Parvatiyar and G. Shainesh (eds) *Customer Relationship Management: Emerging concepts, tools and applications*. New Delhi: McGraw-Hill, pp. 3-25.

- Pragglejaz Group (2007): 'MIP: A method for identifying metaphorically used words in discourse'. *Metaphor and Symbol* 22(1): 1-39.
- Richards, I.A. (1936): *The Philosophy of Rhetoric*. London: Oxford University Press.
- Steen, G. (1999): 'From linguistic to conceptual metaphor in five steps'. In R.W. Gibbs Jr and G.Steen (eds) *Metaphor in Cognitive Linguistic*. Amsterdam: Benjamins, pp. 57–77.
- Steen, G. (2002): 'Towards a procedure for metaphor identification'. .'
In G. Steen (ed.) *Metaphor identification*. Special issue of *Language and Literature* 11(1): 17-34.
- Stefanowitsch, A. (2004): 'HAPPINESS in English and German: A metaphorical-pattern analysis'. In: M. Achard and S. Kemmer (eds) *Language, Culture, and Mind*. Stanford: CSLI, pp. 137-149.
- Stefanowitsch, A. (2007): 'Corpus-based approaches to metaphor and metonymy'. In A. Stefanowitsch and S.Th. Griess (eds) *Corpus-Based Approaches to Metaphor and Metonymy*. Berlin: de Gruyter, pp. 1-16.
- Sunderland, J. (2004): *Gendered Discourses*. Basingstoke: Palgrave.
- Tannen, D. (1990): *You Just Don't Understand: Women and men in conversation*. London: Virago.
- Teng, N.Y. and Sun, S. (2001): 'Metaphoric framing: ROAD as a metaphoric concept in political news in Taiwan'. *Mass Communication Research* 67: 87-112.
- Ungerer, F. (2000): 'Muted metaphors and the activation of metonymies in advertising'. In A. Barcelona (ed.) *Metaphor and Metonymy at the Crossroads: A cognitive perspective*. Berlin: de Gruyter, pp. 321-340.
- van Dijk, T.A. (1993): 'Principles of Critical Discourse Analysis'. *Discourse & Society* 4(2): 249-283.
- van Dijk, T.A. (2003): 'The discourse-knowledge interface'. In G. Weiss and R. Wodak (eds) *Critical Discourse Analysis: Theory and interdisciplinarity*, Basingstoke: Palgrave, pp. 85-109.
- Zimmermann, D. and C. West (1975): "Sex roles, interruptions and silences in conversation". In B. Thorne and N. Henley (eds) *Language and Sex: Difference and Dominance*. Rowley, MA: Newbury House, pp. 105-129.

Martina Pitz
Université Jean Moulin Lyon 3 – CEL, EA 1663

La genèse de la frontière linguistique germano-romane à la lumière des recherches toponymiques et archéologiques

Par son amalgame de langues, de traditions et d'identités culturelles, la population de l'Europe est le résultat de nombreux mouvements migratoires. La période historique qui s'étend de l'expansion de l'Empire romain aux invasions germaniques en est sans doute l'une des plus importantes manifestations. Elle a abouti à une restructuration du continent européen avec l'apparition d'un royaume franc multilingue, mais politiquement uniforme, notamment grâce à l'emploi du latin comme langue administrative. La question des fondements ethniques et linguistiques de ce nouvel empire franc, comprise dans la problématique plus générale de la rupture ou de la continuité culturelle au seuil de l'Antiquité tardive et du haut Moyen âge – le célèbre problème du *clash of cultures* de la recherche anglophone¹ – préoccupent depuis longtemps les spécialistes², car ces fondements constituent les conditions préalables à la formation d'espaces et de frontières linguistiques au cours du haut Moyen âge³.

Par leur expansion dans l'empire romain au cours du IV^e siècle, depuis leurs terres originelles sur la rive droite du Rhin inférieur, les peuplades relativement hétérogènes des Francs entrent en contact avec des populations romanes parlant des idiomes issus du latin provincial des régions concernées. Plus à l'est, les Francs se heurtent à d'autres

¹ Voir notamment : Chrysos/Wood 1999 ; Mathisen 1993 ; Ward-Perkins 2005 ; Wood 1997-2003, 2006.

² Parmi les travaux les plus importants de ces vingt dernières années, on citera notamment Banniard 1989, 2002 ; Geary 2006, 2008 ; Hägermann/Haubrichs/Jarnut 2004 ; Scheibelreiter 1999 ; Werner 1996.

³ Pour cette question, on se reportera à Haubrichs 2007 ; cet article comprend une bibliographie très complète.

peuplades germaniques, installées sur le Rhin moyen. Il convient de souligner que la recherche archéologique contemporaine perçoit ces phénomènes plutôt comme une lente infiltration à partir du milieu du IV^e siècle, qui développa ensuite une dynamique particulière alimentée par les guerres, les conquêtes et les déroutes au cours des deux siècles suivants. On est donc loin de l'idée d'une « invasion barbare » véhiculée par les manuels scolaires⁴, et du point de vue linguistique, le contact des Francs avec les Alamans – et non pas la confrontation de ces deux peuples symbolisée par la bataille de Tolbiac – revêt une importance particulière, du moins pour la genèse du vieux haut allemand⁵. Durant cette période, l'évolution linguistique des régions localisées entre la Somme, la Meuse, la Moselle et le Rhin est caractérisée par une complexité marquante. Ces régions apparaissent pour ainsi dire comme un espace de cristallisation des

⁴ Voir pour cela, entre autres, Brather 2000, 2008. Si mon interprétation des recherches archéologiques et historiques actuelles est correcte, la situation se présente de la manière suivante : depuis la fin du III^e siècle, des groupes de populations relativement hétérogènes sur le plan « ethnique », mais parlant tous des dialectes appartenant au germanique occidental (le « westique » dans la terminologie française) s'emparèrent des terres romaines de la rive gauche du Rhin inférieur où ils finirent par se rassembler en formant la nouvelle tribu des *Franci*. Les attestations historiques permettant de retracer approximativement ces migrations sont rassemblées par de Boone 1954 ; Runde 1998 et Grahn-Hoek 2005, p. 1-5 ; de nouvelles hypothèses sur l'origine de ces populations sont proposées par Wagner 1977 ; Seebold 2000, 2003 et Taayke 2003. L'état actuel des recherches historiques sur la question est résumé chez Grahn-Hoek 2005, qui donne aussi de nombreuses indications bibliographiques. Parallèlement, des pirates d'origine saxonne, dont la langue appartenait également à la branche du « westique », menaçaient les côtes et s'installaient entre autres dans la région de Bayeux et autour de l'embouchure de la Garonne et de la Loire. À partir du IV^e siècle, les Francs s'installèrent durablement sur la Meuse inférieure et aux bords de la rivière de Waal et obtinrent une reconnaissance officielle de leurs conquêtes par l'administration romaine sous Julien l'Apostat, vers 360 ; à partir du témoignage des historiens romains, ce processus est retracé par Drinkwater 1997. Par ailleurs, d'autres types de populations germaniques d'origines très diverses ont été installées par les Romains sur le sol de la Gaule ; d'après les recherches archéologiques récentes, ces groupes ne peuvent plus être considérés comme des *laeti*. Selon Böhme (1998), il s'agirait d'une élite militaire, mais les vraies raisons de sa présence en Gaule n'auraient pas encore été élucidées. De plus, il ne faut surtout pas oublier les nombreux mercenaires servant dans le nouvel *exercitus Gallicanus* issu des réformes de l'armée romaine impulsées par Dioclétien et Constantin. Dès le début du IV^e siècle, d'importants contingents de « barbares » composés notamment de Francs et de Saxons, mais aussi d'autres peuplades germaniques et de cavaliers non germaniques venus des plaines d'Europe orientale, s'infiltrèrent ainsi en Gaule et se répandirent, souvent sous forme de familles entières avec femmes et enfants, dans les provinces romaines de Germanie Seconde, de Belgique Première et Seconde et de Lyonnaise Première. Certains de ces colons militaires furent assez mobiles et retournèrent sur leurs terres d'origine à la fin de leur service ; d'autres s'installèrent durablement, notamment dans les régions situées au nord et à l'est de la Somme, ouvrant ainsi le chemin à une « civilisation mixte », renforcée à partir du milieu du V^e siècle par l'arrivée successive de populations franques.

⁵ Selon nous, l'étude la plus complète sur la question reste celle de Haubrichs 1996 ; ses résultats, et notamment les cartes qui les accompagnent, ont été repris par plusieurs manuels récents destinés aux étudiants.

deux grandes familles linguistiques romane et germanique⁶ : le latin régional des populations autochtones se transforme peu à peu en proto-français, puis en ancien français, et la langue des Francs installés au-delà de la future limite des langues disparaît progressivement – suite à la romanisation de ses locuteurs ou, éventuellement (mais sur ce point, les recherches linguistiques n'en sont qu'au tout début), suite au départ de certains de ces locuteurs vers des régions sises plus à l'est⁷. En même temps, à l'est de la future limite des langues, l'ancien francique évolue également et finit par former un ensemble de dialectes que les germanistes ont coutume d'appeler « l'éventail rhénan », avec une zone bas francique au nord, une zone francique médian au milieu et une zone francique rhénan au sud⁸. On admet aujourd'hui que, du point de vue « ethnique », les Francs n'ont jamais formé un ensemble homogène et que, suite à leurs diverses migrations, les contingents sis plus au nord se seraient petit à petit rapprochés de leurs voisins saxons, frisons et autres installés sur les côtes de la mer du nord, alors que les contingents sis plus au sud se seraient rapprochés des Alamans, ce qui expliquerait la forme curieuse de cet « éventail rhénan »⁹.

Il est facile de démontrer que la notion de *theodisca lingua* utilisée par les sources carolingiennes pour désigner la deuxième langue vernaculaire parlée au sein du royaume franc, parallèlement à la *rustica romana lingua*, transgresse les frontières des ethnies et des aires linguistiques au sens traditionnel du terme¹⁰. Il faut entendre par là les variétés linguistiques de toutes les tribus germaniques qui pouvaient se rencontrer au sein de cet espace, et je reste intimement convaincue que le souci de distinction dans ce domaine apparaît avant tout comme un problème méthodologique des étymologistes, habitués à voir dans les parentés entre les langues une sorte d'arbre généalogique dont les branches se déploieraient du simple vers le

⁶ Voir pour cela, entre autres, Haubrichs 1995 : 186.

⁷ Les données anthroponymiques, et notamment l'analyse des anthroponymes contenus dans les premiers éléments de toponymes d'époque mérovingienne, semblent plaider en faveur de cette hypothèse ; c'était l'un des résultats de nos recherches doctorales (Pitz 1997 : 788-807, 917-927), confirmé depuis par l'analyse d'autres régions, d'autres matériaux et d'autres types toponymiques (Pitz 2000, 2002a, 2002b, etc.), et surtout par une confrontation systématique des données onomastiques et archéologiques (Pitz/Stein 2000, 2008).

⁸ Cf. Meineke/Schwerdt 2008 : 235 ; Pitz/Vollono 2003 ; Ernst 2005 : 92.

⁹ Pour plus de détails, on se reportera à Pitz 2005 : 331.

¹⁰ Une étude très détaillée de l'emploi de ce terme dans les sources latines et vernaculaires est proposée par Haubrichs 2004.

complexe. Cette construction ethnocentriste et évolutionniste ne reflète pas la situation des langues germaniques du haut Moyen Âge, et notamment de l'ancien francique. Il faut, au contraire, envisager cette *lingua theodisca* présente en Gaule romane comme un *continuum* composé d'un certain nombre de « noyaux » de traditions linguistiques qui, un peu à l'image des ensembles enchevêtrés de nos mathématiques modernes, évoluent sous l'effet de processus d'interférence et de différenciation ; mais au sein de cet ensemble, l'intercommunication reste apparemment possible pendant un certain laps de temps, en dépit d'indéniables différences de détail. D'une manière générale, ce type de *continuum* ne se heurte à leurs propres limites que lorsqu'il entre en contact avec d'autres *continua* de langues non apparentées ; et comme nous sommes en droit de supposer que pour les locuteurs de langues germaniques installés sur le territoire de la *Gallo-romania*, ce contact avec la *rustica romana lingua* était pour ainsi dire permanent, notre *continuum* s'est sans doute fragmenté assez vite, tel un habit d'arlequin, en petits îlots linguistiques, ce qui entraîna inévitablement des irrégularités et des innovations. Cet isolement progressif des différents groupes de population franque finit probablement aussi par mettre un frein à l'intercompréhension évoquée plus haut, qui constitue pourtant un préalable indispensable au maintien de la langue maternelle comme marqueur identitaire, ce qui a pu favoriser la romanisation définitive de ces locuteurs.

Certes, cette « mixité » ne favorise guère la tâche du linguiste, d'autant plus que toutes ces variétés apportées dans la *Gallo-romania* par les diverses populations franques et autres depuis leurs territoires d'origine ont dû continuer à évoluer dans leur nouvel environnement roman. Il semble néanmoins assuré que l'identité linguistique de ces régions repose précisément sur la diversité des langues, et notamment sur le plurilinguisme clairement affiché par les couches dirigeantes, réunissant en un véritable « creuset ethnique » des dignitaires francs et des aristocrates romans, avec une population romane d'un côté, largement majoritaire, et une population « théodisque » de l'autre, dans laquelle convergent toutes les bribes de populations germaniques qui avaient été installées en Gaule par l'armée romaine, rejoints par les nouveaux mouvements de « colonisation » impulsés par les Francs. On retiendra que le centre du pouvoir franc se situait originellement à l'ouest, dans la vieille Neustrie ; ce n'est qu'au cours du VI^e siècle que

les Francs s'orientèrent plus vers l'est, vers une Austrasie en gestation encore faiblement peuplée, du moins dans sa frange orientale, par une population gallo-romaine restée au pays. À ce stade de l'évolution, de larges zones bilingues ont dû subsister car avec de fortes diversités régionales, des populations gallo-romaines se sont aussi maintenues à l'est de la future limite des langues où elles formèrent des îlots linguistiques d'importance variée qui ne se dissoudront que progressivement, au fur et à mesure où la « colonisation » franque prendra de l'ampleur. Avec l'établissement d'une frontière linguistique linéaire dont le tracé échappe à toute considération de nature purement géographique, on constate ensuite, au cours des VIII^e et IX^e siècles, une sorte de « démêlage » des langues et un retour progressif au monolinguisme.

Pour les zones germanisées de l'ancien Empire romain, l'étude de ces phénomènes a constitué l'un des axes forts de la recherche en linguistique de contact germano-romane au cours des vingt dernières années. En collaboration étroite avec les archéologues, les linguistes ont développé, pour l'histoire de l'habitat de ces zones germanisées, une approche méthodologique spécifique résolument comparatiste et nourrie d'apports romanistes et germanistes qui permet de saisir de plus en plus précisément le moment où cette ancienne romanité a dû s'effondrer et où s'est effectuée l'assimilation linguistique de la population romane¹¹. Cette méthodologie s'appuie essentiellement sur la chronologie phonétique relative de toutes les langues impliquées dans cette situation de bilinguisme, et devant l'absence de sources écrites, les matériaux analysés sont essentiellement de nature toponymique et anthroponymique, car malgré leurs nombreuses difficultés d'interprétation, les noms propres nous renseignent toujours d'une manière ou d'une autre sur les circonstances extérieures qui ont décidé de leur attribution – en l'occurrence, pour ce qui est des toponymes, sur l'origine des habitats ainsi nommés. Les toponymes peuvent aussi survivre à un, voire à plusieurs changements de langues. La connaissance approfondie de la toponymie d'une région permet ainsi de reconstruire l'évolution linguistique et démographique de manière extrêmement précise, en focalisant les approches de différentes disciplines (géographie, histoire, archéologie, ethnologie, etc.) vers un objectif commun, à savoir l'exploration historique des processus de formation des régions. De façon très sommaire, j'ai tenté

¹¹ Voir pour cela Haubrichs 2003, 2007 ; Pitz 1999, etc.

de rendre compte de ces recherches sur la *Romania submersa* de Rhénanie lors du colloque organisé au Centre d'Études Linguistiques par Brigitte Horiot en décembre 2006¹².

En revanche, les régions galloromanes du nord-est de la France, situées au-delà de la limite des langues, n'ont pas encore fait l'objet d'études approfondies selon cette nouvelle méthodologie. Il s'agit là, avant tout, d'un problème de sources car il n'existe pas de dictionnaires toponymiques véritablement fiables et exploitables tels quels pour la majeure partie de cette zone¹³. Ce type d'études serait pourtant plus que nécessaire car on sait depuis longtemps que l'image des Francs véhiculée durant les premières décennies du XX^e siècle, à savoir celle d'un peuple ayant joué un rôle essentiel pour la genèse des espaces linguistiques au sein même de la *Gallo-romania*, mérite d'être fortement nuancée. En ce qui me concerne, j'adhère de moins en moins à une vision selon laquelle, en dehors du lexique où l'influence du francique est indéniable, la genèse de la langue française serait profondément marquée par l'influence de cette langue. Si cette vision reste encore assez présente aujourd'hui en linguistique historique, c'est bien parce que – tiraillées entre les visions pangermaniques des uns et l'extrême scepticisme des autres –, les recherches sur l'impact de cette situation de bilinguisme n'ont guère avancé depuis des décennies et on n'a pas encore pu déterminer avec suffisamment de précision dans quels domaines de la langue – et par quels types d'interférences – ces influences germaniques ont effectivement pu s'exercer. On concèdera volontiers que les retombées concrètes d'un superstrat ne semblent statistiquement « quantifiables » que pour les seuls secteurs du lexique et de l'onomastique – et encore : si l'existence d'un nombre non négligeable d'emprunts lexicaux et onomastiques aux langues germaniques anciennes n'est plus guère contestée de nos jours, on sait néanmoins que les partisans de la thèse des superstrats ont souvent proposé des solutions étymologiques sans

¹² Cette étude a été publiée sous le titre : « Le couloir romanique du Nord-Est de la Gaule à la lumière des reliquats toponymiques et lexicaux des zones germanisées entre Rhin et Moselle » (pages 227 à 249), dans : Horiot, Brigitte (dir.), *La dialectologie hier et aujourd'hui (1906-2006)*. Actes du colloque de Lyon, 7-9 décembre 2006. Centre d'Études Linguistiques, *Série Dialectologie* 5, 2009, 476 pages.

¹³ En dehors des anciens dictionnaires topographiques, qui renferment de nombreuses erreurs, on ne peut s'appuyer sur des travaux plus récents et mieux conçus que pour une partie de la Picardie (Lambert 1982 ; Malsy 1999-2001), ainsi que sur Gysseling 1960, dont le champ de vision englobe tout le nord-est de la Galloromania et les régions germanophones avoisinantes mais dont les collectes de formes anciennes de toponymes ne comprennent que les attestations antérieures à 1226.

s'appuyer sur un argumentaire suffisamment détaillé, et surtout sans procéder à un examen approfondi de l'étymologie des étymons dans la langue source, donc sans collaborer concrètement avec les germanistes. Depuis la publication, en 1969, du dernier volume du *FEW* consacré aux germanismes, le travail ne progresse qu'à très petits pas et se limite pour ainsi dire à des observations ponctuelles contenues dans une littérature très dispersée, difficile à répertorier. En raison de cet état peu encourageant de la recherche, les manuels de l'histoire de la langue française destinés aux étudiants s'appuient encore très souvent sur les conclusions des travaux de l'avant- et de l'immédiat après-guerre dont les fondements épistémologiques sont pourtant vivement critiqués de nos jours (Pitz 2000, 2005, 2006). Ainsi, une cartographie tirée de Gamillscheg (1938, p. 169), qui avait lui-même repris cette carte dans les *Grundlagen der Volksgeschichte Deutschlands und Frankreichs* d'Adolf Helbok (1937), ouvrage dont les visions pangermaniques ont fait l'objet de critiques acerbes (Oberkrome 1993, p. 130-133 ; Rzoska/Henkes 2003, p. 456-457, etc.) est encore reproduite dans un ouvrage largement utilisé dans l'enseignement supérieur (Lodge 1997, p. 84) ; cette carte montre de vastes zones de colonisation franque plus ou moins continues, idée qui ne semble plus défendable de nos jours. Le récent manuel de Glessgen (2007, p. 320-321) se réfère sans doute directement à Lodge lorsqu'il affirme que les Francs « étaient particulièrement nombreux dans le Nord-Est de la France », ce qui, « si la théorie des superstrats était juste, aurait dû provoquer une différenciation linguistique particulière, nettement plus importante que la simple variation dialectale que l'on peut effectivement observer ». Or, si les recherches récentes en linguistique germano-romane semblent effectivement montrer que l'influence germanique proprement dite n'intervient dans la genèse de ces parlers que dans la seule et unique mesure où des bouleversements sociaux dus à l'intervention de « barbares » ont dû provoquer des attitudes linguistiques bien particulières (Pitz 2003, p. 90), il faut néanmoins garder à l'esprit que du côté des archéologues, l'interprétation des données servant à établir de telles cartographies donne encore lieu à de vifs débats (Brather 2000, 2008 ; Stein 2007, etc.) car ces données sont difficiles à quantifier et difficiles à mettre en rapport avec les données linguistiques. Comme le notent fort judicieusement Patrick Périn et Laure-Charlotte Feffer (1997, p. 195), « il s'agit d'une question délicate » ; « non que les trouvailles

funéraires manquent, bien au contraire, mais parce qu’il importe de savoir avec quelle marge de précision on peut les confronter aux faits historiques ».

Mes recherches personnelles partent de ce constat et tentent de combiner les questionnements des germanistes et des romanistes selon les nouvelles approches de la sociolinguistique historique et de la linguistique de contact¹⁴, en utilisant les reliquats lexicaux, toponymiques et anthroponymiques de la « langue de Clovis » enfouis dans les dialectes gallo-romans, et notamment les germanismes décelables dans la toponymie, dans l’anthroponymie, dans les textes – y compris la documentation en langue latine – et dans les parlers actuels du Nord-Est de la Gaule (domaines lorrain, wallon, picard), comme sources linguistiques pour une tentative de reconstitution – dans ses grandes lignes, car il semble difficile de vouloir espérer plus – de cette langue en vestiges que constituent les variétés « théodisques » des débuts du Moyen Âge en terre romane, définitivement disparues, au plus tard, au cours du X^e siècle avec la romanisation progressive de leurs locuteurs : ces variétés ont-elles connu la seconde mutation consonantique, qui reste le fait phonétique le plus important, mais aussi le plus contesté de l’histoire linguistique de l’allemand (Pitz/Vollono 2003 ; Braune/Reiffenstein 2004, p. 82-95, etc.) ? Ont-elles pu subir un assourdissement précoce de leurs dentales sonores en fin de monème, comme le prétend la recherche récente (Mihm 2004, p. 144-145) et comme on a pu le déceler en ancien bas francique (van Helten 1969, § 52 ; Goossens 1974, p. 66) et en ancien saxon (Gallée/Tiefenbach 1993, § 272) ? Qu’en est-il de la diphtongaison de \bar{o} > *uo* (Braune/Reiffenstein 2004, § 38), du traitement de germ. *ai* (*ibid.*, § 43) et *au* (*ibid.*, § 45), voire éventuellement de la métaphonie devant *i* (*ibid.*, § 51), autres faits phonétiques marquants du vieux haut allemand, ou encore du traitement de *hs* (Franck/Schützeichel 1971, § 114) ou de l’amuïssement de *h* devant *t* (*ibid.*, § 113) qui caractérisent certains dialectes franciques actuels ? En tentant de répondre à ces questions, il est possible de jeter un nouveau regard sur les variétés linguistiques parlées par ces Francs qui ont « envahi » la Gaule, sur leurs origines, leurs déplacements et sans doute aussi sur la genèse proprement dite du vieux haut allemand, à condition d’admettre – ce qui, nous l’avons

¹⁴ Pour le domaine gallo-roman et la linguistique germano-romane, ces recherches n’en sont qu’au tout début, voir p. ex. Banniard 2008, p. 226-227 ; Pitz 2008, p. 229 ; Aquino-Weber/Kristol 2009.

vu, me semble fort probable – que certains contingents de ces Francs occidentaux se sont déplacés depuis la Neustrie vers l’Austrasie et ont donc fortement contribué à construire l’éventail rhénan sus-mentionné, tel qu’il se présente dès l’apparition des premiers textes écrits. Qu’on le veuille ou non : en essayant de dégager les traits phonétiques et morphologiques les plus marquants de cet ancien dialecte westique si mal connu, localisé en quelque sorte aux confins du vieux haut allemand et de l’ancien bas francique, on apporte une pierre essentielle à l’épineux problème de la genèse des langues allemande et néerlandaise actuelles. Ce n’est pas parce que cette passerelle du francique occidental s’est effondrée depuis longtemps et qu’elle ne peut donc être franchie que par ceux qui sont prêts à adopter une approche pluridisciplinaire, qu’on peut indéfiniment continuer à ignorer ces matériaux.

Cette approche comporte donc pour les germanistes un intérêt éminent, tout à fait indépendant du problème de la genèse de l’ancien français qui préoccupe les romanistes, d’autant plus que les étymons franciques de ces transferts, quand ils seront clairement identifiés et analysés selon une perspective conjointe de germaniste et de romaniste, constitueront la seule source valide pour ce dialecte disparu. Il n’existe en effet que peu de témoignages littéraires ; de plus, ils sont assez tardifs et leur interprétation linguistique continue de soulever plus de questions qu’elle n’en résout. Je me contenterai ici de ne citer que deux exemples : la présence, dans un manuscrit conservé au XII^e siècle à Saint Amand (mais dont les différentes mains ne se rattachent pas aux types d’écritures connus pour ce *scriptorium*, cf. Herweg 2002, p. 67-82 ; Balibar 2004), du *Ludwigslied*, chant de la fin du IX^e siècle rédigé dans une langue dans laquelle la seconde mutation consonantique est accomplie et qu’on qualifierait donc volontiers de francique rhénan, constitue un problème non élucidé depuis des décennies, d’autant plus que ce texte a été fixé sur le parchemin par le même scribe, forcément bilingue, à qui nous devons également la *Cantilène de Sainte Eulalie*. D’autre part, la langue des *Pariser Gespräche*, collection de traductions-types en vieux haut allemand de mots et de petites phrases en *rustica lingua romana* essentiels pour la vie quotidienne et jetés sur le parchemin à la fin du IX^e ou au début du X^e siècle par un scribe sans doute issu de la région de Sens (Haubrichs/Pfister 1989), comprend également de nombreux traits qui la relie aux dialectes du nord, donc à l’actuel

néerlandais. Mais comme dans le *Ludwigslied*, la mutation consonantique y est accomplie, ce qui exclut toute localisation dans l'ancien bas francique.

Résoudre l'énigme de ces textes, qui ne correspondent à aucun des dialectes actuels de l'ancien espace habité par les Francs, est une chose, mais quoi qu'il en soit, il semble important de garder à l'esprit que, comparés aux transferts lexicaux et onomastiques, ces textes littéraires sont d'un intérêt mineur pour la reconstitution de la *lingua theodisca* parlée sur le sol de l'ancienne Gaule. Car les emprunts sont par définition les reflets immédiats de la langue parlée ; ils ont été transmis de bouche à oreille, et ce constat reste aussi valable lorsqu'on admet – à juste titre, me semble-t-il – qu'il faut aussi compter sur des variations diastratiques entre les différents types de locuteurs et entre les nombreux îlots linguistiques éparpillés. Il est tout à fait possible que les différents mots qualifiés de germanismes par le *FEW*, le *TLFi* et les différents dictionnaires du latin médiéval nous aient été transmis par des groupes sociaux très différents au sein du magma de la population « théodisque » de Gaule, sans parler des nombreux mots que l'on ne retrouve qu'à travers la microtoponymie ou à travers le dépouillement systématique de textes documentaires (chartes, livres de comptes, rapports de droits, terriers, etc.), non répertoriés par la lexicographie générale. Cela ne change cependant rien à la relation « brisée » des textes littéraires à l'égard de cette oralité historique qui constitue un axiome de base dont on ne tient pas toujours suffisamment compte, y compris dans les approches scriptologiques et variationnelles les plus récentes. Il existe toujours une sorte de filtre linguistique entre les deux, et même si ce filtre agit de manière différente chez chaque scribe et dans chaque texte différent, l'oral et l'écrit ne se situent jamais au même plan, l'écriture impliquant obligatoirement un effort de « normalisation » de la langue. Dans ce contexte, il ne fait aucun doute que les emprunts « non filtrés », mais adaptés au goût et à la bouche de la population romane environnante, constituent une meilleure source pour la reconstitution du système phonétique et morphologique de leur langue source, car les règles de cette adaptation linguistique mise en œuvre lorsqu'on emprunte un terme étranger peuvent être décrites assez facilement, en s'appuyant sur les principes bien connus de la grammaire comparée, alors que le rapport complexe entre l'usage linguistique quotidien et la

scripturalité littéraire reste difficile à cerner sur le plan conceptuel et théorique.

Dans cette démarche résolument comparatiste, je reconnais sans hésiter qu'il est parfois très difficile de prouver de façon concluante l'effet de l'interaction concrète de deux systèmes linguistiques radicalement différents sur le plan phonétique, morphologique, syntactique ou sémantique ; ce faisant, on s'avance, en effet, bien souvent dans une espèce de *terra incognita*, mais il convient aussi de souligner que c'est avant tout en raison d'une approche épistémologique insuffisante que ce type d'études sont encore trop souvent considérées comme un mirage. S'il est vrai qu'en son temps, Antoine Meillet (1925, p. 11) n'avait de cesse de répéter que « la comparaison est le seul instrument efficace dont dispose le linguiste pour faire l'histoire des langues », on admettra volontiers qu'une approche méthodologique résolument menée dans l'optique du contact des langues n'a toujours pas été développée de nos jours, du moins pour ce qui est des dialectes gallo-romans et des différents substrats et superstrats qui les ont enrichis. Quoiqu'elles reconnaissent sans détour l'existence de « trois types de mécanismes », à savoir « l'analogie [...], la réanalyse [...] et l'emprunt » (Marchello-Nizia 2006, p. 67) et admettent comme « facteurs déclencheurs d'un changement » une conjonction particulière entre les « *causes* (objectives, extérieures au locuteur, internes au système parfois) » et les « *motivations* (propres au locuteur dans son usage de la langue) » (*ibid.*, p. 68) d'une innovation linguistique, les recherches actuelles centrées sur le français ont souvent tendance à reléguer à l'arrière-plan cette question du contact ; les *causes* étant considérées comme des facteurs socio-pragmatiques inabordables par le seul biais de la modélisation linguistique, on se concentre sur la question des *motivations*, donc sur le rôle du locuteur « lambda » ou, éventuellement, du scribe, véritable point de mire de la linguistique variationnelle à laquelle ces recherches sont étroitement liées¹⁵. Il me semble qu'il y a là un réel « lack of conceptual unity », pour reprendre une expression due à Frans van Coetsem (2000, p. 51), comme le confirme aussi Christiane Marchello-Nizia (2006, p. 97-98) dans sa synthèse récente : « Le phénomène d'emprunt a été empiriquement repéré sans ambiguïté

¹⁵ Cf. par exemple Völker 2009, p. 45-46 qui, pour répondre à la question de savoir « comment ces niveaux interne et externe se manifestent [...] dans le travail pratique en linguistique variationnelle », évoque les analyses scriptologiques postulant « le contact de deux variétés co-présentes au sein de la conscience linguistique d'un locuteur/scribe ».

dans un grand nombre de langues, il est indubitablement important. Mais comme la notion d'analogie, celle d'emprunt manque d'une assise théorique rigoureuse ; en l'état actuel de son élaboration théorique, sa description reste largement ad hoc. Cependant, quelques jalons sont désormais posés ; il existe des tentatives pour élaborer une échelle d'empruntabilité des termes et des notions [...] et, complémentaiement, pour définir quels sont les traits, dans les langues, qui sont particulièrement stables et donc ne donneront probablement pas lieu à un remplacement par emprunt à un autre système [...]. En effet, [...] la question de l'emprunt ne peut pas être examinée seule ; elle doit l'être en relation avec le caractère plus ou moins modifiable des traits de la langue 'emprunteuse' ; car plus un trait est réputé stable, moins il a de chance d'être remplacé par un trait emprunté ».

Pour pénétrer dans cette *terra incognita* du bilinguisme du haut Moyen âge et de la genèse de la frontière linguistique germano-romane qui a fini par mettre un terme à cette situation si difficile à évaluer, un seul argument est susceptible de nous guider car de nombreuses études menées à partir de situations actuelles de contact des langues en ont clairement démontré la pertinence : toute forme d'innovation, y compris lorsqu'elle est provoquée par des contacts avec une langue seconde, ne se produit qu'à une seule condition : elle doit être utile. Elle répond donc à une nécessité précise et correspond aux intentions culturelles des locuteurs (Pitz 2003, p. 107). Par pragmatisme, l'emprunt apparaît là où les conditions préalables lui sont favorables. Dans un souci d'acculturation et d'assimilation de deux langues et de deux cultures fort éloignées l'une de l'autre, les locuteurs procèdent toujours à une recherche active de similitudes dans les deux systèmes linguistiques impliqués, ce qui permet de jeter des ponts. La plupart du temps, l'interférence s'accroche donc à des phénomènes qui existaient déjà, souvent de manière latente ou isolée, dans l'une et l'autre langue, et les met en valeur d'une manière ou d'une autre. Si ces parallélismes fonctionnent indéniablement comme des vecteurs d'intégration, ils mettent aussi en place des mécanismes de sélection qui renforcent parfois les différences antérieures et finiront donc inévitablement par amener une différenciation plus prononcée des différentes aires linguistiques en fonction de l'intensité de ce contact avec l'autre langue¹⁶. Par conséquent, on ne se trompera

¹⁶ Pour plus de détails, voir Pitz 2007.

sans doute pas en affirmant qu'en dehors d'un important apport lexical – qui, ne l'oublions pas, est un phénomène essentiellement culturel, ne nécessitant aucun bilinguisme actif – le superstrat germanique a surtout renforcé certaines particularités qui existaient déjà dans la langue latine, mais de façon non dominante, et qui prennent ainsi de l'ampleur sous l'effet du contact. À titre d'exemple, on pourra citer le cas de l'antéposition de l'adjectif, phénomène attribué au superstrat germanique par de nombreux manuels d'histoire de la langue française, mais dont l'existence dans certains registres de la langue latine est tout à fait assurée, et ceci pour des périodes bien antérieures aux « grandes invasions » (cf. Pitz 2003, p. 102-107).

On peut en conclure que les recherches futures devront, avant tout, porter sur trois catégories de matériaux : les emprunts lexicaux du français à l'ancien francique, les toponymes et les anthroponymes issus de cette langue. Dans le traitement de ces matériaux, une distinction fondamentale doit être établie entre, d'une part, les formations authentiquement germaniques, dont on peut postuler qu'elles ont été entendues et reprises par les populations romanes lors d'un contact direct avec des locuteurs francs, et, d'autre part, les nombreuses formations hybrides, dans lesquelles on retrouve certains lemmes ou certains éléments de formation à étymologie germanique, mais qui ne sont pas en elles-mêmes des germanismes proprement dits. Pour que l'on puisse le considérer comme un germanisme « authentique », tel ou tel reliquat lexical ou onomastique devra, en effet, revêtir les caractéristiques suivantes :

-dans l'idéal, il est attesté dans plusieurs langues appartenant au westique.

-si tel n'est pas le cas, il doit au moins être grammaticalement correct et ne doit renfermer que du matériel lexical attesté dans une ou plusieurs langues appartenant au westique.

-pour les composés et les dérivés, tous les éléments rentrant dans ces composés doivent être d'origine germanique, et la jonction doit se faire selon les règles morphologiques propres aux langues germaniques.

-ces formations doivent aussi être sémantiquement pertinentes ; dans l'idéal, elles se rangent dans une série de formations parallèles avec lesquelles elles peuvent être comparées sans difficulté.

En revanche – et les conséquences de ce constat pour une approche théorique du problème du superstrat ne sauront être sous-estimées –, les très nombreuses formations qui combinent des éléments d'origine germanique et des éléments d'origine romane, et notamment les nombreux dérivés impliquant des suffixes d'origine romane, doivent toutes être attribuées à la population romane. En aucun cas, ces formes ne sauraient être considérées comme des germanismes proprement dits car leur création a dû s'opérer à partir d'un emprunt déjà complètement lexicalisé dans la langue réceptrice, auquel on a par la suite ajouté une composante romane. Doivent ainsi être imputés à la population romane :

- 1) tous les composés dont un seul des deux thèmes est d'origine germanique, l'autre d'origine romane, comme par exemple les nombreux toponymes en *-ville*, *-court*, etc.
- 2) toutes les formations dont l'élément de dérivation est d'origine romane c'est à dire les formations en *-étu*, *-ariu*, *-ellu*, etc., y compris quand le radical renferme un mot dont l'étymon est d'origine germanique.
- 3) Il en va de même pour les composés dont l'ordre des mots ne correspond pas aux règles de formation lexicale des langues germaniques.

Dans tous ces cas, il ne s'agit donc pas de formations directement attribuables aux Francs, mais les composantes germaniques de ces formations hybrides présentent néanmoins un intérêt indéniable pour les germanistes, dans la mesure où elles prouvent que le lexème en question a bien été en usage chez les Francs occidentaux.

Si l'on s'en tient à ces *a priori* méthodologiques, force est de constater qu'il ne reste pas beaucoup de toponymes à caractère purement germanique en zone romane puisque les nombreux noms en *-iacum*, *-villa*, *-curtis*, etc., forgés sur des anthroponymes germaniques, les toponymes formés sur des emprunts lexicaux, les noms formés à l'aide d'ethnonymes, etc. – tout cela doit être résolument écarté. Ne peuvent finalement être retenus comme formations purement germaniques, véritablement formés par des Francs installés en Gaule, que les toponymes dits « primaires » en *-ingen/-ange*, en *-haim*, en *-lar*, etc. : *Mondelange*, *Herserange*, *Bréhain*, *Anlier*, *Longlier*, etc., ainsi que certains toponymes formés à

base d'hydronymes germaniques et à base d'appellatifs qui ne sont pas attestés comme emprunts lexicaux. À titre d'exemple, on citera les cas d'*Atton* et d'*Harouey* en Meurthe-et-Moselle qui remontent respectivement à **Stadōne* (vha. *stado* « rive ») et **Hari-weg* « chemin emprunté par l'armée » ou *via publica*).

Dans les faits, il faut aussi garder à l'esprit qu'il est souvent difficile d'amener la preuve que les bases de certains toponymes de type appellatif ne sont pas attestées sous forme d'emprunts lexicaux dans la langue romane, car la microtoponymie est encore largement sous-exploitée pour ces questions. L'analyse de la microtoponymie fait régulièrement apparaître de nouveaux germanismes, jamais enregistrés par la lexicographie gallo-romane : Jacques Chaurand (2008, p. 280) a récemment découvert tout à fait par hasard un exemplaire particulièrement fascinant dans la microtoponymie des Ardennes, à savoir le lieu-dit *paure* « marécage desséché ». Le mot germanique ayant servi à former ce toponyme correspond au néerlandais *Polder* « terrain gagné sur la mer », mais on reconnaît facilement que le terme apparaît ici avec un sémantisme bien plus ancien et bien moins spécialisé que ne l'est le terme technique conservé en néerlandais et en allemand, qui se rapporte à une technique d'endiguement très particulière. Avec ce sémantisme très général de « marécage asséché », ce mot n'est pas non plus attesté dans les langues germaniques ; l'emprunt français nous a donc sans doute fourni le seul témoignage subsistant du sens originel de ce terme en ancien francique.

En tant que telle, cette rareté des formations authentiquement germaniques n'a d'ailleurs rien de surprenant, compte tenu de la situation de bilinguisme postulée pour ces régions. Il paraît essentiel de préciser que du point de vue purement linguistique, les nombreuses formations du haut Moyen âge en *-ville*, *-court*, etc., qui fourmillent dans tout le nord-est de la France, sont des formations romanes forgées sur des anthroponymes certes d'origine germanique, mais des anthroponymes empruntés par la population romane, contrairement à une vision encore trop largement répandue qui cherchait à interpréter ces toponymes comme des créations germaniques proprement dites – ce qui reviendrait à postuler, pour la Picardie par exemple, un taux de plus de 70% de noms de lieux d'origine germanique. Mais ce constat ne signifie nullement qu'il n'y ait pas eu d'individus d'origine franque parmi les habitants des localités correspondantes et qu'aucune forme

de bilinguisme n'ait existé dans les secteurs correspondants. Il ne semble pas impossible qu'un bon nombre de ces toponymes aient disposé d'un doublet francique en usage chez les populations franques habitant ces contrées ; les études menées dans les contours immédiats de l'actuelle limite des langues montrent bien que la formation de doublets était monnaie courante. Il faut tout simplement se rendre compte qu'en terre romane, ces doublets franciques (généralement des toponymes forgés sur le même anthroponyme, mais à l'aide d'un déterminé ou d'un suffixe germanique comme *Thionville/Diedenhofen* ou *Bouzonville/Busendorf*, par exemple) ont forcément disparu lors de la romanisation des Francs. On peut imaginer que ces doublets existaient un peu partout, mais d'un point de vue épistémologique, rien ne nous permet aujourd'hui d'en prouver l'existence lointaine si elles ne sont pas expressément attestées par les sources écrites. Il ne nous reste en fait que les formes romanes, celles qui ont subsisté après le retour des anciennes zones bilingues au monolinguisme (Pitz 2005). Même si elles ne nous apprennent rien de précis sur la présence concrète d'une éventuelle population franque, ces formes romanes n'en sont pas moins précieuses, car elles nous permettent d'en savoir un peu plus sur les processus d'appropriation de l'anthroponymie germanique par les populations romanes, voire – qui sait – sur les raisons profondes qui ont pu amener ces populations romanes à abandonner leur propre système anthroponymique hérité des Romains au profit d'un autre, emprunté aux Francs, phénomène d'autant plus surprenant, d'ailleurs, qu'il s'agit là d'un des rares domaines où la population franque, pourtant largement minoritaire, a réussi à imposer son mode de fonctionnement, face à une langue et une culture romanes largement dominantes sur beaucoup d'autres plans.

Pour appréhender ce phénomène d'appropriation d'un système anthroponymique étranger, il paraît important d'insister sur le caractère évolutif du processus linguistique et culturel impulsé par la création de l'empire franc. Le passage progressif d'un système anthroponymique vers un autre n'a pu s'opérer que par le biais de mécanismes de sélection dont nous ne savons à l'heure actuelle que peu de choses. Dans ce contexte, il semble probable qu'avant que les anthroponymes germaniques ne finissent par s'imposer, deux types de noms d'origine gallo-romaine aient été particulièrement privilégiés, à savoir (Pitz 2006 ; Pitz/Stein 2008) :

1) les anciens noms latins, celtiques, grecs ou autres qui s'appuyaient sur des traditions chrétiennes et pouvaient donc servir de régulateurs pour contrecarrer la « barbarisation » de la société impulsée par les Francs. On imagine que ces noms ont dû être particulièrement fréquents dans les régions densément peuplées durant l'Antiquité tardive et caractérisées par une forte continuité romane.

2) les noms disposant de modèles phonétiques, morphologiques ou sémantiques dans les deux langues dont, notamment, certains éléments celtiques. Dans un premier temps, l'innovation par l'interférence s'opérant sous l'effet de superstrats germaniques aurait donc été marquée par un recours des acteurs gallo-romans à des procédés spécifiques, susceptibles de faciliter l'adaptation et l'intégration, que leur offrait leur propre langue. Ce ne serait qu'au cours d'une phase ultérieure qu'on verrait apparaître de véritables noms hybrides matérialisant concrètement cette situation d'interférence, puis les noms composés exclusivement d'éléments germaniques, portés par des couches de plus en plus larges de la population romane.

Pour comprendre ce processus, il convient d'abord de se remémorer comment se composent, morphologiquement parlant, les anthroponymes germaniques. Contrairement au système anthroponymique gallo-romain de l'Antiquité tardive, marqué par un abandon progressif du principe des *tria nomina*, du moins dans la réalité quotidienne, et par un accroissement du nombre de *cognomina* et de *supernomina*, qui pouvaient être puisés dans tous les domaines du lexique – ce système était donc particulièrement riche et extensible et impliquait beaucoup de formes dérivées, notamment des diminutifs –, les langues germaniques connaissent, en tout et pour tout, environ 300 thèmes anthroponymiques, et ces 300 thèmes peuvent être rattachés à seulement quatre champs sémantiques : la guerre et les armes, la puissance et le pouvoir, les valeurs fondamentales, le mythe et le culte (Kaufmann 1965, 1968 ; Pitz/Stein 2008, p. 95-96). Ces thèmes étaient ensuite combinés selon des règles bien précises, suivant en principe une logique assez stricte excluant par exemple toute formation de composés ou de dérivés dénués de sens, ainsi que les formations allitérantes ; par ailleurs, les thèmes trisyllabiques et ceux commençant par une voyelle ne pouvaient pas être utilisés comme

seconds éléments, etc. Certains thèmes étaient aussi beaucoup plus fréquents que d'autres ; certaines bases sont même si rares que leur étymologie reste controversée jusqu'à nos jours. Tout cela permet de reconstruire ces noms de personne d'origine germanique de manière résolument linguistique, y compris lorsqu'ils rentrent comme premiers éléments dans un toponyme roman. Par ailleurs, les individus d'origine germanique portaient tous un nom « officiel », généralement bithématique et porteur de « sens » (*Gérard* « la lance forte » ; *Bernard* « fort comme un ours », *Thierry* « celui qui règne sur le peuple », etc.), ainsi que plusieurs « petits noms » qui étaient employés comme variantes de ce nom officiel ; généralement, il s'agissait de formes abrégées ou dérivées du premier thème du nom « officiel » bithématique : *Gero*, *Gerin*, *Gerik*, *Gerilo*, etc., pour *Gérard*, etc. On suppose que l'utilisation de ces formes abrégées se limitait à certains contextes socio-pragmatiques et à certaines situations de communication.

La recherche onomastique ne s'est guère penchée sur la question de savoir comment les populations gallo-romanes entrées en contact avec des populations germaniques ont pu se familiariser avec ce système anthroponymique si éloigné de leurs propres pratiques car il n'y a, à notre connaissance du moins, aucun témoignage écrit sur ce sujet. En ont-ils compris le fonctionnement ? Ont-ils pratiqué une sélection quelconque, et dans ce cas, quelle place ont-ils accordée aux noms « simples » à un seul thème et à cette multitude prodigieuse de « petits noms » en usage chez les Francs ? Pour répondre à cette question, mon amie archéologue Frauke Stein, de l'université de la Sarre (Allemagne) et moi-même avons développé une approche méthodologique résolument comparatiste, et qui a fait ses preuves dans toutes les micro-régions où nous l'avons mise à l'épreuve, avec des nuances fort intéressantes pour certaines régions, comme la Suisse romande, par exemple, où l'on peut compter sur la présence de plusieurs types de populations germaniques. Pour un espace géographique prédéfini, l'archéologue relève et cartographie d'abord toutes les informations relatives à des sites archéologiques d'époque mérovingienne connus dans la région en question et essaie de dater la fondation de ces habitats à partir d'informations livrées par la littérature sur les objets les plus anciens fournis par les cimetières, en particulier sur la forme des tombes et sur la reproduction graphique ou la description des objets (bijoux, objets de parure, armes et vaisselle),

en tenant compte du fait que la fondation de l'habitat peut, éventuellement, remonter à une génération avant la date fournie par l'analyse du mobilier, si les premiers colons étaient relativement jeunes au moment de la fondation de l'habitat et y ont vécu pendant un certain temps avant d'être inhumés sur le site. Parallèlement, l'onomasticienne recense et cartographie systématiquement tous les toponymes de la zone analysée, y compris les noms de villages disparus identifiables à partir certains types de microtoponymes « parlants » en ce sens (notamment les noms en *-ville*, *-court*, *-villers*, *-ménil*, etc.) relevés au moyen d'un dépouillement exhaustif du cadastre napoléonien. L'onomasticienne pose ensuite l'étymologie de tous ces noms. En s'appuyant sur la méthodologie en usage en archéologie mérovingienne pour rattacher les trouvailles archéologiques à un habitat précis – pas plus de 500 mètres entre le site et l'habitat correspondant, le site se trouvant généralement sur un terrain faiblement incliné au-dessus de l'agglomération (Böhner 1958 ; Périn 1983 ; Stein 1989, etc.) –, l'archéologue et l'onomasticienne procèdent ensuite à une première collaboration, visant à identifier le plus précisément possible les habitats correspondant à ces cimetières. Pour la Lorraine romane, les résultats de cette collaboration ont permis de rattacher près de 80% des nécropoles connues d'époque mérovingienne à un habitat existant ou disparu dont nous sommes en mesure d'indiquer le nom¹⁷. Au total, seules 20% environ des nécropoles attestées n'ont pas pu être reliées à un toponyme faisant clairement allusion à un ancien habitat (*-acum*, *-villa*, *-curtis*, etc.), soit parce que la littérature ne donne pas d'indication précise sur la localisation des vestiges, soit parce que l'examen de la microtoponymie n'a pas fourni d'indices. Ces

¹⁷ Concrètement, ces corrélations ont été établies de la manière suivante : Dans tous les cas où la localisation du village actuel par rapport à l'emplacement de la nécropole correspondait aux critères élaborés par les archéologues, la nécropole a été attribuée à cette agglomération. Pour ce faire, les hypothèses de certains collègues archéologues faisant intervenir la possibilité d'un déplacement des villages entre l'époque de la fondation et les temps actuels ont été longuement discutées, sans que nous puissions toutefois y adhérer ; la solution la plus probable, pour le nord de la France, nous semble être celle d'Édith Peytremann (2003, p. 327-337), selon laquelle des déplacements à courte distance, visant à ériger un nouvel habitat à la périphérie de l'ancien, sont tout à fait probables ; on devrait donc compter sur une certaine dispersion à la fin de l'époque mérovingienne, puis, dans le cadre du développement du système bipartite, un nouveau rapprochement suivi de la disparition d'un certain nombre de ces sites périphériques. Quoi qu'il en soit, nous n'avons rattaché au village actuel que les nécropoles situées à moins de 500 mètres du centre du village ; dès que le site était plus éloigné, il a été procédé à un examen détaillé de la microtoponymie afin de retrouver un nom qui, du point de vue typologique, pouvait correspondre à une localité disparue. Dans beaucoup de cas, cette recherche s'est avérée fructueuse.

nécropoles dont le nom n'est pas connu se répartissent néanmoins de façon régulière sur l'ensemble du territoire ; elles ne sont donc pas susceptibles de modifier sensiblement les résultats de notre étude. Pour la Lorraine, l'attribution des cimetières à certains types de toponymes peut se résumer ainsi : 18 % appartiennent à des agglomérations tirant leur nom d'hydronymes très anciens ou d'autres éléments préromains. Près de 20 % appartiennent à des localités dont le nom renferme le suffixe *-iacum*, caractéristique de l'époque romaine. 25 % se rapportent à des formations en *-villa*, près de 6 % en *-curtis*, 2 % en *-villare*. Les autres toponymes rattachés à des nécropoles, par exemple ceux dérivés de lieux-dits gallo-romans et la majorité des hagiotoponymes ne possèdent aucune caractéristique typologique permettant de les dater précisément.

Notre analyse commune s'est ensuite concentrée sur les toponymes dits « primaires », expressément créés pour dénommer un habitat nouvellement fondé, par opposition aux toponymes dits « secondaires », tirés de noms de forêts, de cours d'eau, etc., et attribués ultérieurement à un habitat fondé sur le lieu en question. Parmi ces toponymes « primaires », nous nous sommes particulièrement intéressés à ceux qu'on considère comme des fondations du haut Moyen âge, à savoir les *-villa*, *-les curtis*, etc., caractérisés par la présence d'un nom de personne dans leurs premiers éléments, sans doute le nom du premier propriétaire. Nous avons aussi inclus dans ce corpus les noms en *-iacum* dont le premier élément est un nom de personne d'origine germanique, car ces formations sont forcément postérieures à « l'invasion » barbare ; elles prouvent en fait que le suffixe *-iacum* était encore productif au haut Moyen âge (Buchmüller-Pfaff 1990, p. 27). Après avoir posé l'étymologie de tous ces noms, ils ont été classés en quatre catégories :

- 1) toponymes forgés sur des anthroponymes germaniques à deux thèmes ;
- 2) toponymes forgés sur des anthroponymes germaniques à un seul thème ;
- 3) toponymes forgés sur des anthroponymes germaniques à caractère hypocoristique, en distinguant clairement les dérivés obtenus à l'aide de suffixes germaniques et ceux contenant des suffixes d'origine romane.

4) pour *-villa* et *-curtis* seulement : les toponymes forgés sur des anthroponymes d'origine gallo-romaine.

C'est là que la collaboration avec l'archéologie devient absolument passionnante. En cartographiant tous ces types et en les confrontant avec la carte des sites archéologiques, on obtient des résultats particulièrement surprenants, résultats qu'il n'aurait pas été possible de réaliser sans cette association entre ces deux disciplines. Il s'avère en fait que les localités en *-villa* et en *-curtis* disposant d'une nécropole sont en majorité des formations avec des noms de personne germaniques à deux thèmes. Un petit nombre de toponymes pourvus de nécropoles contiennent des noms de personne germaniques à un thème, mais les noms de personnes à caractère hypocoristique n'apparaissent jamais. Ce résultat, pour le moins étonnant, est confirmé par la répartition géographique des toponymes : au sein des divers noyaux de peuplement définis par la présence de nécropoles, les formations renfermant soit un anthroponyme germanique à deux thèmes, soit un anthroponyme d'origine gallo-romaine, occupent toujours une place privilégiée : soit, elles sont directement reliées à des nécropoles, soit elles se trouvent dans les environs immédiats de ces nécropoles. En revanche, les formations renfermant un nom de personne d'origine germanique monothématique se retrouvent à la périphérie de ces anciens noyaux de peuplement et se regroupent pour ainsi dire autour des bithématiques. On découvre aussi que les formations impliquant des noms de personne d'origine germanique à caractère hypocoristique sont localisées à l'extérieur de ces zones d'habitat les plus anciennes ; elles ne sont pour ainsi dire jamais reliées à des nécropoles. Il apparaît donc très clairement que les formations renfermant des noms de personne germaniques à deux thèmes sont les plus anciennes : elles apparaissent dès le VI^e siècle. En revanche, les noms de personnes à un thème ne rentrent dans la formation de toponymes qu'à partir du VII^e siècle, et l'adoption d'hypocoristiques semble encore plus récente.

Conclusion

Que faut-il en conclure ? Grâce à cette collaboration scientifique, les onomasticiens pourront désormais fonder leurs conclusions sur les étapes successives de ce processus d'adoption de noms germaniques par des populations romanes sur des bases beaucoup plus solides, et ceci pour des époques pour lesquelles nous manquons cruellement de sources écrites et n'aurions pu, sans cette approche pluridisciplinaire, avancer que des hypothèses plus qu'approximatives. Grâce à l'archéologie, il semble maintenant possible de prouver que pendant un certain laps de temps, les variantes monothématiques et hypocoristiques des anthroponymes franciques qui, chez les Francs, faisaient figure de « petits noms » et avaient un caractère plutôt inofficiel, n'étaient pas employées par les populations romanes, du moins pour former des toponymes. Cela revient à supposer que pendant les premières phases de contact entre les deux communautés linguistiques, ces populations romanes saisissaient très bien les mécanismes de fonctionnement du système anthroponymique des Francs, même si cette connaissance s'est perdue par la suite. Ils ont dû comprendre que dans le système anthroponymique de leurs voisins francs, l'usage des noms hypocoristiques différait sensiblement de celui des bithématiques, contrairement aux règles régissant leur propre système anthroponymique hérité de l'Antiquité tardive, où un nom comme *Maurinus* n'était certainement pas perçu comme « petit nom » pour désigner une personne dont le « vrai » nom était *Maurus*. Il ne semble donc pas impossible que les vraies raisons pour lesquelles les populations romanes ont pu ressentir le besoin d'adopter des noms germaniques soient précisément à chercher dans ces différences fondamentales : le système gallo-romain ne disposait pas de noms bithématiques pouvant faire figure de formes « officielles », et aux yeux de l'aristocratie franque, les nombreux dérivés en usage au sein des populations romanes étaient sans doute facilement compris comme des « petits noms », dénués de tout prestige.

Il s'agirait donc d'un phénomène d'adaptation culturelle par nécessité concrète qui concernerait d'abord les noms à deux thèmes et partait sans doute de groupes divers amenés à établir des contacts directs avec leurs voisins « théodisques », pour s'étendre ensuite à d'autres types de populations. L'adoption des noms germaniques à un thème – *Bero*, *Hugo*, *Lando*, *Chlodo*, *Berta*, etc. – par les populations

romanes ne se fera finalement qu’au VII^e siècle. Comme j’ai tenté de le prouver dans une autre étude (Pitz 2009), ce phénomène est sans doute lié au fait qu’à ce moment-clé de l’évolution de la langue romane que constitue le VII^e siècle, en raison de l’amuïssement des voyelles finales en proto-français¹⁸, il n’était plus possible de distinguer de façon claire et nette certains noms masculins d’origine gallo-romaine de leurs correspondants féminins – chose inimaginable pour une population qui se trouvait en contact régulier avec les traditions et le mode de vie très « guerriers » des dirigeants francs. En effet, suite à cet amuïssement des voyelles finales, *Maximus* et *Maxima* devinrent *Maxime* au cas régime, *Leontius* et *Leontia* devinrent *Leonce*, *Eustachius* et *Eustachia* devinrent *Eustache*, etc., et dans cette situation sans doute très embarrassante, ce fut le recours à des formes hypocoristiques en *-o* et en *-io* (*Felicio*, *Paulo*, *Leo*, *Urso*, etc.), renforcées par les nombreuses formes analogues qui existaient chez les Francs et qu’on s’empressait alors d’emprunter (*Hugo*, *Bruno*, *Audo*, etc.) qui sauva les populations romanes de l’embarras, car ces noms prenaient au cas régime une désinence en *-ōne* ou *-āne*, clairement différenciée selon les sexes. Toutes ces observations ne nous incitent pas seulement à nous interroger sur le degré de bilinguisme de ces populations romanes, qui ne se débrouillaient apparemment pas si mal en ancien francique ; elles permettent aussi de poser un nouveau regard sur un problème très délicat de la morphologie historique du français, à savoir la genèse de la soi-disant « déclinaison germanique » *ber/baron* ou *nonne/nonnain*, qui n’est finalement pas germanique du tout puisque le modèle *Felicio*, *Felicione* ou *Felicia*, *Feliciane* existait bel et bien dans la langue latine (Pitz 2009). Une fois de plus, le rôle concret des Francs se cantonnait plutôt dans la mise en place de nouvelles normes linguistiques, découlant de nouveaux comportements sociaux liés à l’inévitable « cohabitation » des deux groupes de populations, normes auxquelles les anciens noms romano-chrétiens répondaient de moins en moins. Mais les processus de romanisation proprement dits de ces anthroponymes germaniques devront faire l’objet d’une autre étude, qui permettra de mettre à l’épreuve les approches comparatistes dont les fondements méthodologiques ont été exposés ici.

¹⁸ Pour la datation de cette évolution, voir Pitz 2008.

Références bibliographiques

- Aquino-Weber, Dorothee/Kristol, Andres (dir.). 2009. *Sociolinguistique historique du gallo-roman. Enjeux et méthodologies*. Berne : Lang.
- Balibar, Renée. 2004. Eulalie et Ludwig. Le manuscrit de la bibliothèque de Valenciennes : colinguisme et prémices littéraires de l'Europe. Cortil-Wodon : E.M.E.
- Banniard, Michel. 1989. Genèse culturelle de l'Europe. V^e-VIII^e siècles. Paris : Seuil.
- . 2002. *Langages et peuples d'Europe. Cristallisation des identités romanes et germaniques (VII^e-XI^e siècles)*. Toulouse : Université de Toulouse II Le Mirail.
- . 2008. La latinophonie ; caractères d'une dynamique vitale. In : Greule, Albrecht *et al.* (dir.). *Studien zu Literatur, Sprache und Geschichte in Europa. Wolfgang Haubrichs zum 65. Geburtstag gewidmet*. St. Ingbert: Röhrig, 215-227.
- Böhme, Horst Wolfgang. 1998. Franken und Romanen im Spiegel gallrömischer Grabfunde im nördlichen Gallien. In: Geuenich, Dieter (dir.). *Die Franken und die Alemannen bis zur Schlacht bei Zülpich*. Berlin/New York : de Gruyter, 31-58.
- Böhner, Kurt. 1958. *Die fränkischen Altertümer des Trierer Landes*. 2 vol., Berlin : Mann.
- Boone, Jan Willem de. 1954. *De Franken van hun eerste optreden tot de dood van Childerik*. Amsterdam: Laporte & Dosse.
- Brather, Sebastian. 2000. Ethnische Identitäten als Konstrukte frühgeschichtlicher Archäologie. *Germania* 78, 139-177.
- . 2008. *Zwischen Spätantike und Frühmittelalter. Archäologie des 4. bis 7. Jahrhunderts im Westen*. Berlin/New York : de Gruyter.
- Braune, Wilhelm/Reiffenstein. Ingo. 2004. *Althochdeutsche Grammatik I : Laut- und Formenlehre*. Tübingen : Niemeyer.
- Chaurand, Jacques. 2008. Compte-rendu de : Guelliot, Octave. 1997-2007. Dictionnaire historique de l'arrondissement de Vouziers. Édition mise au point par Michèle Sanzé et Michel Tamine. 10 vol., Charleville-Mézières : Terres Ardennaises. NRO 49-50, 276-281.
- Coetsem, Frans van. 2000. *A General and Unified Theory of the Transmission Process in Language Contact*. Heidelberg : Winter.
- Chrysos, Evangelos/Wood, Ian (dir.). 1999. *East and West : Modes of communication*. Leiden : Brill.

- Drinkwater, John F. 1997. Julian and the Franks and Valentinian and the Alamanni : Ammianus on Romano-German Relations. *Francia* 24, 1-15.
- Ernst, Peter. 2005. *Deutsche Sprachgeschichte*. Tübingen : Francke.
- Franck, Johannes/Schützeichel, Rudolf. 1971. *Altfränkische Grammatik*. Göttingen : Vandenhoeck & Ruprecht.
- Gallée, Johan Henrik/Tiefenbach, Heinrich. 1993. *Altsächsische Grammatik*. Tübingen : Niemeyer.
- Gamillscheg, Ernst. 1938. *Germanische Siedlung in Belgien und Nordfrankreich*. Berlin : Verlag der Akademie der Wissenschaften.
- Geary, Patrick. 2006. *Quand les nations refont l'histoire*. Paris : Flammarion.
- . 2008. *Franks, Northmen and Slavs. Identities and state formation in the early medieval Europe*. Turnhout: Brepols.
- Glessgen, Martin-Dietrich. 2007. *Linguistique romane. Domaines et méthodes en linguistique française et romane*. Paris : Armand Colin.
- Grahn-Hoek, Heike. 2005. Salii – Franci ipsi – (gentes) qui et Franci. Zur Ethnogenese der Franken und den Anfängen der fränkischen Südwestbewegung bis zum Ende des 4. Jahrhunderts. *Rheinische Vierteljahrsblätter* 69, p. 1-69.
- . 2006. Franci und Francia im 6. Jahrhundert. Zu den historischen Ursachen einer sprachlichen Entwicklung. In: Geuenich, Dieter/Runde, Ingo (dir.). *Name und Gesellschaft im Frühmittelalter*. Hildesheim: Olms, 173-218.
- Gysseling, Maurits. 1960. *Toponymisch woordenboek van België, Nederland, Luxemburg, Noord-Frankrijk en West-Duitsland (vóór 1226)*. 2 vol., Bruxelles : Belgisch Interuniversitair Centrum voor Neerlandistiek.
- Hägermann, Dieter/Haubrichs, Wolfgang/Jarnut, Jörg (dir.). 2004. *Akkulturation. Probleme einer germanisch-romanischen Kultursynthese in Spätantike und frühem Mittelalter*. Berlin/New York: de Gruyter.
- Haubrichs, Wolfgang. 1995. Volkssprache und volkssprachige Literaturen im lotharingischen Zwischenreich (9.-11. Jahrhundert). In: Herrmann, Hans-Walter/Schneider, Reinhard (dir.). *Lotharingia. Une région au centre de l'Europe autour de l'an Mil*. Sarrebruck : sdv, 181-244.

- . 1996. Sprache und Sprachzeugnisse der merowingischen Franken. In: Wiczorek, Alfred/Périn, Patrick (dir.). *Die Franken – Wegbereiter Europas*. Mannheim : Philipp von Zabern, 559-573.
- . 2003. « Die verlorene Romanität im deutschen Sprachraum ». In : Ernst, Gerhard *et al.* (dir.). *Histoire linguistique de la Romania*. Berlin/New York : de Gruyter, 695-709.
- . 2004. *Theodiscus*, Deutsch und Germanisch – drei Ethnonyme, drei Forschungsbegriffe. Zur Frage der Instrumentalisierung und Wertbesetzung deutscher Sprach- und Volksbezeichnungen. In: Beck, Heinrich *et al.* (dir.). *Zur Geschichte der Gleichung „germanisch-deutsch“*. Berlin/New York : de Gruyter, 199-227.
- . 2007. L'espace physique, l'histoire, la langue. L'élaboration des zones de contact et des frontières linguistiques entre Romania et Germania, entre la Suisse et le Luxembourg. In : Lienard, Thomas (dir.). *Construction de l'espace au Moyen Âge : pratiques et représentations*. Paris : Publications de la Sorbonne, 167-191.
- ./Pfister, Max. 1989. *In Francia fui. Studien zu den romanisch-germanischen Interferenzen und zur Grundsprache der althochdeutschen ‚Pariser (Altdeutschen) Gespräche nebst einer Edition des Textes*. Wiesbaden : Steiner.
- Helbok, Adolf. 1937. *Grundlagen der Volksgeschichte Deutschlands und Frankreichs*. Berlin : de Gruyter.
- Helten, Willem Lodewijk van. 1969. *Die altostniederfränkischen Psalmenfragmente. Die Lipsius'schen Glossen und die altsüdmittelfränkischen Psalmenfragmente*. New York, London: Johnson Reprint Corporation.
- Herweg, Matthias. 2002. *Ludwigslied, de Heinrico, Annolied. Die deutschen Zeitdichtungen des frühen Mittelalters im Spiegel ihrer wissenschaftlichen Rezeption und Erforschung*. Wiesbaden : Reichert.
- Kaufmann, Henning. 1965. *Altdeutsche Rufnamen*. Munich : Fink.
- . 1968. *Ergänzungsband zu Ernst Förstemann : Personennamen*. Munich : Fink.
- Keller, Rudi 2003. *Sprachwandel*. Tübingen : Francke.
- Lambert, Emile. 1982. *Dictionnaire topographique du département de l'Oise*. Amiens : Musée de Picardie.
- Lodge, R. Antony. 1997. *Le français. Histoire d'un dialecte devenu langue*. Paris : Fayard.

- Malsy, Jean-Claude. 1999-2001. *Les noms de lieu du département de l'Aisne*. 3 vol., Paris : Société Française d'Onomastique.
- Marchello-Nizia, Christiane. 2006. *Grammaticalisation et changement linguistique*. Bruxelles : De Boeck.
- Mathisen, Ralph Whitney. 1993. *Roman Aristocrats in Barbarian Gaul. Strategies for Survival in an Age of Transition*. Austin: University Press.
- Meillet, Antoine. 1925. *La méthode comparative en linguistique historique*. Paris : Champion.
- Meineke, Eckhard/Schwerdt, Judith. 2001. *Einführung in das Althochdeutsche*. Paderborn : Schöningh.
- Mihm, Arend. 2004. Zur Geschichte der Auslautverhärtung und ihrer Erforschung. *Sprachwissenschaft* 29, 133-206.
- Oberkrome, Willi. 1993. *Volksgeschichte. Methodische Innovation und völkische Ideologisierung in der deutschen Geisteswissenschaft 1918-1945*. Göttingen : Vandenhoeck & Ruprecht.
- Périn, Patrick. 1983. Le peuplement du diocèse de Reims à l'époque mérovingienne. In : Lohrmann, Dietrich (dir.). *Villa – curtis – grangia. Landwirtschaft zwischen Loire und Rhein von der Römerzeit zum Hochmittelalter*. Munich : Artemis, 519-533.
- Périn, Patrick/Feffer, Laure-Charlotte. 1997. *Les Francs*. Paris : Armand Colin.
- Peytremann, Edith. 2003. *Archéologie de l'habitat rural dans le Nord de la France du IV^e au XII^e siècle*. Saint-Germain-en-Laye : Association française d'archéologie mérovingienne.
- Pitz, Martina. 1997. *Siedlungsnamen auf –villare (-weiler, -villers) zwischen Mosel, Hunsrück und Vogesen. Untersuchungen zu einem germanisch-romanischen Mischtypus der jüngeren Merowinger- und der Karolingerzeit*. 2 vol., Sarrebruck : sdv.
- . 1999. Toponymie zwischen den Sprachen. Ortsnamen als Instrumente landes- und siedlungsgeschichtlicher Forschung im lothringischen Sprachgrenzraum, in: Brogiato, Heinz-Peter (dir.). *Geographische Namen in ihrer Bedeutung für die landeskundliche Forschung und Darstellung*. Trèves : Dokumentationszentrum für Deutsche Landeskunde, 67-95.
- . 2000a. Personennamen in frühmittelalterlichen Siedlungsnamen. Methodische Überlegungen am Beispiel der –villare-Namen des Saar-Mosel-Raumes, in: Heinrich TIEFENBACH/Heinrich LÖFFLER

- (dir.), *Personennamen und Ortsnamen*. Basler Symposion 6./7. Oktober 1997, Heidelberg: Winter, 143-188.
- . 2000b. Le superstrat francique dans le Nord-Est de la Gaule. Vers une nouvelle approche philologique et toponymique, in : *Nouvelle Revue d'Onomastique* 35/36 (2000), 69-85.
- . 2002a. Nouvelles données pour l'anthroponymie de la Galloromania. Les toponymes mérovingiens du type *Avricourt*, in: *RLiR* 66, 421-449.
- . 2002b. Namenbildung in Interferenzräumen. Romanische und 'westfränkische' Personennamen in den ältesten Schichten fränkischer Siedlungsnamen des lothringischen Sprachgrenzgebiets, in: ANA ISABEL BOULLÓN AGRELO (dir.). *Actas do XX congreso internacional de ciencias onomásticas, Santiago de Compostela, 20-25 Setembro 1999*. Coruña 2002, p. 1219-1235.
- . 2003. Innovations du centre et archaïsmes du nord-est: fruits du contact des langues en Gaule mérovingienne? Considérations sur le nord-est du domaine d'oïl dans la perspective d'une linguistique de contact, in : *Vox Romanica* 62, 86-113.
- . 2005a. Kulturelle Umorientierung als Motor sprachlicher Innovation? Überlegungen zur Einführung ‚südlicher‘ Formen in der deutschsprachigen Urkundenüberlieferung des 14. Jahrhunderts im Saar-Mosel-Raum, in: Kurt GÄRTNER / Günter HOLTUS (dir.), *Überlieferungs- und Aneignungsprozesse im 13. und 14. Jahrhundert auf dem Gebiet der westmitteldeutschen und ostfranzösischen Urkunden- und Literatursprachen*. Referate des dritten internationalen Trierer Urkundensprachenkolloquiums vom 20.-22. Juni 2001, Trèves: Philipp von Zabern Verlag, 2005, 323-358.
- . 2005b. Franz Petri et l'idée du nivellement toponymique en Gaule mérovingienne, in: *Rivista Italiana di Onomastica* 11, 341-370.
- . 2006. Les toponymes en *-iācum*, *-iāca(s)* forgés sur des anthroponymes germaniques : Approche méthodologique d'un phénomène controversé. L'exemple de l'Auxerrois, in: Eva BRYLLA / Mats WAHLBERG (dir.), *Proceedings of the 21st International Congress of Onomastic Sciences*, vol. 2, Uppsala: Språk- och folkminnesinstitutet, 2006, 245-268.
- . 2007. Superstratsprachen. *Reallexikon der germanischen Altertumskunde*. 35, 38-42.

- . 2008. L'amuissement des voyelles finales en galloroman. Nouveaux indices chronologiques fournis par la toponymie des zones de contact germano-romanes, in : Johannes KRAMER *et al.* (dir.), *Zur Bedeutung der Namenkunde für die Romanistik*. Akten des XXV. Romanistischen Kolloquiums am 24. und 25. Juni 2006 in Trier, Tübingen: Narr, 2008, 171-194.
- . 2009. Le type *Charles-Charlon* ou *Eve-Evain* en ancien français – vraiment un reflet du superstrat germanique ? *NRO* 51, 175-200.
- /Stein, Frauke. 2000. Genèse linguistique d'une région frontalière: Les environs de Forbach et de Sarreguemines. *Les Cahiers Lorrains*, 365-412.
- . 2008. L'adoption d'anthroponymes germaniques par les populations romanes de la Gaule mérovingienne à la lumière des données archéologiques : L'exemple de la Lorraine romane, in : *NRO* 49-50, 83-114.
- /Vollono, Maria. 2003. Die ‚zweite oder hochdeutsche Lautverschiebung‘ – ein obsolet gewordener Begriff für ein allzu komplexes konsonantisches Phänomen? Anmerkungen zu einer neuen Publikation. *Rheinische Vierteljahrsblätter* 67, 313-332.
- Pohl, Walter. 1998. Telling the Difference. Signs of Ethnic Identity. In: *id./Reimitz, Helmut* (dir.). *Strategies of Distinction. The Construction of Ethnic Communities*, 300-800. Leiden : Brill, 17-69.
- . 2004. *Die Suche nach den Ursprüngen. Von der Bedeutung des frühen Mittelalters*. Vienne: Österreichische Akademie der Wissenschaften.
- Runde, Ingo. 1998. Die Franken und die Alemannen vor 500. Ein chronologischer Überblick. In: Geuenich, Dieter (dir.). *Die Franken und die Alemannen bis zur ‚Schlacht bei Zülpich‘*. Berlin/New York : de Gruyter, p. 656-690.
- Scheibelreiter, Georg. 1999. *Die barbarische Gesellschaft. Mentalitätsgeschichte der europäischen Achsenzeit. 5.-8. Jahrhundert*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Seebold, Elmar. 2000. Wann und wo sind die Franken vom Himmel gefallen? *Beiträge zur Geschichte der deutschen Sprache und Literatur* 122, 40-56.
- . Die Herkunft der Franken, Friesen und Sachsen. In: Taayke, Ernst *et al.* (dir.) *Essays on the Early Franks*. Eelde : Barkhuis, 24-34.

- Stein, Frauke. 1989. Die Bevölkerung des Saar-Mosel-Raumes am Übergang von der Antike zum Mittelalter. Überlegungen zum Kontinuitätsproblem aus archäologischer Sicht. *Archaeologia Mosellana* 1, 89-195.
- . 2007. Ergebnisse zur Interferenz zwischen Franken und Romanen im frühen Mittelalter anhand des Gräberfeldes bei Audun-le-Tiche (F, dép. Moselle). In: Pitz, Martina *et al.* (dir.). *Grenzen ohne Fächergrenzen. Interdisziplinäre Annäherungen*. St. Ingbert : Röhrig, 403-438.
- Völker, Harald. 2009. La linguistique variationnelle et la perspective intralinguistique. *RLiR* 73, 27-76.
- Wagner, Norbert. 1977. Die Herkunft der Franken. *Frühmittelalterliche Studien* 11, 218-228.
- Ward-Perkins, Bryan. 2005. *The fall of Rome and the end of civilization*. Oxford: University Press.
- Werner, Karl-Ferdinand. 1996. La ‚conquête franque‘ de la Gaule. Itinéraires historiographiques d’une erreur. In : *Bibliothèque de l’École des Chartes* 154, 7-45.
- Wood, Ian. 1997-2003. *Transformation of the Roman World*. 14 vol., Leiden: Brill.
- . 2006. Transformation of the Roman World. *Reallexikon der germanischen Altertumskunde* 31, 132-134.

Manuel Jobert
Université Jean Moulin – Lyon 3 – CEL, EA 1663
& CREA – EA 370 (Université Paris 10 Nanterre)

Estuary English (EE) : rupture ou continuité ?

Estuary English [...], this modern version of London speech is the pantomime villain for parents who are desperate for their children to speak 'properly' and for earnest defenders of regional accents alike.
Andrew Taylor (2006 : 212).

Introduction

On estime que l'anglais compte aujourd'hui entre 320 et 380 millions de locuteurs natifs (Crystal : 1997) auxquels il faut ajouter entre 150 et 300 millions de locuteurs dans les pays où l'anglais joue un rôle institutionnel (pays du Commonwealth). On imagine aisément la variété dialectale et les différences d'accents qu'un tel nombre de locuteurs peut représenter.

Si l'on se concentre sur l'Angleterre, en laissant de côté les pays celtes, on constate qu'entre 50 et 60% (Wells : 1982) de la population parle avec des traces d'accent du nord et qu'entre 3% à 5% seulement parle avec un accent dit « standard » ou R.P. (*Received Pronunciation*) (Hughes *et al* : 2005). Les Anglais ont une capacité très grande à localiser l'origine géographique de leurs compatriotes (le *Geordie* de Newcastle, le *Brummie* de Birmingham, le *Scouse* de Liverpool, le *Tyke* du Yorkshire et, bien sûr, le *Cockney* de Londres pour ne citer que les plus connus). Il existe de nombreuses séries télévisées explicitement fondées sur cette appartenance linguistico-géographique : *Eastenders*, la plus ancienne (qui débuta en 1985), (je te propose « diffusée à partir de 1985 ») pour le Cockney et *Two Pints*

of *Larger and a Packet of Crisps*, la plus récente (débutée en 2001), (idem) pour l'accent de Runcorn, au sud-est de Liverpool. Une telle exposition médiatique de la variation contribue bien sûr à une plus grande tolérance linguistique sans toutefois la généraliser. Tant que l'accent régional est confiné à des programmes bien particuliers, la tolérance est de mise. Si le présentateur du journal télévisé avait un accent régional très marqué, des lettres de protestation indignées afflueraient à la BBC.

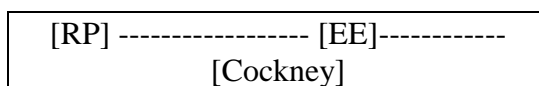
Londres, contrairement à New York, a toujours été une source d'innovations linguistiques et culturelles. L'*Estuary English* y trouve son origine. On tentera d'établir ce qu'est ce nouvel accent par rapport à la R.P. - dont certains prétendent qu'il est le remplaçant - et au Cockney dont il serait une version plus acceptable. Dans un deuxième temps, on tentera de tracer les contours de l'EE d'un point de vue syntaxique, lexical et phonologique. Enfin, on évaluera son influence sociolinguistique à Londres par rapport à la prononciation standard.

1. R.P., Cockney, Estuary English

Le terme *Estuary English* fut créé par David Rosewarne dans un article publié dans le *Times Education Supplement* du 19 octobre 1984 et véritablement popularisé par Paul Coggle dans un livre destiné au grand public de 1993, *Do You Speak Estuary ?* Pour Rosewarne (1984), l'anglais de l'estuaire de la Tamise est un accent régional mélangeant l'accent dit standard ou R.P. et des traits caractéristiques des accents du sud-est de l'Angleterre.

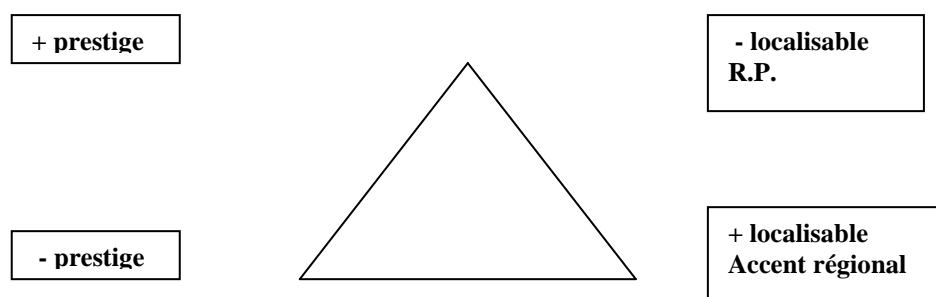
Estuary English is a variety of modified regional speech. It is a mixture of non-regional and local south-eastern English pronunciation and intonation. If one imagines a continuum with R.P. and London speech at either end, EE speakers are to be found grouped in the middle ground.

On peut représenter la situation linguistique de Londres et des alentours à l'aide d'un *continuum* avec, d'un côté, l'accent le plus prestigieux qu'est la R.P. et le Cockney de l'autre. L'EE se situerait entre les deux :

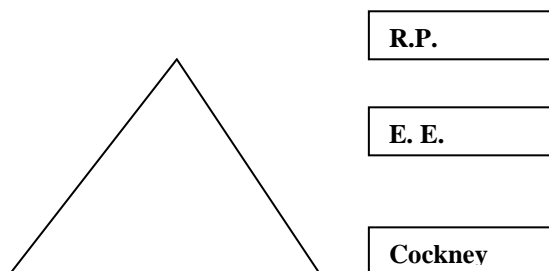


Cette présentation de la situation linguistique implique l'existence autonome de l'EE comme accent distinct à la fois de la R.P. et du Cockney. Cette conception de l'EE a été remise en cause de manière assez timide par la communauté linguistique, à l'exception de Maidment (1998), et le terme de EE, bien qu'il soit critiqué, fait maintenant partie du langage courant de l'anglais, comme le prouve son introduction dans le LPD – 1990 de Wells.

Il est, à ce stade, bon de rappeler la manière dont accent et statut social sont corrélés au Royaume-Uni :



Appliqué à Londres et à sa région, on aurait la représentation suivante :



Il convient, dans un premier temps, de définir la R.P. et le Cockney, l'EE étant présenté comme une variété intermédiaire. Le terme *Received Pronunciation* (R.P.) fut inventé par Alexander Ellis en 1869 et popularisé par Daniel Jones, l'auteur du premier dictionnaire de prononciation de l'anglais en 1913. Jones (cité dans le EPD [2003 : v]) explique que cet accent est :

most usually heard in everyday speech in the families of Southern English persons whose menfolk have been educated at the great public boarding schools.

Historiquement, cet accent s'est développé à la fin du Moyen-âge dans la région de Londres et du sud-est. C'était l'accent de la cour et de l'aristocratie. C'est au 19^{ème} siècle, après avoir évolué, qu'il est devenu l'accent des *public Schools* (dont les plus prestigieuses se situent dans le sud de l'Angleterre, comme Winchester, Harrow etc.) Il se répandit dans la haute fonction publique de l'Empire britannique et dans les forces armées. Il incarnait la culture, le pouvoir et l'autorité.

Le choix du terme pour renvoyer à l'accent standard et les caractéristiques phonétiques qu'il recouvre restent une question qui anime les phonéticiens contemporains. Dans son édition du EPD (2003: v), Peter Roach écrit:

For this edition a more broadly-based and accessible model accent for British English is represented. The time has come to abandon **the archaic name Received Pronunciation**. The model used for BE is what is referred to as **BBC English**. This the pronunciation of professional speakers employed by the BBC as newsreaders and announcers on BBC1 and BBC2 television, the World Service and BBC radio 3 and 4. Of course, one finds differences between such speakers – individual broadcasters all have their own personal characteristics, and an increasing number of broadcasters with Scottish, Welsh and Irish accents are employed. However, the accent described here is typical of broadcasters with an English accent, and there is a useful degree of consistency in the broadcast speech of these speakers. Their speech does not carry for most people the connotations of high social class and privilege that PSP [Public School Pronunciation] and RP have had in the past.

Quant à John Wells, dans le LPD (2008 : xix), il commente :

The model of British English pronunciation recorded in LPD is a modernized version of the type known as Received Pronunciation, or RP. In England and Wales, RP is widely regarded as a model for correct pronunciation, particularly for educated formal speech. It is what was traditionally used by BBC newsreaders – hence the alternative name BBC pronunciation, although now that the BBC admits regional accents among its announcers this name has become less appropriate. [...] RP itself inevitably changes as the years pass. There is also a measure of diversity within it. Furthermore, the democratization undergone by English society during the second half of the century means that

it is nowadays necessary to define RP in a rather broader way than was once customary. (xix).

Par delà la bataille éditoriale à laquelle se livrent les auteurs de ces deux dictionnaires, on voit que la définition de la R.P. est problématique. Les éditeurs du ODP (2001) conservent le terme RP et reprennent la distinction faite par Wells (1982) dans *Accents of English* entre *U-RP* et *Mainstream R.P.* qu'ils préfèrent appeler « R.P. non marqué » (*unmarked*). Selon Wells (1982), le RP se caractérise par sa variabilité. Si l'on reprend sa présentation, on a là aussi affaire à un *continuum* composé de quatre variétés identifiables :

- La ***U-R.P.*** est l'accent le plus prestigieux. La distinction *U / non-U* vient des années 50. On peut également parler de R.P. « conservateur », de *advanced RP* ou de « RP marqué ». Pour cette variété de R.P, il est parfois difficile de distinguer la description objective du stéréotype.

- La ***Mainstream R.P.*** est un accent R.P. plus neutre (non marqué) dont sont gommées les marques trop visibles de l'élitisme. On parle parfois de *general R.P.*

- La ***Adoptive R.P.*** est l'accent des individus ayant consciemment ou non adopté la prononciation R.P. sous la pression sociale. Ces locuteurs gardent leur capacité à parler avec leur accent d'origine. Phonétiquement, ils se distinguent du *Mainstream R.P.* dans le sens où ils ont tendance à utiliser moins de formes réduites, d'assimilations et de r de liaison, phénomènes qu'ils perçoivent comme des marques de parlure relâchée.

- Enfin, la ***Near R.P.*** est un accent perçu comme relativement standard mais dans lequel des marques d'accent régional demeurent. Cette variété est difficile à définir positivement. Par exemple, c'est l'accent d'un locuteur du nord de l'Angleterre ayant le même inventaire phonémique que la *Mainstream R.P.* à l'exception de la distinction /æ/ /ɑ:/ dans des mots comme *bath*.

D'autres termes furent ou sont toujours employés pour renvoyer à la R.P. : le *Queen's English*, *Oxford English*, *Public School accent* etc. qui mélangent accent et dialecte. Les termes *English English* ou

Southern British English se rencontrent aussi de plus en plus. La question est donc complexe et on verra que le choix terminologique se révèle instructif dans la présentation de l'EE.

Le Cockney, pour sa part, présente d'autres difficultés de description. La légende voudrait que soit (et parle) Cockney celui qui est né dans un rayon d'où l'on pouvait entendre les cloches de l'Eglise de St Mary-le-Bow dans le quartier de Cheapside (City of London), qui n'est pas dans l'East End. L'étymologie du terme n'est pas sûre et l'on avance soit le mot français « cocagne » (pour désigner Londres et ses environs), soit un enfant gâté (*to cocker*). Aujourd'hui, on considère que tous les habitants de l'East End sont des Cockneys mais que l'accent se rencontre de manière beaucoup plus large. On peut, ici aussi, invoquer la notion de *continuum* avec :

- Le dialecte Cockney (caractérisé par des différences grammaticales et lexicales – *Cockney Rhyming Slang*).
- L'accent Cockney (relativement standard au niveau grammatical mais ayant les caractéristiques phonétiques très nettes (élision du /h/ à l'initiale, glottalisation du /t/ et diphtongues différentes de la R.P.).
- Le Cockney du théâtre ou *stage Cockney* (accent stéréotypé pour évoquer les ouvriers et les artisans sur scène) ainsi que le *Mockney* (*Mock / Cockney*) qui désigne un accent Cockney outré et non authentique comme on le rencontre par exemple dans *The Book of Dave* de Will Self. On y trouve une référence particulièrement claire chez Ruth Rendell¹.

On voit que la notion de *continuum* est prépondérante et l'EE serait le chaînon manquant entre RP et Cockney pour la région du grand Londres :

U-RP – Mainstream RP – Adoptive RP – Near RP / EE – London accent / Cockney
--

¹ “You know how you sometimes hear an actor talking cockney on TV and it sounds good – they learn it from tapes and they've learnt well – but at the same time it's not genuine, it's not the real thing, **it's telly cockney** that we've got used to an accept. Well, that's what his voice was like, someone who'd learnt his cockney from a tape, and dropped his voice and took the inflections out of it. Altogether too much of a good thing, if you get my meaning”. (Ruth Rendell 1997: 122).

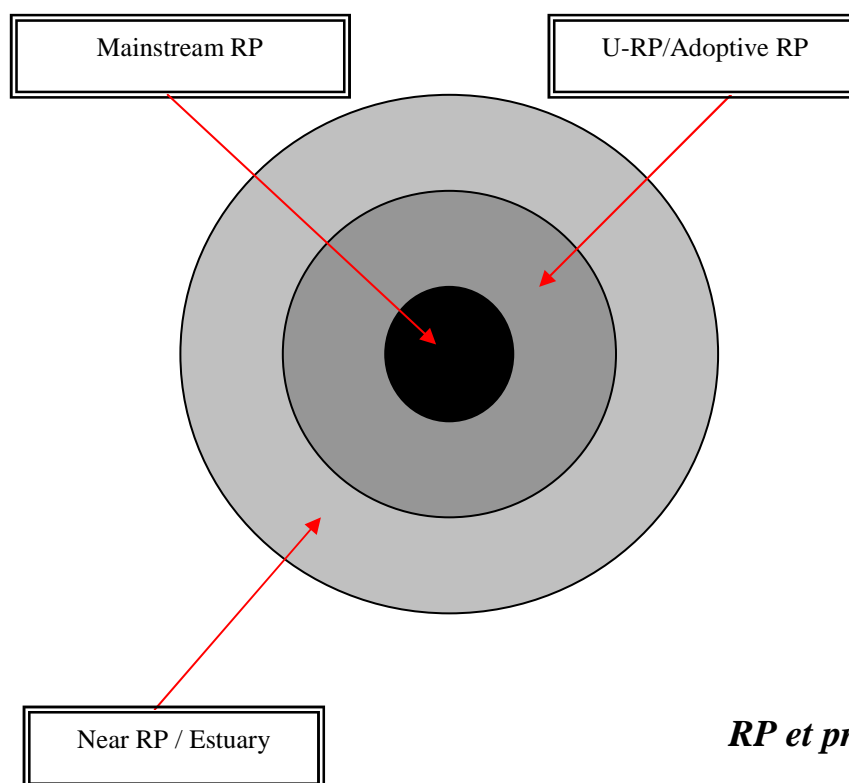
L'EE serait ce que Wells (1982 : 303) appelle *London Regional Standard* (LRS), c'est-à-dire un accent proche de la R.P. mais contenant des traits régionaux. Wells (1982 : 303) commente :

It must be remembered that labels such as “popular London”, “London Regional Standard”, do not refer to entities we can reify but to areas along a continuum stretching from broad Cockney [...] to R.P.

Chevillet (1991 : 105) propose la représentation suivante de la situation londonienne en la corrélant avec les différentes strates de la société :

Vrai Cockney (<i>broad</i>)	Accent londonien populaire	<i>London Regional Standard</i>	R.P.
Classe populaire		Classe moyenne	Classe supérieure

On peut néanmoins adopter un autre point de vue et tenter de raisonner en termes de prototypes. On aurait la R.P. non marquée (*Mainstream*) comme accent prototypique, c'est-à-dire comme point de référence cognitif. La structure radiale permet de positionner la *U-RP* et la *Adoptive RP* dans le second cercle. Toutes deux possèdent des attributs communs à la R.P. non marquée ainsi que d'autres.



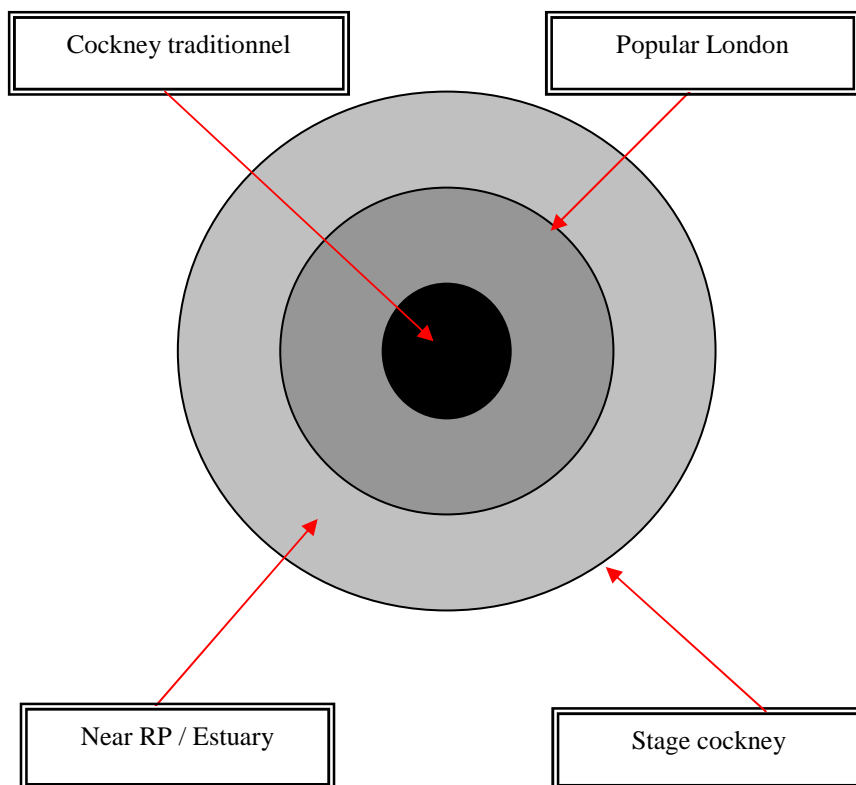
RP et prototype

Enfin, en zone périphérique, on peut situer la *Near RP* ayant elle aussi des attributs communs aux autres variétés mais des spécificités nettes. Pour reprendre l'exemple du locuteur du nord ayant adopté la R.P. mais gardant certains phonèmes nordiques, il est moins RP que les autres locuteurs car sa variété d'anglais, contrairement à toutes les autres, demeure localisable.

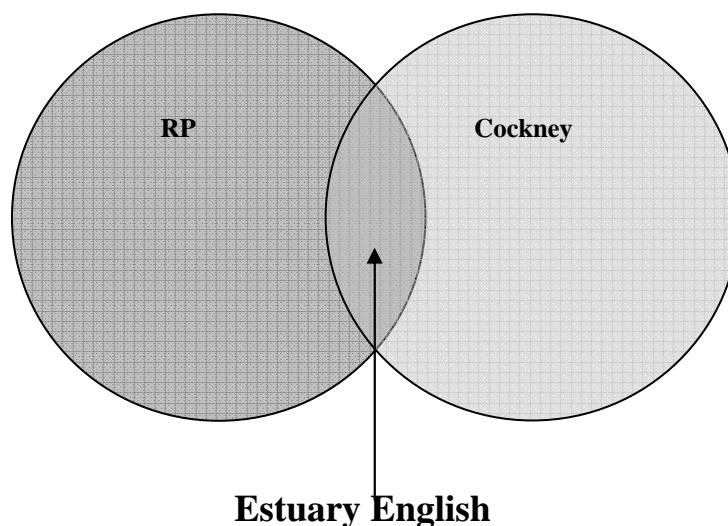
L'intérêt de cette présentation est de rompre avec l'idée implicite à la présentation pyramidale qui est de présenter la *U-RP* comme le modèle le plus prestigieux (acrolecte), donc celui à atteindre. La société anglaise reste très hiérarchisée, mais tend à estomper les marques trop évidentes d'appartenance à une élite voire, dans le cas de la *U-RP*, à l'aristocratie. On a donc la représentation prototypique suivante :

Si l'on applique le même principe au Cockney, on a :

Cockney et prototype



Si l'on en revient à l'EE, il se situerait au point de contact entre la présentation prototypique du RP et la représentation prototypique du Cockney :

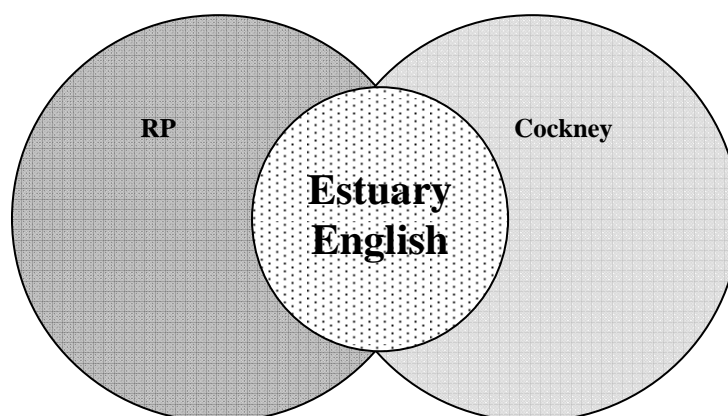


Cette présentation ne remet pas en cause des descriptions de l'EE qui existent mais a l'avantage de ne pas préjuger de son autonomie comme accent prototypique. Certes, le terme semble désormais faire partie de la conscience linguistique des anglophones comme le prouve cette référence très explicite chez Barbara Vine (1998 : 155) :

'Jason, please. I don't think anyone's called me Mr Thague before. Ever.' His wasn't a Suffolk voice, **rather the accent dubbed in the eighties Estuary English.**

Toutefois, cela ne constitue pas une garantie suffisante de son existence. On peut en effet s'interroger sur l'existence d'un prototype *Estuary* dans la conscience linguistique anglaise (voir schéma suivant). La surexposition médiatique tend à indiquer qu'un tel prototype existe et le terme *Estuary English* fait désormais partie du langage courant.

Or, si ce modèle prototypique existe, il est fondé sur une description précise des caractéristiques de l'EE.



Représentation d'un prototype de l'Estuary English, issu de la R.P. et du Cockney

2. Présentation de l'*Estuary English*

Il convient donc maintenant de décrire ce que recouvre l'appellation *Estuary English*.

2.1. EE : accent ou dialecte ?

La première interrogation est de savoir si l'EE est un accent ou un dialecte. Pour de nombreux linguistes, la question ne se pose pas et l'EE est présenté comme un accent. Or, le terme utilisé implique que l'EE est plus qu'un accent. En effet, on parle de EE et non pas de *Estuary Accent* ou de *Estuary English Accent*.

Dans la *Cambridge Encyclopedia of the English Language*, David Crystal (1995 : 327) donne plusieurs exemples de traits syntaxiques propres à l'EE :

- a. La systématisation du *question-tag* de confrontation : *I said I was going, didn't I.*
- b. L'utilisation fréquente de l'adverbe *never* pour renvoyer à une seule occurrence : *I never said I was going.*
- c. L'omission de la flexion adverbiale *-ly* dans : *You're turning it too slow / You're turning it too slowly.*
- d. La généralisation de la désinence de la troisième personne du présent simple, surtout dans le discours narratif : *I gets out of the car ; We was walking down the road.*

e. L'utilisation particulière de certaines prépositions : *I got off of the bench ; I looked out the window.*

Dans une perspective plus lexicale, Rosewane (1994) et Coggle (1993) notent également :

f. L'utilisation de *cheers* pour remplacer *Thank you* et *Good-bye*.

g. L'utilisation systématique du vocatif *mate* pour éviter *Sir* perçu comme marqueur de déférence.

h. L'adverbe *basically* perd son sens originel pour devenir une copule d'oralité.

i. Plus généralement, l'EE se caractériserait par une plus grande perméabilité aux américanismes du type *there you go (here you are)*, *Excuse me (I'm sorry)*, *sure (certainly)* etc.

Tous ces traits n'ont évidemment pas tous la même importance et nombre d'entre eux relèvent de la variation stylistique. L'anglais contemporain est soumis à l'influence des américanismes et il est difficile d'évaluer la perméabilité de l'EE. On considère généralement que les jeunes britanniques intègrent plus volontiers les américanismes que leurs aînés, dans le sud-ouest comme ailleurs. Cette plus grande porosité n'est donc pas propre à l'EE. Il en va de même pour les évolutions lexicales : *Cheers Mate* est tout aussi nordique que londonien. La perte de la flexion adverbiale (-ly) est elle aussi très répandue. Le seul trait plaçant véritablement pour l'EE comme un dialecte particulier est la désinence de la troisième personne à toutes les personnes. Or, David Crystal est le seul à associer ce trait, très répandu dans de nombreux dialectes anglais, à l'EE.

Si l'on met de côté ce dernier point, qui reste à prouver, les marques lexico-syntaxiques présentées relèvent donc plus d'un style conversationnel ou de l'évolution naturelle de l'anglais que de fonctionnements particuliers de la syntaxe ou de l'utilisation d'un lexique spécifique. Tous les mots cités plus haut appartiennent à l'anglais standard et c'est leur fréquence qui est pointée dans le cas de l'EE. On pourra donc (John Wells 1994 ; 1997) considérer que l'EE est conforme à la grammaire et à l'usage de l'anglais standard.

2.2. Description phonétique du EE

La description phonétique / phonologique de l'EE a plus retenu l'attention de la communauté linguistique et de nombreux travaux existent désormais, par exemple, Wells (1994, 1997), Coggle (1993), Altendorf (1999), Haenni (1999) ou Maidment (1998). On peut tenter la synthèse suivante en ne sélectionnant que les traits les plus souvent repris par ces différents auteurs :

- Vocalisation du L

Ce trait existe en Cockney depuis au moins un siècle. Selon Wells (1982 : 259), il a de fortes chances de se propager en anglais standard au cours du 21^{ème} siècle, ce que Hughes *et al* (2005) confirment. Le [o] est une variation allophonique de /l/. Ainsi, dans certains contextes, le /l/ est vocalisé : *milk* devient ['miok] au lieu du RP /'mɪlk/ et *Football* devient ['fʊʔbɔ:ɔ] au lieu de /'fʊtbɔ:l/. Ainsi, selon Wells [1994], on a :

	R.P.	E.E
field	'fi:ld	'fi:od
build	'bi:ld	'bi:od
wealth	'welθ	'weoθ
canal	kə'næl	kə'næo
Charles	'tʃɑ:lz	'tʃɑ:oz
doll	'dɒl	'dɒo
pull	'pʊl	'pʊo
bulk	'bʌlk	'bʌok
mile	'maɪl	'maio

- *Palatalisation*

L'EE se caractérise par une forte propension à la palatalisation. Ainsi /tj/ → /tʃ/ et /dj/ → /dʒ/ (surtout devant /u:/ sous accent) alors que la R.P. tolère la palatalisation mais surtout dans les syllabes non accentuées. On reviendra sur ce phénomène plus loin. Théoriquement, on a :

R.P. /'tju:sdeɪ/ EE /'tʃu:sdeɪ/

R.P. /'ætɪtju:d/ EE /'ætɪtʃu:d/

R.P. /'dju:k/ EE /'dʒu:k/

R.P. /rɪ'dju:s/ EE /rɪ'dʒu:s/

Le Cockney tend à éliminer le /j/, ce qui rend la palatalisation impossible (*tube, news*). L'effacement de /j/ se propage aussi en R.P. (*absolute, suit, salute, assume, consume, presume*). On assiste donc à une disparition du /j/ dans ces contextes, soit par élision (*yod-elision*) soit par palatalisation (*yod-coalescence*).

- *Modification des diphtongues*

Le Cockney et les accents du sud-est de l'Angleterre sont connus pour avoir une série de diphtongues distinctives. L'EE n'échappe pas à la règle et se trouve dans une situation intermédiaire entre la R.P. et le Cockney :

	RP	EE	Cockney
face	eɪ	ʌɪ	aɪ
price	aɪ	ɑɪ	ɒɪ
loaf	əʊ	ʌʊ / ɒʊ	aʊ
mouth	aʊ	ɑʊ	æ:

On laisse de côté la différence entre la R.P. et l'EE pour des mots comme *goal* et *roll* /ɒʊ/ devant un l sombre, en fait devant /o/ en EE, et /ʌʊ/ devant un l clair.

- Niveau prosodique

Aucune étude n'a été menée de manière systématique au niveau prosodique. Deux caractéristiques, qui restent à confirmer, sont parfois évoquées :

- L'EE aurait tendance à faire porter un accent nucléaire sur des prépositions et des auxiliaires (*Let's get TO the point, the house is ON fire*), mais ce phénomène bien connu est généralement expliqué par le fait qu'à la radio et à la télévision, les présentateurs utilisent ce type d'accent contrastif à des fins de clarté ou en raison d'une oralisation à partir d'un téléprompteur. Cette tendance ne serait donc pas une spécificité de l'EE.

- L'amplitude fréquentielle serait moins grande en EE qu'en R.P.

Curieusement, le *High Rise Terminal* (HRT) ou *uptalk* n'est pas évoqué comme caractéristique de l'EE alors que ce trait prosodique, venu d'Australie, se répand très rapidement parmi les jeunes générations. Mais c'est peut-être le manque d'étude systématique qui explique cette absence.

On peut donc résumer ainsi les traits qui distinguent l'EE du Cockney et de la R.P. :

Marqueur Phonétique	R.P.	Cockney	Estuary English
Vocalisation du L	+	+	+
Glottalisation + C	+	+	+
Glottalisation V-V	-	+	-
/i/ /ɪ/ /i:/	+	+	+
Palatalisation	+	+/-	+
Transformation des diphtongues	- ?	+	+
TH-fronting	-	+	-

Les phénomènes présents en EE semblent aussi toucher la R.P. telle qu'elle évolue. Les diphtongues de la R.P. ne sont pas celles de l'EE mais subissent également une évolution. La vocalisation du /l/ est amorcée en R.P. La glottalisation est présente et la palatalisation n'est

plus une exception. Seuls le TH-fronting (/f/ au lieu de /θ/ et /v/ au lieu de /ð/) et la glottalisation à l'intervocalique ne se trouvent qu'en Cockney. La question est donc de savoir si l'EE est amené à remplacer la R.P. ou est seulement un accent régional.

3. Quel avenir pour l'EE ?

Dans le LPD, Wells (2000 ; 2008) fait explicitement référence à l'EE en proposant les variantes allophoniques /o/ pour /l/ et /ʔ/ pour /t/ dans les contextes définis plus haut, ce qui donne une légitimité institutionnelle à l'EE. Les opposants à l'idée que l'EE deviendrait la nouvelle R.P. mettent en avant son caractère localisable alors que la R.P. est définie par l'absence de localisation possible.

Cet argument est difficilement recevable. On a vu que la R.P. trouvait son origine à Londres et dans ses environs et ce n'est que parce qu'elle est devenue l'accent du pouvoir, de la culture et de l'autorité qu'il a perdu son caractère localisable. Le facteur social a pris le pas sur la variable géographique. Par ailleurs, on voit que la R.P. évolue et que cette évolution suit, pour certains traits, ce que l'on observe en EE.

La palatalisation est un phénomène d'assimilation très courant en R.P. en syllabe non accentuée. On peut comparer les différentes transcriptions proposées par le LPD depuis 1990 afin d'évaluer le taux de pénétration de la palatalisation en RP en s'intéressant aux syllabes accentuées. On obtient les résultats suivants :

Mot	LPD 1 (1990)	LPD 2 (2000)	LPD 3 (2008)
Duke	dju:k, *dʒu:k	dju:k, *dʒu:k	dju:k → dʒu:k
Tuesday	'tju:zdeɪ, *'tʃu:z-	'tju:zdeɪ, *'tʃu:z-	'tju:zdeɪ, → 'tʃu:z-
Tune	tju:n, *tʃu:n	tju:n (64%) , *tʃu:n (35%)	tju:n (44%), → tʃu:n (54%)
Reduce	rɪ'dju:s * - dʒu:s	rɪ'dju:s * - dʒu:s	rɪ'dju:s → - dʒu:s
Induce	ɪn'dju:s * - dʒu:s	ɪn'dju:s * - dʒu:s	ɪn'dju:s → - dʒu:s
Deduce	dɪ'dju:s * - dʒu:s	dɪ'dju:s * - dʒu:s	dɪ'dju:s → * - dʒu:s
Dune	dju:n, *dʒu:n	dju:n, → *dʒu:n	dju:n, → dʒu:n
During	'dʒuərɪŋ, 'dʒuə-, 'dʒɔ:-, 'dʒɔ:-	' dʒuərɪŋ 67% , 'dʒuə- 37%, 'dʒɔ:-, 'dʒɔ:-	' dʒuə- (54%) 'dʒuə.(46%) 'dʒɜ:r 'dʒɜ:r, dʒɔ:r; dʒɔ:r

- * indique que cette prononciation n'est pas R.P.
- → indique une variante due à l'application d'une règle générale.

On note la disparition de la notation du fait que la prononciation n'appartient pas à la R.P. (*) entre l'édition de 2000 et celle de 2008, sauf pour *deduce*. De manière plus intéressante, *during* et *tune* ont comme première prononciation, la forme palatalisée. Pour *tune* la statistique passe à 68% pour /tʃ/ pour les locuteurs nés après 1981. Pour *During*, la statistique passe à 67% pour /dʒ/ pour les locuteurs nés après 1981. Cela laisse à penser que la tendance devrait se confirmer. Il convient bien sûr de garder à l'esprit que ces résultats ne sont que ceux donnés par le LPD et qu'il ne s'agit pas de véritables statistiques de terrain. Néanmoins, on sait qu'il y a généralement un laps de temps assez grand entre l'apparition d'un trait phonétique et son intégration dans les dictionnaires. On peut donc imaginer que la situation est encore plus nette que ce que les statistiques laissent à penser.

Est-ce à dire que l'EE va remplacer le RP ? La première question qui doit être posée est celle de l'existence même de l'EE (Maidment : 1998). En effet, s'agit-il d'un *Near-RP* ayant reçu une publicité disproportionnée ou d'une évolution naturelle de la R.P., amenée à le remplacer. Trudgill (2001) opte pour la première possibilité. En effet, quelle est la différence entre le *London Regional Speech* (LRS) de Wells (1982) et l'EE sinon son impact médiatique ? Trudgill souligne que le terme EE est mal adapté car il suggère que l'EE serait une nouvelle variété d'anglais, ce qui n'est pas le cas. Par ailleurs, il laisse à penser qu'il s'agit d'un accent limité à l'estuaire de la Tamise. Pour Trudgill, il s'agit d'un accent des classes moyennes (*lower-middle class*) typique des Home Counties (Essex, Kent, Surrey, Berkshire, Buckinghamshire et Hertfordshire). Trudgill s'inscrit en faux par rapport à ceux qui prétendent que le EE se répand dans tout le pays et propose trois explications :

- De nombreux locuteurs qui auraient autrefois adopté la R.P. (*adoptive RP*) conservent des marques régionales dans leur accent. Le sud-ouest étant la région la plus peuplée, nombre de ces locuteurs accèdent aux médias et on a donc une impression de forte représentation de l'EE.
- La mobilité sociale s'est accrue et des locuteurs de la classe moyenne ont désormais accès à des postes de responsabilité

Conclusion

Tout le monde s'accorde aujourd'hui sur le fait que le terme EE est mal choisi. Son influence sur le grand public est toutefois considérable et il serait vain de tenter de revenir à une appellation du type LRS, bien moins alléchante sur le plan du marketing. Si, à l'origine (articles de Rosewarne et livre de Coggle), la confusion a régné, personne ne prétend plus aujourd'hui que le EE soit un dialecte particulier et sa description phonologique est assez précise, même si des travaux restent à faire sur le plan prosodique.

Le Royaume-Uni a, depuis la fin des années 80, subi de nombreuses mutations socio-économiques dont la médiatisation de l'EE est un avatar linguistique. Les plus conservateurs y voient un appauvrissement, voire une faillite de la langue ; les plus libéraux y voient un nivellement linguistique nécessaire dans une société encore très hiérarchisée. Des débats semblables existent dans de nombreux pays.

La réalité linguistique, telle qu'on peut l'observer aujourd'hui et sans pouvoir pour autant prédire l'avenir, est nuancée. La R.P., quel que soit le nom que l'on souhaite lui donner, évolue et les linguistes anglicistes ont la chance d'avoir à leur disposition trois dictionnaires de prononciation concurrents qui, pour des raisons de précision descriptive autant que de marketing, auront à cœur de noter les évolutions au plus près. Certaines évolutions de la R.P. se retrouvent dans les accents/dialectes modernes dont parle Trudgill, ce qui explique que certains traits de l'EE se retrouvent en R.P. contemporain. Enfin, même si l'on sait aujourd'hui que les innovations linguistiques ne se propagent pas par vagues mais qu'elles se transmettent des grandes villes aux villes moyennes, puis des villes moyennes aux campagnes, certaines isoglosses, comme celle allant de l'estuaire de la Severn à celui de la Wash (/æ/ /ɑ:/), restent, à l'heure actuelle, des barrières linguistiques difficiles à franchir, même pour un accent aussi médiatisé que ce qu'il convient maintenant d'appeler *Estuary English*.

Références bibliographiques

Dictionnaires de Prononciation

- Jones, Daniel (1917). *Cambridge English Pronouncing Dictionary*. Cambridge, Cambridge University Press, 2003, 3rd edition, Peter Roach, James Hartman & Jane Setter (eds).
- Upton, Clive, Kretzschmar, William & Konopka, Rafal (2001). *Oxford Dictionary of Pronunciation for Current English*. Oxford, Oxford University Press.
- Wells, J-C. (1990). *Longman Pronunciation Dictionary*. London, Longman, 2008, 3rd edition.

Ouvrages et articles

- Altendorf, Ulrike (1999). « Estuary English. Is English Going Cockney? ». <http://www.phon.ucl.ac.uk/home/estuary/altendf.pdf> (consulté le 20/01/08).
- Chevillet, François (1991). *Les Variétés de l'anglais*. Paris, Nathan.
- Cogle, Paul (1993). *Do You Speak Estuary ?* London, Bloomsbury.
- Crystal, David (1997). *English as a Global Language*. Cambridge, Cambridge U.P.
- _____ (2003). *The Cambridge Encyclopedia of The English Language*. Cambridge, Cambridge U.P. (2nd edition).
- Haenni, Ruedi (1999). *Esturay English: Supposed Evidence and a Perceptual Approach* (unpublished dissertation). <http://www.phon.ucl.ac.uk/home/estuary/haenni1999.pdf> (consulté le 20/01/08).
- Hugues et al. (2005). *English Accents and Dialects*. London, Hodder Arnold (4th edition).
- Maidment, J.A. (1998). "Estuary English. Hybrid or Hype?". <http://www.phon.ucl.ac.uk/home/estuary/maidment.htm> (consulté le 20/01/08).
- Rendell, Ruth (1997). *Road Rage*. London, Penguin.
- Rosewarne, David (1984). "Estuary English". *Times Educational Supplement*, 19 (October 1984).
- Self, Will (2007). *The Book of Dave*. London, Penguin.

- Taylor, Andrew (2006). *A Plumb in your Mouth*. London: Haper Collins.
- Trudgill, Peter (2001). "The Sociolinguistics of Modern R.P."
- Vine, Barbara (1998). *The Chimney Sweeper's Boy*. London, Penguin.
<http://www.phon.ucl.ac.uk/home/estuary/trudgill.htm> (consulté le 20/01/08).
- Wells, John (1994). *Accents of English*. Oxford, O.U.P. 1982 (3 volumes).
- _____. "Transcribing Estuary English".
<http://www.phon.ucl.ac.uk/home/estuary/transcree.htm> (consulté le 20/01/08).
- _____ (1997). "What is Estuary English".
<http://www.phon.ucl.ac.uk/home/estuary/whatis.htm> (consulté le 20/01/08).

Brigitte Horiot
Université Jean Moulin Lyon 3 – CEL, EA 1663

Autour du francoprovençal : intérêt diachronique et repères dialectologiques

Contrairement à ce que pourrait laisser entendre le terme « francoprovençal », il ne s'agit pas d'un mélange de français et de provençal. « Le francoprovençal est un produit de la latinisation de la Gaule non méridionale qui, en refusant les innovations linguistiques de l'époque carolingienne, s'est détaché du domaine d'oïl. » (Tuailon 2007 : 20).

Au Moyen Âge le gallo-roman¹ était ainsi divisé en trois langues : la langue d'oïl et la langue d'oc bien connues, et le francoprovençal, ainsi dénommé seulement en 1873 par le linguiste italien Ascoli : « J'appelle franco-provençal un type linguistique qui réunit, en plus de quelques caractères qui lui sont propres, d'autres caractères dont une partie lui est commune avec le français et dont une autre lui est commune avec le provençal, et qui ne provient pas d'une tardive confluence d'éléments divers, mais au contraire atteste sa propre indépendance historique, peu différente de celle, par lesquelles se distinguent entre eux, les autres principaux types romans. » (Ascoli 1878 : 61 [traduction du texte italien]). Cet acte de naissance contenait, comme l'a noté Gaston Tuailon (1972 : 297), « plus d'une proposition téméraire ». Grâce aux linguistes qui consacreront une grande partie de leurs travaux, notamment leurs thèses, à l'étude de ce domaine², le francoprovençal est aujourd'hui solidement défini et décrit. Dans le cadre des « mercredis » du Centre d'Études Linguistiques, notre intervention, destinée à des étudiants non spécialistes du francoprovençal, s'appuiera constamment, à l'aide de

¹ Au sens de « langue romane née en Gaule » (Tuailon 2007 : 21).

² Pour les références, voir *infra*, sous 3.1.

nombreuses citations, sur les travaux de deux grands spécialistes : M^{gr} Pierre Gardette et Gaston Tuailon.

1. L'espace linguistique francoprovençal et son environnement

Le domaine du francoprovençal prend la forme d'une ellipse dont les deux foyers sont Lyon et Genève et il s'étend sur trois états : la France (65% du domaine), la Suisse (25%) et l'Italie (10%). Voir la carte 1 en fin d'article.

1.1. L'espace linguistique francoprovençal

Le domaine francoprovençal est à cheval sur trois pays : la France, la Suisse et l'Italie. En France ce domaine s'étend sur les départements de la Loire (moins la région de Saint Bonnet-le-Château), du Rhône, de la Saône-et-Loire (Mâconnais et sud-est de la Bresse loughannaise), du Doubs (arrondissement de Pontarlier), du Jura (sauf le pays de Dôle), de l'Ain, de la Haute-Savoie, de la Savoie, sur la plus grande partie du département de l'Isère et sur l'extrémité nord de l'Ardèche et de la Drôme. En Suisse, ce sont les cantons de Genève, de Vaud (chef-lieu Lausanne) et de Neuchâtel, ainsi que la partie romande des cantons de Fribourg et du Valais (chef-lieu Sion) qui se rattachent au domaine francoprovençal. L'Italie appartient à ce même domaine par la vallée d'Aoste (à l'exception des communes de Gressoney et d'Issime), par les petites vallées montagnardes situées au sud du Grand-Paradis (Orco, Stura, Viù) et par les communes de la vallée de la Cenischia entre le Mont-Cenis et Suse. Ainsi délimité, le domaine francoprovençal ne constitue ni une unité politique, ni une unité géographique, ni une unité historique.

A ce domaine s'ajoute une colonie francoprovençale en Italie méridionale, dans le nord de la région des Pouilles. Les deux communes de Celle di San Vito et de Faeto (le point 715 de l' AIS) « ont été repeuplées par des gens venus du pays francoprovençal. Leurs descendants vivent toujours dans ces deux communes, où ils parlent aujourd'hui encore le francoprovençal. Ce francoprovençal a été fortement influencé par le dialecte italo-roman des Pouilles et par la langue italienne. » (Tuailon 2007 : 110). A la suite d'une

conférence faite par Gaston Tuillon à Nantua (Bugey), un historien local a indiqué à ce dernier « que des pièces d'archives locales signalent qu'au XV^e siècle, des Bugistes sont partis pour l'Italie méridionale, sous la conduite d'un seigneur très influent de la région bugiste, le comte Alaman. Les archives du Bugey détiennent peut-être la solution du problème [...]. » (Tuillon 2007 : 112).

1.2. Les voisins du francoprovençal

« Quatre domaines linguistiques entourent le domaine francoprovençal : à l'est, l'alémanique de Suisse et de la vallée du Lys en Vallée d'Aoste ; au sud-est, le piémontais ; au sud et à l'ouest, depuis la vallée du Sangone en Piémont jusqu'à la région de Thiers en Auvergne, l'occitan ; au nord, la langue d'oïl. Entre alémanique et francoprovençal, l'opposition est brutale, elle se superpose à une grande partie de la limite qui sépare la Suisse romande de la Suisse alémanique. Entre le piémontais et le francoprovençal, la limite s'est compliquée par des siècles de bilinguisme : le piémontais est la langue de la plaine et le francoprovençal celle de la montagne. [...] ; ce bilinguisme a quelque peu altéré le francoprovençal et compliqué la délimitation, surtout pour la partie francoprovençale du Piémont. L'opposition avec l'occitan est assez nette, tandis qu'au nord, la délimitation avec la langue d'oïl est la plus compliquée. » (Tuillon 2007 : 26). En effet, avec l'appui de la langue française, les patois d'oïl n'ont cessé de progresser vers le sud. « Cette avancée peut connaître des phases d'accélération brutale, comme celle qu'ont subie le Mâconnais, le Roannais et le Beaujolais à partir de la fin du XIX^e siècle. » (Tuillon 2007 : 129).

2. L'originalité du francoprovençal

Pour expliquer l'originalité du francoprovençal, à la suite de Walther von Wartburg, on a d'abord fait intervenir l'influence burgonde. Le superstrat burgonde aurait joué dans la création du francoprovençal un rôle analogue à celui qu'aurait exercé le superstrat francique dans la formation du français. Les travaux d'Ernest Schüle, notamment sa communication au colloque de dialectologie francoprovençale de Neuchâtel en septembre 1969 (Schüle 1971 : 27-

47), ont conduit à abandonner cette explication. Les Burgondes ne sont pas venus très nombreux, « le bilinguisme né de leur faible présence n'a pas été assez fort pour créer des conditions favorables à d'importantes modifications linguistiques. Aussi le francoprovençal est-il resté une langue proche du latin, d'autant plus que, autour de Lugdunum, la capitale de la Gaule lyonnaise, la latinisation avait été établie sur des bases solides, que les Burgondes n'ont pas ébranlées. » (Tuailon 2007 : 160).

C'est actuellement l'hypothèse de M^{gr} Gardette, suivant laquelle le francoprovençal n'est peut-être que le pays des routes lyonnaises vers l'Italie, qui retient l'attention. La formule « le francoprovençal, langue des routes » est de P. Gardette et elle date de 1950, d'un compte rendu (in *Le Français Moderne* XVIII, p. 146-147) de l'ouvrage de Konrad Lobeck, *Die französisch-frankoprovenzalische Dialektgrenze zwischen Jura und Saône* (Genève - Zurich, « Romanica Helvetica 24 », 1945).

Dix-sept ans plus tard, au 5^{ème} congrès international de langue et littérature d'oc et d'études franco-provençales (Nice, 6-12 septembre 1967), P. Gardette terminait sa communication intitulée : « Le francoprovençal. Son histoire, ses origines » par cette phrase : « Le domaine franco-provençal n'est peut-être que le pays du Rhône moyen et de ses affluents, le pays des routes lyonnaises vers l'Italie. » (Gardette 1983 : 584). C'est dans un exposé (« Brève présentation du francoprovençal ») fait le 14 mai 1971 aux chercheurs du Centre de Philologie romane de l'Université de Strasbourg, à l'occasion de leur visite à l'Institut de Linguistique romane des Facultés catholiques de Lyon, que P. Gardette a développé sa formule, l'illustrant d'une carte (voir carte 2 en fin d'article). Voici un rapide résumé de l'exposé (Gardette 1983 : 607-612) et, plus généralement, de la pensée de Gardette sur cette question.

Le domaine francoprovençal se présente, pour l'essentiel, comme une unité sociale et économique, avec pour centre de rayonnement la ville de Lyon, toute tournée vers l'est. La forme triangulaire de ce domaine s'inscrit dans le dessin d'une double voie romaine à partir de la commune d'Aoste en France. La première voie remontait le cours du Rhône jusqu'à Genève, contournait le Léman par le nord, remontait le haut Rhône jusqu'à Martigny, passait les Alpes au Grand Saint-Bernard et aboutissait à l'autre Aoste en Italie. La deuxième voie rejoignait le cours de l'Isère, passait les Alpes au Petit Saint-

Bernard, et aboutissait elle aussi à Aoste. Cette double voie était, à l'époque romaine, la route la plus rapide vers l'Italie depuis Lyon, centre de l'activité économique gallo-romaine. Le francoprovençal est donc, à l'origine, la langue d'une communauté sociale qui s'est unifiée le long des routes qui empruntaient le cours moyen du Rhône et la haute vallée de l'Isère et qui, de tous temps, ont fait communiquer Lyon et l'Italie. A l'ouest de Lyon, traits et mots francoprovençaux s'arrêtent très vite, sans même aller jusqu'à la frontière naturelle des monts du Forez qui séparent Lyon de l'Auvergne et des pays du Centre orientés vers Paris. Aujourd'hui encore la région Rhône-Alpes reconstituée, grosso modo et sans le savoir, le domaine francoprovençal.

3. Les caractéristiques du francoprovençal

C'est encore dans sa communication intitulée « Le francoprovençal. Son histoire, ses origines » que P. Gardette a montré que l'individualisation du francoprovençal relevait de ses choix à chaque étape de son histoire linguistique (Gardette 1983 : 569-584).

La première étape, c'est la romanisation, et l'on sait que Lugdunum, fondée en 43 avant Jésus-Christ, fut une ville aimée des empereurs, le lieu de naissance de l'empereur Claude, la Rome des Gaules et, par là même, un centre important de la romanisation. Mais Lugdunum, tout en acceptant, pour une part, le latin de la Narbonnaise, affirme également l'indépendance de ses choix dans le sens du purisme, de l'archaïsme et de la poésie.

La deuxième étape se situe à la période romane primitive, période qui va être marquée par cinq grandes évolutions phonétiques dans la langue d'oïl et qui voit le francoprovençal en accepter quatre : la diphtongaison des voyelles moyennes accentuées en syllabe ouverte, la palatalisation de A précédé de palatale, la sonorisation des consonnes intervocaliques, la palatalisation de C, G devant A. Seule l'évolution de A tonique libre à [é]³ sera refusée, de même que l'affaiblissement des voyelles inaccentuées à la finale. Ces refus permettent à G. Tuailon de donner du francoprovençal la définition suivante : « Le francoprovençal est un produit de la latinisation de la

³ L'alphabet phonétique utilisé ici est celui auquel l'abbé Rousselot a donné son nom et qui a servi aux auteurs d'atlas linguistiques français, notamment pour les deux atlas auxquels nous renvoyons dans cet article : l'ALLy et l'ALJA.

Gaule du nord (Gaule lyonnaise et Gallia Belgica) ; mais ce produit de la latinisation de la Gaule non méridionale (Narbonnaise et Aquitaine) s'est détaché, à la fin de l'époque carolingienne des autres langues gallo-romanes du nord, en refusant une (au moins) innovation phonétique du IX^e siècle, l'affaiblissement des voyelles inaccentuées, qui a commencé à cette époque. » (Tuailon 2007 : 20).

La troisième étape, qui va du Moyen Âge à l'époque contemporaine, marque la désagrégation du francoprovençal sous l'influence de la langue de Paris. C'est dans ses patois, du moins ceux qui persistent, et dans le français régional qu'il faut maintenant aller chercher les traits phonétiques et morphologiques caractéristiques du francoprovençal et surtout son lexique. Seule la région autonome de la vallée d'Aoste, de par son bilinguisme officiel français/italien, conserve encore au francoprovençal la valeur de langue.

3.1. Les phénomènes linguistiques les plus originaux

Les phénomènes linguistiques les plus originaux du francoprovençal ont souvent été relevés et étudiés, citons notamment les noms de Louis Gauchat, Jules Jeanjaquet, Bengt Hasselrot, Hans-Erich Keller, André Devaux, Antonin Duraffour, Pierre Gardette, Gaston Tuailon⁴. Jean-Baptiste Martin, pour le tome V, 1, 1990 du *Lexikon der Romanistischen Linguistik* (« Francoprovençal », p. 671-685), a largement utilisé les études de Gardette et Tuailon mais il a ignoré nombre de publications, suisses notamment, consacrées au sujet, et il n'a pas toujours donné de précisions aréales, ce qui lui a valu un compte rendu très critique de P.-H. Liard dans le *Glossaire des patois de la Suisse romande. 93^e rapport annuel 1991* (p. 17-18). Pour éviter d'encourir de semblables reproches et compte tenu du fait que notre présentation se situe dans le cadre d'un bref exposé, destiné à des auditeurs qui n'ont pas forcément déjà entendu parler du francoprovençal, nous tenons à préciser que nous limitons notre description à l'énumération des phénomènes linguistiques les plus originaux et à leur illustration à l'aide d'exemples choisis dans deux textes utilitaires lyonnais, l'un du XIV^e siècle (le *Testament de Jean de la Mure*) et l'autre du XV^e siècle (un *Journal de recette du péage de Belleville*), textes que nous avons eu l'occasion d'étudier (Horiot

⁴ Nous indiquons dans la bibliographie quelques textes, notamment les thèses, de ces chercheurs.

2002 et Horiot, du Pouget 1990). Des renvois à des cartes choisies dans les deux atlas linguistiques francoprovençaux du territoire français : celui du Lyonnais (*ALLy*) et celui du Jura et des Alpes du Nord (*ALJA*), complètent cette présentation, sans pouvoir, le temps imparti ne le permettant pas, donner des précisions aréales quant aux cartes citées.

3.1.1 Phonétique

A atone final et A tonique libre.

C'est à Bengt Hasselrot que nous devons le critère qui permet de distinguer un parler francoprovençal d'un parler d'oïl et d'un parler d'oc. Dans un article déjà ancien, Hasselrot définissait le francoprovençal « comme l'ensemble des parlers où a final précédé de palatale devient *i* (*ə*, *e*) mais se conserve dans tous les autres cas [...]». Le premier critère sert à distinguer le franco-provençal du provençal, le second du français. » (Hasselrot 1938 : 80). L'évolution de A tonique libre suit les mêmes règles que A final : non précédé de consonne palatale, il se conserve tandis qu'il devient [yé] lorsqu'il est précédé de palatale.

Le *Testament de Jean de la Mure*, daté d'octobre 1361 (Durdilly 1975), nous servira pour illustrer ce traitement de A atone final > [a] :

[...], ju recomando m'arma a Nostrun Segnour Jhesus Cript et a sa chera mare, Nostra Dama, [...] (§ 1, p. 589) ;

ou [i] dans le cas de A final précédé de palatale :

Item, a chacon languan ['gisant'] en siveri en la cita de Lion, [...] (§ 6, p. 590) ; [...] el ['il'] lour leset la vigni dous Hovrors et la vigni dou Colombier [...] (§ 18, p. 592).

Ainsi que le traitement de A tonique non précédé de palatale :

[...] ma volunta et que mos ers ['mon héritier'] paiet, chaque an, [...] (§ 13, p. 591) ;

ou précédé de palatale :

[...] el non fit balier que .II^c ['il n'en fit donner que 200 florins'] (§ 28, p. 594).

L'ALLY et l'ALJA nous montrent que dans les patois recueillis il y a un demi siècle ou plus, non précédé de consonne palatale le A tonique libre reste [a] ou se vélarise en [o] tandis que précédé de consonne palatale, il se palatalise et le résultat actuel de cette palatalisation est, selon les régions ou les cas, [yé], [i], [ya], [é] (ALLY c. 1245 chanter, 1246 manger ; ALJA c. 1190 manger).

À remarquer que dans les cas de A atone final du pluriel (-AS) et de la désinence verbale de la deuxième personne du singulier de l'indicatif présent des verbes du premier groupe, il y a affaiblissement en [ê], comme l'attestent le *Testament de Jean de la Mure* : siveres (§ 11, p. 591), vignes (§ 26, p. 593), choses (§ 36, p. 596), ... et les cartes de l'ALLY (c. 1216 tu chantes) et de l'ALJA (c. 1828 venue ; venues) où l'on relève même, dans certaines localités, la disparition de [ê].

A suivi de consonne nasale.

Il conserve généralement son timbre, exemples dans le *Testament de Jean de la Mure* : man 'main' (§ 2, p. 589), chapellan 'chapelain' (§ 6, p. 590), pan 'pain' (§ 14, p. 591) ainsi que dans l'ALLY (c. 419) et l'ALJA (c. 859), la carte pain.

U atone final.

U atone final se maintient dans les proparoxytons, sous les graphies *ou* ou *o*. Il en est de même pour O désinence verbale de la première personne du singulier. Ainsi, dans le *Testament de Jean de la Mure* on relève malado 'malade' (§ 6, p. 590), chenevo 'chanvre' (§ 23, p. 593), leso '(je) laisse' (§ 4, p. 589) et voir ALLy c. 1143 et ALJA c. 401 chanvre ; ALLy 1215 je chante.

L'accent de mot.

Un autre critère permet de distinguer parler d'oïl - parler francoprovençal : c'est la place de l'accent. Alors que dans la langue d'oïl, par suite de la chute des voyelles atones finales et de l'évolution de -A final à e muet, tous les mots sont accentués sur la dernière syllabe prononcée, en francoprovençal, l'accent peut porter sur la dernière syllabe (phénomène appelé oxytonisme) ou, en raison du maintien de voyelles atones finales, sur l'avant-dernière syllabe (paroxytonisme). Avec **chenevo** 'chanvre' (§ 23, p. 593) et **Jaqueme** 'Jacques' (§ 36, p. 596), le *Testament de Jean de la Mure* nous fournit

deux exemples. Le *Journal de recette du péage de Belleville*, tenu en 1424, consigne également la forme francoprovençale du représentant de CANNAPU : [...] a passé [...] deux balon ... de **chenevez** (§ 2, p. 97) ; Item plus passé [...], deux quintaulx **cheneve**, [...] (§ 148, p. 108).

À la question "du chanvre" (c. 1143), les informateurs de l'*ALLY* ont massivement donné les formes francoprovençales ou occitanes, avec déplacement de l'accent sur la pénultième ; il en est de même pour ceux de l'*ALJA* (c. 401).

Palatalisation de U.

U s'est palatalisé en [u] à date récente car nous savons que le francoprovençal demeura fidèle à la prononciation de l'Ū latin lorsque, dans toute la France, se fit jour la tendance à palataliser cette voyelle en [u], probablement dès le début de la période romane. Aujourd'hui encore, il demeure fidèle au timbre vélaire *ou* « dans de petites aires témoins situées dans le Bugey, le Haut-Valais, la région de Chamonix, une petite région auvergnate qui confine avec le Forez. De plus, dans des zones beaucoup plus vastes, les parlers francoprovençaux conservent l'U en hiatus, par exemple dans la finale des participes passés féminins en -UTA devenue -*wa*, en -UTAS devenue -*wé*, et aussi dans les cas où l'U est en contact avec une consonne nasale. Le refus de la palatalisation de l'U latin est une caractéristique du francoprovençal qui le sépare des parlers d'oïl, exception faite du wallon, et en général des parlers d'oc qui n'ont pas au même degré conservé la prononciation latine. » (Gardette 1983 : 573). L'*ALLY* (c. 1247 vendu ; -ue ; -ues) et l'*ALJA* (c. 1828 venue ; venues) attestent ça et là de la conservation des finales [-wa] < -UTA et [-wé] < -UTAS.

U long + N > [õ] 'un' : *ALLY* c. 1175 et *ALJA* c. 1560.

Diphtongaison des voyelles moyennes.

Si le francoprovençal a accepté la diphtongaison spontanée des voyelles moyennes libres, il est à remarquer qu'il a parfois aussi diphtongué les voyelles entravées. On peut, surtout dans les documents anciens, relever *puerc*, *buec*, *uert* ... pour *porc*, *bois*, *ort* ... (Hafner 1955 : 52). Dans le *Testament de Jean de la Mure*, on relève une forme *cuer* < CORDU :

[...] .II. linceuz de lin et .I. coysotz et trimelleres de **cuer** buly [‘deux draps de lin et une paire de cuissardes et de jambières de cuir bouilli’] (§ 47, p. 598).

Cette diphtongaison, venue de la langue d’oïl, s’est répandue peu à peu en francoprovençal, par irradiation du Nord vers le Sud. Ainsi la *Somme du Code*, texte écrit à Uriage, tout près de Grenoble, dans la 1^{ère} moitié du XIII^e siècle, ne connaît pas cette diphtongaison. Les graphies du *Journal de recette du péage de Belleville* témoignent encore parfois de l’existence d’anciennes formes, principalement pour le prénom Pierre : le filz **Perre** (§ 80, p. 103 ; § 86, p. 104), **Perre** Nausier (§ 65, p. 102), ..., la saint **Perre** (§ 138, p. 108), mais aussi [...] et ung navey **nové** [‘un bateau neuf’] (§ 251, p. 116).

Les résultats actuels de ces diphtongaisons varient considérablement d’une région à l’autre. Voir les cartes porc (*ALLY* 321 ; *ALJA* 731), bois (*ALLY* 439 ; *ALJA* 552) et jardin (*ALLY* 253 ; *ALJA* 407).

La nasalisation.

Comme en français, les voyelles sont ordinairement nasalisées lorsqu’elles sont suivies d’une consonne nasale intérieure devant consonne, ou devenue finale. Toutefois le francoprovençal a conservé les voyelles nasales placées devant consonne nasale intervocalique (ex. : LANA > [lãna], *ALLY* c. 315 et *ALJA* c. 719. Voir également Hafner 1955 : 71 et 88) et il se distingue encore du français en gardant, le plus souvent, à la voyelle nasalisée son timbre originel (voir *supra*, A suivi de consonne nasale) :

PANE > [pã] ‘pain’ (*ALLY* c. 419, *ALJA* c. 859) ;
 VENTU > [vẽ] ‘vent’ (*ALLY* c. 766, *ALJA* c. 39) ;
 VENDERE > [vẽdr] ‘vendre’ (*ALLY* c. 1248*).

3.1.2 Morphologie

L’article défini.

En ancien francoprovençal :

	masculin singulier	masculin pluriel	féminin singulier	féminin pluriel
Cas sujet	<i>li</i>	<i>li</i>	<i>li</i>	<i>les</i>
Cas régime	<i>lo</i>	<i>los</i>	<i>la</i>	<i>les</i>

Voici quelques exemples relevés dans le *Testament de Jean de la Mure* :

Ju, Johan de la Mura, **li** plus ayna [‘l’aîné’] (§ 1, p. 589) ;
 [...] et a tos **los** sainz et sayntes de **la** cort de paradis [...] (§ 1, p. 589) ;
 Item, vuel et ordeno que seant meis el pont de Sauna en cel an, se **les** egues
 sont en poin [‘si les eaux le permettent’] et quan on pora, [...] (§ 9, p. 590) ;
 [...] que **li** .c. flur. se metant en pavir **lo** pont de Sauna [...] [‘que les cent
 florins soient employés à paver le pont de Saône] (§ 10, p. 590) ;
 [...] **li** entencion dou testour est que, [...] (§18, p. 592).

En francoprovençal moderne, c’est *lo* au masculin singulier, *lou* au masculin pluriel ; *la* au féminin singulier, *lé* ou *le* au féminin pluriel.

Voir cartes 1171, 1172, 1173, 1174 pour l’*ALLY* et cartes III, IV (volume I) pour l’*ALJA*.

Le pronom personnel neutre.

Les formes (sujet et complément) sont ordinairement différentes de celles du pronom personnel masculin singulier :

	neutre	masculin singulier
sujet	<i>o, é</i>	<i>i</i>
complément	<i>o, ou, i</i>	<i>lo</i>

Voir *ALLY* c. 776 et 777, *ALJA* c. 1592 pour *il* ; *ALLY* c. 1238 et 1239, *ALJA* c. 1607 pour *le*.

Le pronom démonstratif neutre.

Il présente une forme accentuée [sɛ̃] qui comporte une voyelle nasalisée, remontant probablement au latin INDE. La forme francoprovençale s’oppose aux formes des autres familles gallo-romanes (français *ce*, occitan *aquo*) qui sont des composés du latin HOC. Voir *ALLY* c. 1204, *ALJA* c. 1587, et Martin 1976.

En ancien francoprovençal, neutre sujet et régime : *co* (*so, czo*), *cen*. Ex. : *Testament* § 7, p. 590 [...] et paya **so** que ju leso ; § 19, p. 592 [...] non obstant **czo** que el ne ly deit pas [...].

Pour le francoprovençal moderne, et pour les démonstratifs en général, la répartition des formes en ECCE- et en *ACCU-, voir Gardette 1941 (*Géographie morphologique*) : 33-38.

L'adjectif possessif.

Dans les anciens textes comme dans les patois, il présente, pour certaines formes, une particularité intéressante. Voici le tableau en ancien francoprovençal, pour les unipossessifs :

	masculin singulier	masculin pluriel	féminin singulier	féminin pluriel
Cas sujet	<i>mos, tos, sos</i>	<i>mi, ti, si</i>	<i>mi, ti, si</i>	<i>mes, tes, ses</i>
Cas régime	<i>mon, to, son</i>	<i>mos, tos, sos</i>	<i>ma, ta, sa</i>	<i>mes, tes, ses</i>

Ainsi dans le *Testament de Jean de la Mure* :

et vuel que **mos** cors ['mon corps'] seit seveli avoy **mon** pare ['soit enterré avec mon père'] (§ 1, p. 589) ; Item, vuel et ordino que **mi** clan et **mes** lesses seant payez ['que mes dettes et mes legs', traduction approximative des termes latins clamor et eleemosyna] (§ 2, p. 589) ; [...] ou a **soz** eirs ['ou à ses héritiers'] (§ 39, p. 596).

Pour les pluripossessifs, et c'est là que réside la particularité du francoprovençal, sur *mon, ton, son*, on a les formes analogiques *notron, votron* bien étudiées par Bengt Hasselrot (1938 : 80 et 1966). L'analogie a été facilitée par le fait qu'en francoprovençal *notre* et *votre* ont une finale *-o*. P. Gardette note que tout le Forez francoprovençal dit [notrõ] tandis que la partie occitane du même Forez dit [nòtr] (Gardette 1941 *Géographie morphologique* : 42).

Le *Testament de Jean de la Mure* nous fournit un exemple : Ju recomando m'arma a **Nostrun** Segnour Jhesus Cript [...] (§ 1, p. 589).

Voir *ALLY* c. 1183 *notre* (jardin) ; *votre* (jardin) et *ALJA* c. 1569 *notre* (père), c. 1569* *votre* (père).

Les formes verbales :

À la 1^{ère} personne de l'**indicatif présent**, le francoprovençal a conservé la désinence [o] de la 1^{ère} conjugaison latine et il l'a étendue à l'ensemble des conjugaisons (ex. : *ALLY* c. 1215 *je chante*, *ALLY* c. 1256 et *ALJA* c. 1689 *je finis*) et à d'autres temps (ex. : *ALLY* c. 1271 *je chantais*, *ALJA* c. 1757 *je faisais*). Pour l'ancien francoprovençal, le *Testament de Jean de la Mure* nous fournit de nombreuses formes : *recomando* (§ 1, p. 589), *ordino* (§ 2, p. 589), *dono* (§ 35, p. 595), ...

Au singulier de l'indicatif présent des verbes des groupes II et III le francoprovençal possède une structure 1^{ère} personne \neq 2^{ème} personne = 3^{ème} personne (ex. pour le verbe "venir" : *veno* / *vẽ* / *vẽ*), qui le différencie à la fois de la structure française (1^{ère} personne = 2^{ème} personne = 3^{ème} personne) et de la structure occitane (1^{ère} personne \neq 2^{ème} personne \neq 3^{ème} personne).

À la 5^{ème} personne de l'indicatif présent du groupe I, le francoprovençal emploie généralement deux types de désinences : *a* (ou *o*) : *ALLY* c. 1219 *chanta* '(vous) chantez' ; *ie* (ou *i*) : *manji* '(vous) mangez'. Cette caractéristique n'est que la conséquence du traitement de A tonique libre en francoprovençal (voir *supra*, A atone final et A tonique libre).

À l'**indicatif imparfait**, le francoprovençal, comme l'occitan, fait une distinction entre les verbes qui comportent un suffixe en *v* issu des formes latines en -ABAM (ex. : *ALLY* c. 1271 *chantavo* 'je chantais' ; *Testament* § 28, p. 594 *montave* 's'élevait à', § 22, p. 593 *montavon* 's'élevaient à' (en parlant d'une somme d'argent), § 13, p. 591 *restavon* 'restaient' ; *Journal de recette du péage de Belleville* § 22, 83, 84, 167 202, 223 *menave* 'menait', § 100 *menavez* 'menait', § 43 *menevet* 'menait'), et les verbes qui, normalement, ne comportent pas ce suffixe, car ils sont issus des formes vulgaires en -EAM, -IAM (ex. : *fazyà* 'tu faisais', *venya* 'tu venais'). Mais, dans la plupart des régions, le suffixe en *v*, plus expressif, s'est, par analogie, introduit dans de nombreux imparfaits des groupes II et III (ex. : *venivo* 'je venais', *fazivo* 'je faisais').

Au **passé simple**, les désinences les plus fréquentes étaient en *i* (ex. : *chanti* 'je chantai'), mais ce temps a disparu de la plus grande partie du domaine. Dans le *Testament*, *costiet* 'coûta' (§ 27, p. 594), *baliè* 'donna' (§ 28, p. 594) et dans le *Journal de recette du péage de Belleville*, *passit* relevé deux fois (§ 3 et 6) contre une attestation *passa* (§ 180), et *chargit* (§ 65 et 146).

À l'**infinitif**, les verbes du 1^{er} groupe comportent deux désinences différentes, -ARE ayant évolué différemment selon qu'il était, ou non, précédé de consonne palatale (voir *supra*, A atone final et A tonique libre).

Même remarque pour le **participe passé** masculin du groupe I, mais à noter que dans la majeure partie du francoprovençal, palatale + -ATU a donné [ya], alors que palatale + -ARE a donné [yé], [i], [é].

3.1.3 Morpho-syntaxe

Les adjectifs déverbaux.

Gaston Tuillon remarque que chaque parler francoprovençal a une dizaine ou une vingtaine d'adjectifs déverbaux, qu'on appelle ordinairement "participes tronqués". Ces adjectifs déverbaux accentués sur le radical ont émergé dans le français local, où ils demeurent très vivaces (ex. : Je suis tout *trempe* 'je suis tout mouillé' ; ma montre est *arrête* 'ma montre est arrêtée'). On distingue ainsi le participe, accentué sur la désinence, de l'adjectif verbal (Tuillon 1983 : 63, article *arrête*). L'adjectif déverbal est employé pour marquer l'état, le participe passé est utilisé pour indiquer l'action. Voir également Fréchet 1995 : 37, 1.5 et DSR : article *arrête*, remarque.

4. Un texte patois du XXI^e siècle

Voici un petit texte, écrit en 2008 et publié dans *Lo Conte*, bulletin de l'Association Vaudoise des Amis du Patois. À ce bulletin, rédigé en patois, est joint, sous forme de feuilles volantes, une traduction en français de tous les textes. Notre document est en page 3 du bulletin et sa traduction en page 1 des feuilles volantes. Cet écrit contient nombre de traits phonétiques et morpho-syntaxiques décrits *supra* et, par conséquent, il ne nécessite pas de commentaire linguistique, si ce n'est pour le terme *tsalande* 'Noël' (*ALLY* c. 883, *ALJA* c. 1520), bien étudié par Jakob Jud (« Sur l'histoire de la terminologie ecclésiastique de la France et de l'Italie ») qui, carte à l'appui, a localisé le terme dans la Suisse romande et tout le Sud-Est de la France, c'est-à-dire dans les parlers provençaux et francoprovençaux (Jud 1973 : 218-219 carte, 232). Le mot *chalendes* est relevé six fois dans le *Journal de recette du péage de belleville* (§ 2, 69, 101, 104, 159, 234), toujours dans le même contexte, pour servir de date. Ex. : [...] le mardi avant **chalendes** [...] (§ 101).

La forme *treimblyotant*, avec la désinence *-ant*, n'est pas caractéristique du francoprovençal, on la trouve également (*-ant/-ont*)

dans le domaine d'oïl, à l'exception de la Normandie et de l'Île-de-France.

Sapalla

Cein que y'âmo ein tot premî dein lo tsalande de l'Amicâla, l'è sa sapalla, sta pucheinta sapalla que son coutset va quasu gatoillî lo pliafond dâo grand pâilo de Pouâidâo et que sè tsandèlè sant dâi vèritâbliè, avoué lâo petioûtè flianmè que treimblyotant, et pas dâi z'èlètriquè. Lâi a on mouî de mot ein patois po translata "sapin". Et po sti yâdzo noutron patois l'è quasu fèministe du que lo premi mot que no z'ein, l'è la sapalla ; clli mot va atant po lo vouârgno que po la fiva et la pèssa.

Onna petita sapalla, l'è on sapalet âo on sapalon.

Voici la traduction de ce texte :

Sapin

Ce que j'aime en tout premier dans le Noël de l'Amicale, c'est son sapin, ce puissant sapin dont la cime va presque chatouiller le plafond de la grande salle de Puidoux et parce que ses bougies sont de vraies bougies, avec des petites flammes qui tremblotent, et pas des électriques.

Il y a une quantité de mots en patois pour traduire « sapin ». Et pour cette fois notre patois est presque féministe puisque le premier mot que nous en avons, c'est la sapelle ; ce mot va aussi bien pour le sapin blanc que pour l'épicéa ou sapin rouge.

Un petit sapin c'est un sapinet ou un sapalon.

5. Le lexique francoprovençal

Pierre Gardette a montré que parmi les mots spécifiques du francoprovençal, une partie remontait au latin classique et poétique diffusé par Lugdunum. En dépouillant les volumes du *FEW*, Christian Schmitt a établi que le francoprovençal possède « 136 unités lexicales d'origine latine (que j'appellerai le vocabulaire spécifique) qui ne se trouvent nulle part ailleurs dans la Galloromania [...]. » (Schmitt 1977 : 94). Mais force est de constater que parmi tous les mots propres au francoprovençal, bien peu touchent l'ensemble du domaine, et même en couvrent une partie importante.

Parmi les mots dont l'aire occupe tout le domaine francoprovençal : *molar* 'talus', 'tertre' (*ALLY* c. 834, *ALJA* c. 1545), mot longuement étudié par Gardette puisqu'il lui a consacré un article : « Francoprovençal *molar* 'tertre, talus, tas de pierres, tas de terre, montagne' » (Gardette 1983 : 373-394) et qu'il en a parlé dans d'autres articles, notamment celui consacré à l'histoire et aux origines

du francoprovençal (Gardette 1983 : 569-584) où il remarque : « Parfois Lugdunum fit cavalier seul et choisit des mots ou des formes indépendamment de toutes les autres villes de Gaule. Nous en avons un précieux exemple dans le frpr. *molar* "tertre, colline", parfois "montagne" qui, conservé aujourd'hui surtout comme nom de lieu, est répandu dans tout le domaine francoprovençal, à l'exclusion de toutes les autres régions de la Gaule. Il semble bien que l'origine de ce mot soit l'adjectif latin MOLARIS, employé comme substantif au sens de "grosse pierre", "rocher" par les poètes Virgile et Ovide et par un Lyonnais du V^e siècle, le précieux Sidoine Apollinaire. » (Gardette 1983 : 571).

Un autre nom, celui de la noisette, a également retenu longtemps l'attention de P. Gardette. Dans un article publié en 1964 dans la *Revue de Linguistique Romane* : « En marge des atlas linguistiques du Lyonnais, du Massif Central, du francoprovençal du Centre. Les influences des parlers provençaux sur les parlers francoprovençaux » (Gardette 1983 : 711-723), Gardette a étudié quelques mots dont celui de la noisette. Un an avant, en 1963, sa contribution aux *Festschrift Alwin Kuhn* (Gardette 1983 : 769-773) était également consacrée à la noisette : « Les dénominations gallo-romaines de la noisette ». L'étude mérite d'être rappelée ici.

Le latin « n'avait qu'un mot pour la noix et la noisette : NUX, qui pouvait désigner aussi généralement tous les fruits à amande. La noisette cultivée à Abella en Campanie était renommée ; on l'appelait naturellement NUX ABELLANA. Peu à peu, suivant une habitude bien connue, on se mit à dire simplement ABELLANA, ou avec un suffixe allongé *ABELLANIA ; on en arriva même à désigner ainsi non seulement la noisette d'Abella, mais toute espèce de noisette. De la sorte on supprimait l'amphibologie de NUX "noix" et "noisette" : on pouvait réserver NUX pour la noix et employer ABELLANA, -IA pour la noisette. Le latin tardif remédia d'une autre façon encore à l'amphibologie de NUX, il fabriqua pour la noisette des diminutifs : NUCELLA, et NUCICULA qui n'apparaît pas dans les textes mais seulement dans les gloses et dans Isidore de Séville. » (Gardette 1983 : 769). A l'époque où la Gaule apprenait à parler latin, il y avait donc « deux manières de nommer la noisette, si l'on voulait éviter l'amphibologie des deux sens de NUX : on pouvait user d'un diminutif comme NUCICULA ou *NUCEOLA ; on pouvait aussi dire ABELLANA ou *ABELLANIA. » (Gardette 1983 : 769). Grâce aux

textes médiévaux, grâce surtout aux patois, Gardette esquisse la carte des noms gallo-romains de la noisette. Au Moyen-Âge, dans la France du Nord, pays de langue d'oïl, nous trouvons trois façons de nommer la noisette : on dit *noisille*, *nois de coudre* et *noiz menue*, on dit enfin *noisette*. « Dans la France d'oc, l'usage est différent. On dit *avelanas*, *avilana*, *aulanas*, *aulaigna* [...], mots qui remontent évidemment à ABELLANA, -IA. » (Gardette 1983 : 769). Quant à l'ancien francoprovençal, c'est grâce à des noms de lieux en -ARIU, -ARIA que nous pouvons connaître le choix fait par ce domaine. Dans le département du Rhône (fichier dictionnaire topographique), nous avons des *Allognet*, des *Allognières*, des *Allogniers*. La plus ancienne mention (*Aloygnetum*) est de 1300. Dans le département de la Loire, ces noms sont également nombreux (Dufour 1946). Nous rencontrons également des noms de personnes comme *Ollagnon*, *Ollagnet*, ...

Ainsi, dès le XIII^e siècle au moins, le francoprovençal connaissait les noms actuels de la noisette : *olagne*, *alogne*, *alagne*. Ces noms permettent à Gardette « de ranger le domaine francoprovençal à côté du domaine provençal et de remarquer qu'au Moyen-Âge la langue d'oïl d'une part, la langue d'oc et le francoprovençal d'autre part, présentaient des dénominations fort différentes pour la noisette. Cette opposition des deux moitiés de la France est corroborée par les patois, qui nous offrent, grâce aux atlas [*ALLY* c. 479 et *ALJA* c. 479], des documents nombreux, sûrs, bien localisés. Ils nous montrent que les mots du Moyen-Âge sont les mêmes aujourd'hui. » (Gardette 1983 : 770).

Conclusion

J'emprunterai la conclusion à Gaston Tuailon (2001 : 12) : « Le francoprovençal est une langue romane qui est restée au stade naturel de langue géographiquement variable, exactement comme l'était au Moyen Âge la langue d'oïl, avant que la langue du Roi, c'est-à-dire la variété parisienne de la langue d'oïl, ne propose ou n'impose une forme unitaire et codifiée, qui devint, au fil des siècles, le français. »

Tout est dit dans ce « qui est restée au stade naturel de langue géographiquement variable » et l'on ne peut que s'étonner de rencontrer des partisans, voire des militants d'une langue qui, pour en être une, doit être artificiellement créée.

Orientation bibliographique

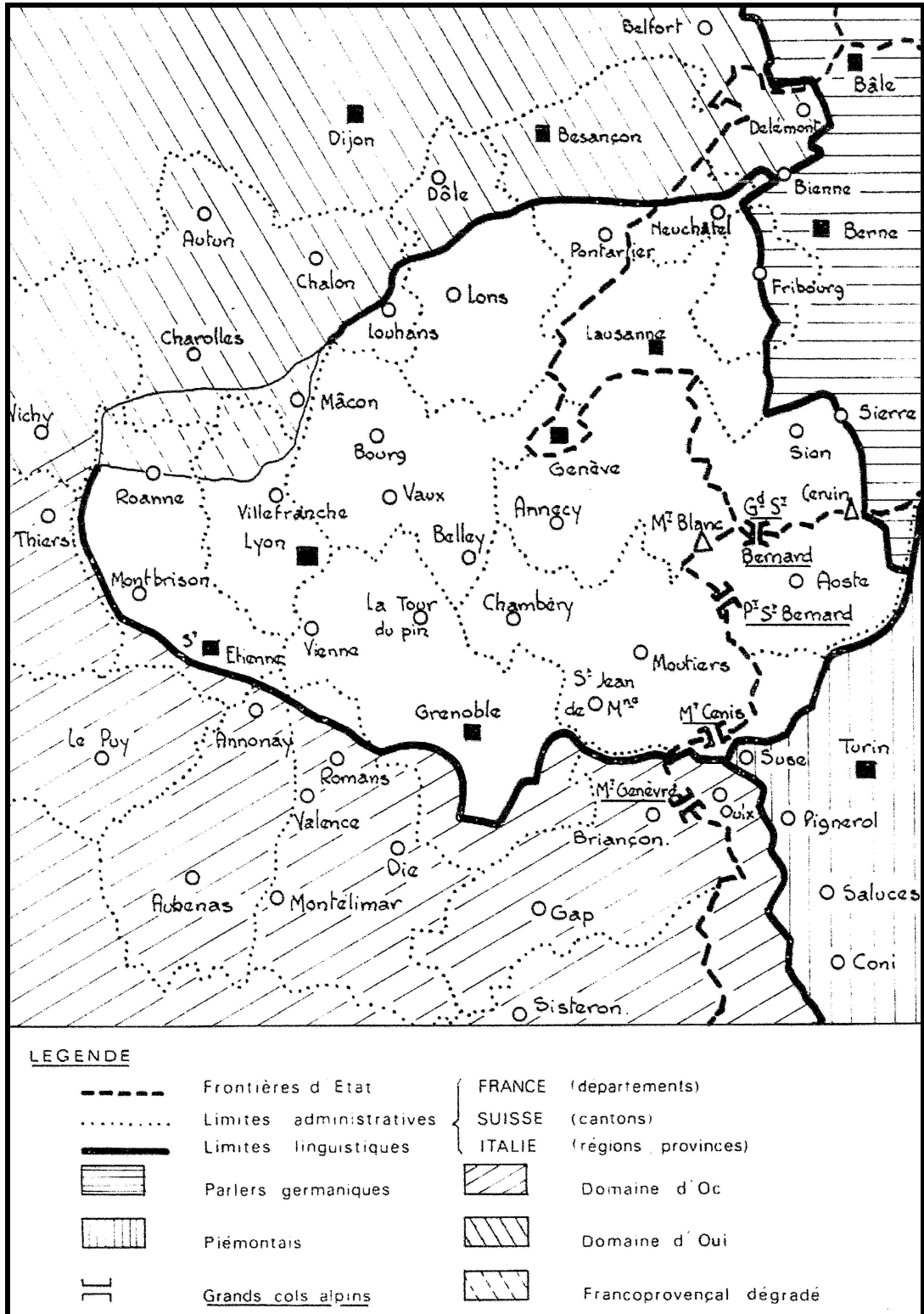
- AIS* : *Sprach- und Sachatlas Italiens und der Südschweiz* von Karl Jaberg und Jakob Jud, 8 volumes, Zofingen : Ringier & Co., 1928-1940.
- ALJA* : *Atlas linguistique et ethnographique du Jura et des Alpes du Nord* par Jean-Baptiste Martin et Gaston Tuaille, Paris : éditions du CNRS, 3 volumes, 1971-1978.
- ALLy* : *Atlas linguistique et ethnographique du Lyonnais* par Pierre Gardette, Paris : éditions du CNRS, 5 volumes, 1950-1976.
- Ascoli Graziadio-Isaia, 1878. « Schizzi franco-provenzali », in : *Archivio glottologico italiano* III, 1878, p. 61. La date de 1878 donnée à ce tome est celle du groupement sous la même reliure d'articles parus précédemment.
- Devaux André, 1892. *Essai sur la langue vulgaire du Dauphiné septentrional au moyen âge*, Paris-Lyon. [Genève : Slatkine reprints, 1968]. [thèse]
- DSR* : *Dictionnaire suisse romand. Particularités lexicales du français contemporain. Une contribution au Trésor des vocabulaires francophones*, conçu et rédigé par André Thibault, sous la direction de Pierre Knecht. Carouge - Genève : éditions Zoé, 1997.
- Dufour J.-E., 1946. *Dictionnaire topographique du Forez et des paroisses du Lyonnais et du Beaujolais formant le département de la Loire*, Macon : imprimerie Protat frères (« Fondation Guichard »).
- Duraffour Antonin, 1932. *Phénomènes généraux d'évolution phonétique dans les dialectes franco-provençaux d'après le parler de Vaux-en-Bugey (Ain)*, Grenoble ; *Description morphologique avec notes syntaxiques du parler franco-provençal de Vaux (Ain) en 1919-1931*, Grenoble. [thèses]
- Durdilly Paulette, 1975. *Documents linguistiques de la France, série francoprovençale, II Lyonnais*, Paris : éditions du CNRS.
- FEW* : *Französisches Etymologisches Wörterbuch. Eine darstellung des galloromanischen sprachschatzes* par Walther von Wartburg et al. Bonn ; Leipzig ; Bâle, 1922-2002, 25 volumes.
- Fréchet Claudine, 1995. *Le français parlé à Annonay (Ardèche)*, Paris : Klincksieck (« Matériaux pour l'étude des régionalismes du français ; 9 »).

- Gardette Pierre, 1941. *Géographie phonétique du Forez*, Macon : imprimerie Protat frères ; *Études de géographie morphologique sur les patois du Forez*, Macon : imprimerie Protat frères. [thèses]
- 1983. *Études de géographie linguistique* publiées par les soins de Brigitte Horiot, Marie-Rose Simoni, Georges Straka. Strasbourg : Société de Linguistique Romane (en dépôt : Klincksieck, Paris).
- Gauchat Louis, Jeanjaquet Jules, Tappolet Ernest, 1925. *Tableaux phonétiques des patois suisses romands*, Neuchâtel : P. Attinger.
- Gauchat Louis, Jeanjaquet Jules, Tappolet Ernest, Muret Ernest, 1924-. *Glossaire des patois de la Suisse romande*, Neuchâtel/Paris : P. Attinger ; Genève : Droz.
- Hafner Hans, 1955. *Grundzüge einer Lautlehre des Altfrankoprovenzalischen*, Berne : A. Francke, « Romanica Helvetica 52 ».
- Haselrot Bengt, 1937. *Étude sur les dialectes d'Ollon et du district d'Aigle (Vaud)*, Uppsala/Paris. [thèse Uppsala]
- 1938. « Sur l'origine des adjectifs possessifs *nostron*, *vostron* en franco-provençal », in : *Studia Neophilologica* XI, p. 62-84.
- 1966. « Les limites du francoprovençal et l'aire de *nostron* », in : *Revue de Linguistique romane* XXX, p. 257-266.
- Horiot Brigitte, 2002. « Voyage dans le temps d'un dialectophone lyonnais », in : *Revue belge de philologie et d'histoire*, tome 80, fascicule 3, p. 789-799.
- Horiot Brigitte, du Pouget Marc, 1990. « Un journal de recette du péage de Belleville au XV^e siècle », in : *Revue de Linguistique Romane* 54, p. 93-142.
- Jeanjaquet Jules, 1931. « Les patois valaisans : caractères généraux et particularités », in : *Revue de Linguistique Romane* VII, p. 23-51.
- Jud Jakob, 1973. *Romanische Sprachgeschichte und Sprachgeographie*, Zürich und Freiburg : Atlantis.
- Keller Hans-Erich, 1958. *Études linguistiques sur les parlers valdôtains*, Berne : A. Francke, « Romanica Helvetica 66 ». [thèse]
- Martin Jean-Baptiste, 1976. « Une caractéristique du francoprovençal : le pronom démonstratif neutre », in : *Mélanges de langues et de littératures romanes offerts à Carl Theodor Gossen*, Berne : A. Francke ; Liège : Marche Romane, tome II, p. 541-554.

- 1990. « Francoprovençal », in : *Lexikon der Romanistischen Linguistik* édité par Günter Holtus, Michael Metzeltin, Christian Schmitt, Tübingen : Max Niemeyer, volume V, 1, p. 671-685.
- Schmitt Christian, 1977. « A propos de la formation linguistique du domaine francoprovençal », in : *Revue de Linguistique Romane* 41, p. 91-103.
- Schüle Ernest, 1971. « Le problème burgonde vu par un romaniste », in : *Actes du colloque de dialectologie francoprovençale* (Neuchâtel, 23-27 septembre 1969) publiés par Zygmunt Marzys, avec la collaboration de François Voillat. Neuchâtel : Faculté des Lettres (« Recueil de travaux publiés par la Faculté des Lettres de l'Université de Neuchâtel, XXXIV ») ; Genève : Droz, p. 27-47.
- Tuaillon Gaston, 1972. « Le francoprovençal : progrès d'une définition », in : *Travaux de linguistique et de littérature* X, 1, p. 293-339.
- 1983. *Les régionalismes du français parlé à Vourey, village dauphinois*, Paris : Klincksieck (« Matériaux pour l'étude des régionalismes du français ; 1 »).
- 2001. *La littérature en francoprovençal avant 1700*. Grenoble : ELLUG.
- 2007. *Le francoprovençal*. Tome premier : définition et délimitation ; phénomènes remarquables. Quart (Vallée d'Aoste) : Musumeci éditeur.

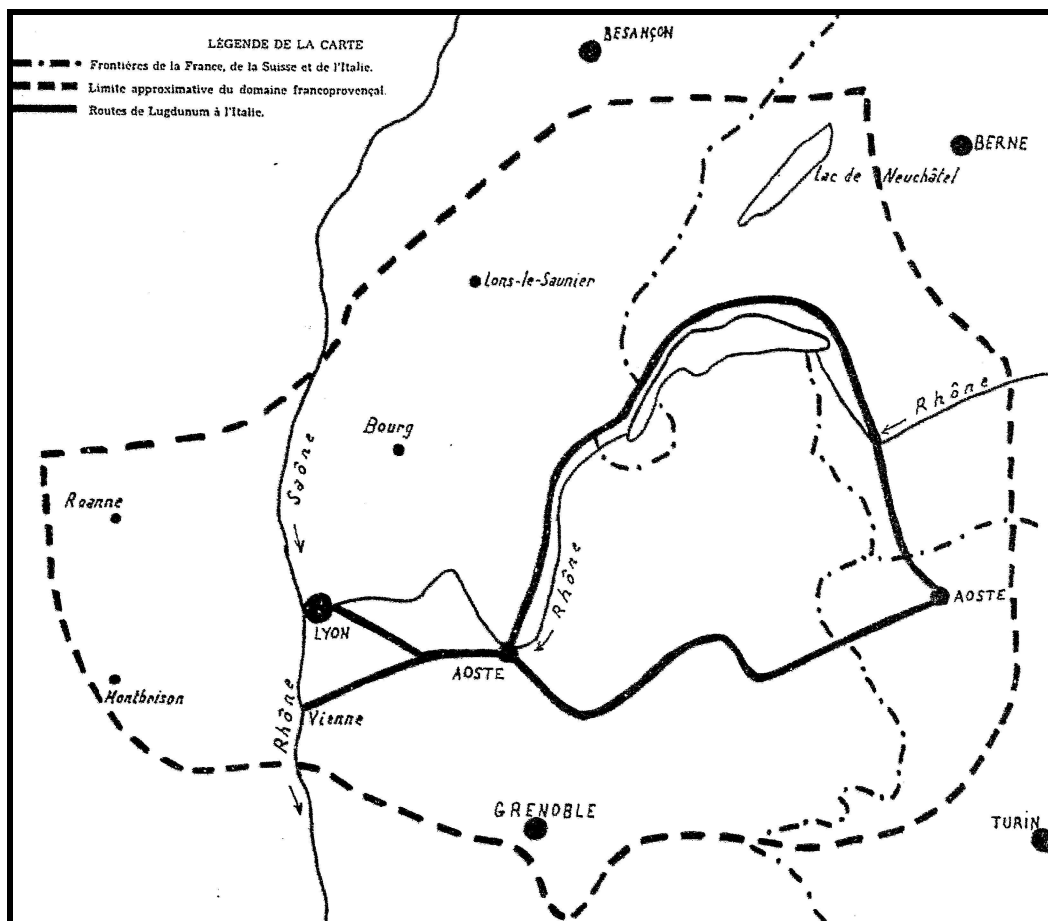
Carte 1

Gaston Tuailleon, « Le francoprovençal : progrès d'une définition », in : *Travaux de linguistique et de littérature* X, 1, 1972, p. 337.



Carte 2

Pierre Gardette, « Brève présentation du francoprovençal », in : *Etudes de géographie linguistique*, Strasbourg, 1983, p. 612.



*James Costa, ICAR
(UMR CNRS 5191, INRP & Université de Lyon)*

**Du local au global : essai de clarification
idéologique préalable**
*Discours concurrents et revitalisation linguistique en
Provence¹*

Introduction

La défense d'une langue vue comme menacée, comme tout phénomène social, suppose un certain nombre de prises de position nécessairement situées, socialement ou moralement autant que politiquement. Ces phénomènes, nécessairement discursifs, sont le fait de plusieurs types d'acteurs : linguistes, militants, locuteurs, enseignants etc. Nous faisons ici l'hypothèse, à la suite des travaux critiques de Deborah Cameron autour de la notion d'hygiène verbale² (Cameron, 1995, 2007), que les discours autour de la « mort des langues » sont la manifestation de questions sociétales n'ayant en réalité que très peu à voir avec des questions linguistiques, et qu'à travers eux s'expriment des conflits idéologiques liés aux enjeux sociaux d'une époque donnée.

Cet article, centré sur le contexte de la revitalisation linguistique en Provence, se propose d'étudier de manière critique (Blommaert & Bulcaen, 2000 ; Fairclough, 1992 ; Heller, 2002) trois documents militants participant à ce que Fairclough (1992) nomme une « chaîne

¹ Cet article doit beaucoup aux remarques et commentaires de Patricia Lambert (UMR ICAR) et Michel Bert (UMR ICAR & UMR DDL). Toute erreur ou omission est de ma seule responsabilité.

² L'hygiène verbale est définie par Cameron (1995) comme l'ensemble des tentatives délibérées d'agir sur une variété langagière, sous forme de prescriptions, de jugements, de propositions, d'alternatives etc.

intertextuelle », un dialogue suivi entre différents niveaux de textes, ici entre des discours locaux et, nous en faisons l'hypothèse, d'autres discours à visée plus globale. Les documents retenus visent à faire advenir une certaine vision de la langue régionale de Provence, un certain découpage du monde social. Nous essaierons de montrer comment le débat en cours sur une scène réduite sert de miroir à d'autres débats qui ont cours à un niveau plus global, et comment certains mouvements militants résistent de manière locale à d'autres discours qu'ils associent, à tort ou à raison, à des processus de globalisation (économique, culturelle) plus large.

Une telle démarche suppose « une identification des sources de pouvoir et de valeur, des manières de les distribuer, et des stratégies des acteurs autour de ces ressources, aussi bien que des conséquences » (Heller, 2002, p. 76).

Centré sur trois textes en particulier, notre article se base sur des informations recueillies en Provence par un travail de terrain de plusieurs années, comme militant et enseignant d'abord, puis comme sociolinguiste. Ce travail a débouché sur une thèse de doctorat en sociolinguistique sur les idéologies liées aux questions de revitalisation linguistique en Provence et en Ecosse.

En Provence, il existe aujourd'hui un vif débat idéologique à propos de la langue régionale qui y est parlée, entre militants comme entre linguistes (*cf.* par exemple l'échange entre P. Blanchet et D. Sumien dans la revue *Glottopol* : Blanchet, 2008 ; Sumien, 2009, pour un exemple de débat entre linguistes). Ainsi, une idéologie pan-occitane, qui tend à considérer que le provençal serait une variante locale d'une langue parlée sur l'ensemble du Sud de la France ainsi que dans certaines parties de l'Espagne et de l'Italie, interagit avec³ des idéologies qui traitent le provençal comme une langue à part entière, séparée de l'occitan et faisant partie d'un ensemble qui serait celui des langues d'oc, au pluriel. On peut considérer que cette confrontation idéologique prend sa source au 19^e siècle au moins, au temps de la Renaissance provençale autour du Félibrige et de Frédéric Mistral.

Cette confrontation, particulièrement violente depuis le début du 21^e siècle, concerne des enjeux de nomination et de délimitation de variantes linguistiques dans un contexte de déclin linguistique avancé.

³ Pour reprendre la terminologie de Field & Kroskrity, 2009, p. 25.

Un trait commun les rapproche néanmoins : les discours tenus aujourd'hui autour de ces langues cherchent à s'inscrire dans un contexte global de « mort des langues », ou de « langues en danger », pour reprendre la terminologie de l'Unesco (2009).

Notre travail se situe par ailleurs dans la lignée des travaux de Kroskrity (2009) sur la « clarification idéologique préalable », moment réflexif nécessaire avant toute réflexion en termes d'aménagement linguistique. Il s'agit, pour Kroskrity, d'identifier les questions liées à la langue et soumises à une contestation d'ordre idéologique⁴ dans un contexte donné. Plus spécifiquement en ce qui concerne cet article, nous cherchons à identifier les enjeux idéologiques fondateurs dans les processus de différenciation des mouvements militants linguistiques en Provence, et, de ce fait, dans la formation de discours concurrents concernant la revitalisation de la langue régionale. Nous cherchons en outre à comprendre à quels discours globaux ils se rattachent.

Pour ce travail, qui se veut exploratoire et ne prétend nullement à l'exhaustivité, nous analyserons les documents suivants (que nous reproduisons en annexes) :

- l'« appel à la manifestation⁵ » de Carcassonne (que nous appellerons désormais T1 dans le texte) émis en juin 2009 par une confédération de cinq associations militantes : Institut d'Estudis Occitans [IEO], Fédération des Enseignants de Langue et Culture d'Oc, la Fédération des Écoles Calandretas, « écoles occitanes associatives et laïques », Òc-Bi (« association de parents d'élèves pour l'enseignement bilingue public »), le Félibrige (« mouvement de défense de la Langue d'Oc fondé en 1854 par Frédéric Mistral »). Cet appel est disponible dans les six dialectes occitans reconnus par des linguistes comme Bec (1973), ainsi qu'en français. C'est cette dernière version que nous utiliserons pour nos analyses.
- l'appel à une « Grande manifestation pour la langue et la Culture Provençales⁶ » à Arles (désormais T2), diffusé par le Collectif

⁴ Par idéologie, nous entendons un ensemble de croyances socialement partagées, acceptées comme évidentes et universellement vraies, liées à des intérêts de pouvoir, social, politique, économique, matériel ou symbolique (Woolard, 1998).

⁵ <http://www.manifestarperloccitan.com/>, consulté le 08/09/2009.

⁶ <http://www.collectifprovence.com/spip.php?article141>, consulté le 08/09/2009.

Prouvènço, en vue d'un événement devant se tenir le 3 octobre 2009, ainsi que le « Manifeste pour la Provence » (2007), produit par la même association.

- Les pages d'ouverture du *Livre Blanc pour l'avenir des langues niçoise et provençale*, (désormais T3) préparé par la Fédération des Associations du Comté de Nice et l'Union Provençale (2003, pp. 2-3).

Les trois textes sont donnés en annexe à la fin de cet article.

Fonder une légitimité

Dans cette première partie, nous cherchons à montrer comment les principaux mouvements militants provençaux s'organisent les uns par rapport aux autres.

Ces textes ont en commun leur caractère rassembleur, en ce sens qu'ils prétendent faire converger un grand nombre de personnes autour d'idées fédératrices en termes de planification linguistique dans le but de revitaliser la langue dont ils se réclament, à travers deux manifestations publiques concurrentes, T1 et T2. Ils représentent par ailleurs un nombre important de militants, puisque T1 représente une fédération des principaux organismes militants historiques (dont le Félibrige et l'IEO), tandis que T2 se réclame de « près de 8000 adhérents, 150 associations et plus de 100 communes et communautés de communes en Provence⁷ » ; enfin, l'Union Provençale (T3) se présente comme regroupant 85 associations culturelles⁸.

Les deux premiers textes datent de 2009 mais ils reprennent un argumentaire soigneusement élaboré pour de précédentes manifestations, en 2005 et 2007 pour le premier document, en 2007 pour le second. Le troisième document, plus ancien, est néanmoins toujours d'actualité puisqu'il est encore présent sur divers sites internet, et par ailleurs repris en grande partie dans l'argumentaire en ligne de l'Union Provençale⁹. Ces textes nous semblent donc bien

⁷ Site du Collectif Prouvènço, <http://www.collectifprovence.com/spip.php?rubrique55>, consulté le 08/09/2009.

⁸ Site de l'Union Provençale : <http://prouvenco.presso.free.fr/mouvemen.html#up>, présentation de l'association, consulté le 08/09/2009.

⁹ Site de l'Union Provençale : <http://prouvenco.presso.free.fr/prouvencau.html>, présentation du provençal, consulté le 08/09/2009.

refléter l'ensemble des sensibilités principales officielles dans le débat actuel en Provence autour de la question de la langue régionale.

La question de la légitimité des mouvements militants est centrale en ce sens qu'elle conditionne le succès de leurs appels. Si le Félibrige ou l'IEO peuvent se revendiquer d'une légitimité historique, littéraire ou scientifique, les mouvements plus récents, comme le Collectif Prouvènço, doivent prouver la leur autrement.

La différence en termes de légitimation entre les documents émanant des associations plus anciennes (T3 et T1) ne tient pas à leurs références intertextuelles : ces deux textes introduisent leurs propos de manière classique dans le domaine des langues en danger par une référence de type écologique à la disparition annoncée de la diversité linguistique (thème récurrent depuis l'article de Hale, et al., 1992, et analysé par Cameron, 2007). Dans les deux cas, il s'agit de justifier une action par des références à un discours circulant de manière globale, et rendus de plus en plus disponible par leur disponibilité dans les médias.

Cette référence est présupposée connue dans T1 (la phrase : « La préservation de la diversité linguistique et culturelle dans le monde est un enjeu majeur pour les années à venir », est ici un simple rappel), mais moins dans T3, plus ancien, qui développe plus : « Il n'en reste pas moins que le provençal et le niçois sont en danger : ils sont mentionnés comme tels dans l'*Atlas des langues en péril dans le monde*, de l'UNESCO ». Dans le premier cas, l'accent est mis sur un processus, ici nominalisé : « la préservation » (cf. Fairclough, 1992). Dans le second cas, ce sont les langues qui sont mises en avant dans leur état de mise en danger. Dans l'un comme dans l'autre cas, les procédés stylistiques employés permettent d'éviter la question des agents responsables de ce fait, peut-être pour pouvoir désigner l'État comme responsable *in fine*. Il s'agirait dans ce cas de faire prendre en charge par l'institution la question du renversement de la substitution linguistique en cours, dans le cadre d'une réparation d'un dommage historique.

T3 et T1 se positionnent ainsi de manière similaire, idéologiquement parlant, face à l'État. La différence principale entre T3 et T1 porte donc surtout sur le choix des limites de langues et des désignations : quel(s) territoire(s) pour quelle(s) variété(s) langagières(s) ? Et par delà cette question : le provençal, langue autonome ou dialecte d'une autre langue ? Cette dernière question en

induisant nécessairement d'autres quant au rapport au monde des défenseurs de l'une ou l'autre position : les termes « dialecte » et « langue » sont en effet particulièrement connotés en français, négativement pour le premier, positivement pour le second.

Par ailleurs, dans les deux cas, nous sommes confrontés à des mouvements qui ont intégré la dimension moderniste de l'idéologie dominante qui iconise le lien entre langue et territoire : ainsi, si en France on parle français, en Provence on doit parler provençal, et à l'inverse, l'existence d'une langue occitane justifie l'existence de l'Occitanie.

Or, pour se développer, un mouvement militant doit être pensé en termes d'entitativité¹⁰ forte, particulièrement sur un marché fortement concurrentiel, où par ailleurs les profits symboliques et matériels sont relativement modestes. Dans ce contexte, le discours global, assimilant la question de la diversité linguistique à une question écologique, associant la problématique provençale à la problématique des langues en danger, peut permettre une première légitimation, mais il ne peut suffire à distinguer les acteurs, du fait de son acceptation partagée au sein des divers mouvements.

Le marché du militantisme provençal est déjà largement saturé : plusieurs associations sont installées depuis plusieurs dizaines d'années (1854 pour le Félibrige, 1945 pour l'Institut d'Études Occitane (cf. Abrate, 2001 ; Martel, 1987), pour ne citer que les plus anciens) et sont organisées autour de positions désormais établies à la fois en termes de choix graphiques normés, et de désignations des variétés linguistiques à l'intérieur de limites linguistiques communément admises. L'Union Provençale est certes plus récente, mais elle a néanmoins près de 30 années d'activité derrière elle.

Un nouvel acteur sur la scène militante : Le Collectif Prouvènço

Les questions qui se posent autour de l'entrée sur le marché discursif linguistique d'un nouvel acteur (le Collectif Prouvènço a été

¹⁰ L'entitativité (*entitativity*) est une notion développée en psychologie sociale par Hogg, et définie comme les propriétés d'un groupe qui le font apparaître en tant qu'entité cohérente, distincte et clairement structurée (Hogg, 2005, p. 206).

créé en 2000 : T3, p.10) sont donc autant « pourquoi ? » que « comment ? ». Les liens entre les mouvements porteurs de T3 et de T2 sont néanmoins aisément démontrables. Outre la présence du Collectif comme association mentionnée dans T3 (p.10), le site du Collectif Prouvènço propose à la signature du public, depuis 2007, un « Manifeste pour la Provence¹¹ » pour « obtenir la reconnaissance de la langue provençale » et « promouvoir le provençal comme langue de France à part entière ». Ce document s'ouvre de la manière suivante :

La langue provençale est vivante, à l'image de la culture qu'elle porte. La vitalité de l'édition, de la chanson, de la presse d'expression provençale, du théâtre ou des associations culturelles en est la démonstration.

Or, T3 s'ouvre de la même manière sur une phrase similaire :

La langue niçoise et la langue provençale sont bien vivantes, à l'image de la culture qu'elles portent. La vitalité de l'édition, de la chanson, de la presse d'expression provençale ou niçoise, du théâtre ou des associations culturelles en est la démonstration (p. 2).

Une première manière d'entrer sur le marché consiste ainsi à utiliser des références intertextuelles claires à destination des sympathisants de l'une des principales parties en présence (ici les sympathisants de T3, mais aussi plus largement ceux attachés à une définition locale de leur identité), en d'autres termes de construire une similarité avec l'existant.

Il s'agit dans un second temps de construire la différence, et de la faire accepter. La difficulté principale provient du fait que l'autorité (celle de la tradition, de l'histoire, de la littérature) se trouve, en ce qui concerne la Provence, largement établie du côté du Félibrige.

Cette légitimité est donc à aller chercher ailleurs. Ainsi, dans T2, la légitimité provient de la Loi elle-même, plus précisément de la Constitution, donc du peuple à travers ses représentants. Le document s'ouvre ainsi sur la phrase suivante :

¹¹ Manifeste pour la Provence, site du Collectif Prouvènço : http://www.collectifprovence.com/spip.php?article1&debut_signatures=90, consulté le 07/09/2009.

« En 2008, le congrès a modifié la Constitution, désormais : les langues régionales appartiennent au patrimoine de la France (art 75.1 de la Constitution française) ».

Cette reconnaissance en appelle une autre, puisqu'il s'agit de demander « la reconnaissance de la langue provençale comme l'une des LANGUES D'OC au sein des langues de France ». La juxtaposition des deux phrases implique ainsi une causalité, presque une nécessité, entre les deux événements, l'un fondant la nécessité de l'autre.

Cependant, l'autorité provient également, symboliquement, « des Provençaux » eux-mêmes, qui « sont majoritairement attachés au patrimoine linguistique et à l'identité de leur région provençale », et qui complèteraient par leur revendication, au niveau local, la décision de la représentation nationale.

La phrase suivante précise :

Pour se faire entendre et faire aboutir des revendications légitimes, l'Assemblée Générale du Collectif Prouvenço, réunie à Saumane de Vaucluse le dimanche 22 mars 2009 lance un appel à manifester le 3 octobre prochain entre les villes de Beaucaire et de Tarascon.

Un lien, dont la nature n'est pas précisée, est établi entre "les Provençaux" et « l'Assemblée Générale du Collectif Prouvenço », permettant de laisser penser que l'appel à manifester est la conséquence de l'attachement des « Provençaux » à leur langue. D'autre part, ce lien implicite laisse entendre que l'association, et son Assemblée Générale, auraient une légitimité à parler au nom des « Provençaux », sans que soit précisé qui est inclus sous cet ethnonyme. Ainsi ici, l'autorité ne viendrait donc pas directement de la tradition ou de l'existence d'une littérature prestigieuse et ancienne, comme pour le Félibrige, mais du peuple lui-même. Dans ce but, les « Provençaux » sont nécessairement homogénéisés, et un lien iconique de type moderniste est également créé entre la « région provençale », la langue, et le peuple.

Dans les documents sélectionnés, la différence entre mouvements militants concurrents n'est pas construite sur l'opposition la plus saillante, à savoir une distinction basée sur une utilisation de codes graphiques différents (*cf.* Bayle, 1968 ; Vernet, 1987). Le Félibrige et l'Institut d'Études Occitanes sont en effet regroupés dans le collectif organisant la manifestation de Béziers (T1) malgré des choix

graphiques différents, le Félibrige ayant tendance à favoriser le système mis au point au 19^e siècle lors de la Renaissance provençale félibréenne, l'IEO préférant un système basé sur une reconstruction d'un occitan administratif médiéval. Ce qui fonde ici la différence principale entre T2 et T1, c'est bien le cadre géographique de référence – la région administrative pour T3 et T2 (fait suffisamment rare en France pour être souligné) et un espace supra-régional pour T1 –, ainsi que les noms utilisés pour désigner la langue : provençal vs. occitan / langue d'oc.

Enjeux de la différence

Il est particulièrement intéressant de constater que les enjeux symboliques que recouvrent ces discours, semblent exclure la question de la bonne communication entre membres d'un groupe, qui est pourtant habituellement centrale dans les revendications des mouvements d'hygiène verbale (Cameron, 1995, pp. 23-25). Ceci peut se concevoir dans la mesure où la langue qu'il s'agit de défendre ici n'est plus le moyen exclusif de communication d'aucun groupe social, son déclin la confinant à des fonctions de plus en plus symboliques.

De fait, il est à noter que les locuteurs eux-mêmes sont singulièrement absents des textes de notre corpus. Les seuls acteurs sociaux représentés sont d'éventuels utilisateurs de l'occitan (T1 : « **ceux** qui le désirent [doivent pouvoir] l'utiliser », nous soulignons), les « promoteurs de la langue provençale » (T3, p. 2), les « enfants [des langues provençales et niçoises] » (T3, p. 2), « l'ensemble de la société provençale/niçoise » (T3, p. 3), les « Provençaux » (T2), et un « nous » indéterminé (Provençaux ou militants) dans T2, que l'on retrouve à plusieurs reprises : « nous sommes et nous serons », « Nous nous rassemblerons à Beaucaire ». Face à ces groupes non clairement identifiés, on trouve néanmoins, bien identifiée quant à elle, la puissance publique.

Cette absence de locuteurs, et ce peu de référence à une pratique linguistique semblent indiquer que s'il ne s'agit pas à proprement parler de revendiquer une langue de peur que la communication entre ses locuteurs ne devienne imparfaite au point de risquer la dislocation d'un groupe. Il s'agit pourtant bien ici de communiquer, selon nous, non pas *avec quelqu'un*, mais *quelque chose*.

Cet élément à communiquer nous semble tenir à la fois d'une dimension visant à imaginer le groupe que l'on cherche à faire exister, pour reprendre la terminologie d'Anderson (1983), en même temps que d'une conception politique de l'avenir de ce groupe.

T3 et T1, de par leurs références à la diversité linguistique, et de par leurs revendications adressées au principal centre du pouvoir (Woolard, 2004) en France, semblent *a priori* s'inscrire dans un rapport classique de revendication périphérie vs. centre, ou, pour reprendre la terminologie de Martel, dans la problématique Nord vs. Sud, idée récurrente dans le discours politique et historique français (Martel, 2002).

Le discours de T2 nous paraît quant à lui plus complexe. Nous avons vu qu'il utilise des stratégies de distinction basées sur une légitimité qui serait populaire plutôt qu'historique ou littéraire, tout en s'inspirant largement du discours de T3. L'antagonisme entre militants « provençalistes » et « occitanistes » est certes ancien en Provence (cf. Bayle, 1968, 1979, 1982), mais le phénomène de manifestations concurrentes est, lui, nouveau. La signification même de l'acte de manifestation comme acte de magie sociale (Bourdieu, 1980) appelle une analyse plus poussée.

Le discours de T2 qui s'inscrit dans cette continuité de la querelle graphique et passe par un processus de construction discursive d'un Autre négatif et potentiellement menaçant. Considérons à présent l'extrait suivant de T2 :

« La reconnaissance des Langues d'Oc est plus que jamais à l'ordre du jour. La **globalisation** occitane ne peut conduire qu'à l'échec de la mise en œuvre d'une politique régionaliste **authentique** » [l'ensemble de ce passage est en gras dans le texte original, mais les termes ici soulignés sont de notre fait].

Ce passage clef, souligné par l'utilisation de caractères en gras, nous semble en fait le cœur même du texte, et le cœur même du débat. Il permet de définir ce que doit être une bonne politique « régionaliste », à savoir une politique « authentique ». Si ni l'un ni l'autre de ces termes ne sont définis, ils sont cependant opposés dans le texte à « la globalisation occitane », qui mettrait en échec cette « politique régionaliste authentique ». L'utilisation du terme « globalisation » fait écho à un phénomène plus connu en français sous le nom de mondialisation. De fait, l'opposition « authenticité » vs. « mondialisation » est une opposition politique désormais

classique qui présuppose une vision d'une authenticité enracinée dans le local et la tradition face à une mondialisation vue comme désincarnée et menaçante, source de désordre, de décomposition (cf. Ali-Khodja & Boudreau, 2009) et de renversement de valeurs. D'autres textes du mouvement à l'origine de T2 soutiennent par ailleurs le refus du port de la bombe pour les cavaliers en Camargue et défendent le port du costume arlésien comme signe valorisant d'une identité provençale rhodanienne authentique. Il ne s'agit naturellement pas pour nous de chercher à stigmatiser un mouvement ou des pratiques culturelles spécifiques, mais de chercher à comprendre à quels discours ceux-ci se rattachent. La langue semble bien, dans ce contexte, vue comme menacée par des pratiques considérées comme extérieures, de la même manière que le sont d'autres pratiques culturelles.

On pourrait interpréter ces positionnements comme les avatars modernes du débat gauche / droite, l'autre type de discours central dans le traitement discursif de la question « occitane » selon Martel (2002). Cette explication nous semble néanmoins réductrice par rapport à des clivages en cours de redéfinition face à des enjeux internationaux forts et encore largement sous-étudiés.

La « globalisation » invoquée dans ce texte convoque ainsi un certain nombre d'images négatives (délocalisations, pertes d'emploi, oubli des traditions, pertes de repères traditionnels, disparition de certains modes de vie). En associant ce terme à celui d'« occitan », T2 peut espérer rallier certaines oppositions à une « globalisation » économique et culturelle anonyme en désignant et en identifiant un ennemi commun. Potentiellement, le texte cherche à en rallier les victimes (ouvriers, paysans, salariés) face aux supposés bénéficiaires, les classes moyennes et/ou supérieures, que représenterait la tendance inverse (le mouvement occitaniste est décrit sur le site du Collectif Prouvènço comme un mouvement de « bobos »¹²). Ailleurs, le Collectif Prouvènço reproche par exemple au mouvement occitaniste d'avoir « horreur du costume provençal et de son folklore¹³ ». La

¹² Cf. « Nous en avons assez de ces "Bo-b'Oc" qui vous disent que la Provence n'est qu'une partie de la grande Occitanie dont la langue se dit "langue d'oc" au singulier, ou "provençal occitan », <http://www.collectifprovence.com/spip.php?article107>, consulté le 04/01/2010.

¹³ Revue *Me dison Prouvènço* (publiée par l'association Collectif Prouvènço) n°16, Janvier 2007, p. 12.

distinction avec les autres discours militants s'opère donc en partie par la recontextualisation d'un discours politique contemporain disponible sur la scène nationale, susceptible d'attirer une sympathie large parmi les classes populaires, premières victimes des phénomènes décrits plus hauts.

Si l'on peut penser que ce type de discours représente ce que Fairclough appelle une « stratégie manipulatrice » (Fairclough, 1992), en vue d'imposer un classement du monde particulier sur un marché limité, nous voudrions ici suggérer une possible seconde explication. Ce discours rappelle en effet ce que Cameron nomme « l'horreur de l'homogénéisation » :

[w]hat animates today's concerns about language endangerment and 'death' might be thought of as a 'horror of homogeization', rooted [...] in contemporary geopolitical realities / [c]e qui anime les préoccupations contemporaines sur les langues en danger et leur mort peut être pensé comme une 'horreur de l'homogénéisation', qui prend racine dans les réalités géopolitiques contemporaines (2007, p. 282).

Le débat sur la question linguistique cristallise, et permet d'exprimer, des peurs et des angoisses, qui, bien que réelles, ne sont pas immédiatement liées à des questions linguistiques. Le domaine de la langue permet l'expression d'un débat qui, dans un autre contexte, devrait s'exprimer par d'autres conceptualisations peut-être plus délicates à manier, et qui reflète les opinions de certains groupes percevant la mondialisation, phénomène complexe et multiforme (Fairclough, 2006), comme profondément anxiogène.

Ces groupes investissent en conséquence des domaines comme la langue et ses rapports à l'identité, dans le cadre d'une idéologie dominante (tant au sein du discours autour de la langue dominante qu'au sein des divers discours militants en présence) liant de manière iconique langue et identité. Il ne s'agit donc pas pour eux de remettre en cause cet aspect de l'idéologie dominante, mais de demander une reconnaissance perçue comme source d'apaisement face à des phénomènes globaux jugés incontrôlables et angoissants dans leur remise en cause d'éléments d'un mode de vie plus traditionnel et authentique.

On peut voir dans ces discours une manière de reprendre prise sur ces phénomènes, par l'association discursive occitan / mondialisation, le mouvement occitan symbolisant, en tant que bouc émissaire, l'ensemble de ces processus, réels ou supposés. Cette association,

purement symbolique compte tenu des faibles enjeux véhiculés par les langues régionales en France, permet donc à moindre frais de s'attaquer à un phénomène autrement large, complexe et potentiellement effrayant dans lequel la disparition des langues locales n'est qu'un élément parmi d'autres.

Paradoxalement néanmoins, cette peur de l'homogénéisation face à ce qui peut apparaître comme des mouvements anonymes passe également par une homogénéisation de l'*in-group* : T2 précise ainsi que « les provençaux sont majoritairement attachés au patrimoine linguistique et à l'identité de leur région provençale ». Cet argument, présenté comme une vérité partagée et généralement admise, a pour principal effet de présenter les Provençaux comme un groupe homogène, possédant une identité stable reconnue à l'intérieur comme à l'extérieur de ce groupe.

Lorsque Cameron (2007) parle de peur de l'homogénéisation et de la mondialisation, elle fait référence à des réactions occidentales face à la perte de la diversité linguistique mondiale vécue comme un phénomène abstrait et lointain. Ici, c'est d'un phénomène sensiblement différent qu'il s'agit : nous sommes confrontés à des groupes qui ressentent comme une perte réelle la disparition de certains usages traditionnels, et qui, par le discours, expriment leurs angoisses et tentent de résister et de construire un avenir alternatif.

Ces discours sont donc particulièrement intéressants en ce sens qu'ils sont un indice de peurs qui traversent la société française, et sans doute européenne, de manière plus générale ; il nous semble particulièrement important de ne pas les ignorer, mais au contraire de les étudier plus en détail afin de mieux cerner les tensions dont ils sont représentatifs.

Conclusion

Nous avons vu dans cet article comment une langue régionale est mise en scène de diverses manières, ici concurrentes, nécessairement complexes, mêlant diverses voix et discours. Cet acte de représentation aboutit ici à une représentation de la langue sur la scène publique à travers l'organisation de manifestations.

Nous avons pu mettre en évidence néanmoins, que les luttes idéologiques présentes dans ces discours étaient symptomatiques de préoccupations bien éloignées des questions uniquement linguistiques

qu'elles prétendent mettre en avant. On voit se refléter dans les oppositions discursives que nous avons mises en évidence des questions sociales profondes et complexes, ainsi que des réactions d'angoisse face à des phénomènes tout aussi complexes de mondialisation économique.

A travers les discours sur la langue, ce sont aussi potentiellement des enjeux de classes qui apparaissent, face à une mondialisation vue par certains acteurs comme largement favorable aux classes moyennes ou dominantes. La langue offre par ailleurs un terrain discursif particulièrement favorable à ce type de discours en ce sens qu'elle permet l'expression de points de vue concurrents sur l'authenticité, de fait sur l'identité de ses locuteurs, et sur les instances ou les personnes habilitées à dire qui est un locuteur authentique, un bon locuteur, qui est un provençal ou pas.

Ces enjeux, dont le repérage relève bien de ce que Kroskrity (2009) nomme « clarification idéologique préalable », ne relèvent donc pas exclusivement de problématiques linguistiques. Par conséquent, il n'est pas certain que, malgré ce qu'écrit Kroskrity (2009, p. 80), l'identification d'idéologies concurrentes, et de leurs enjeux, permette d'éviter ou de minimiser les conflits du type de celui que nous avons analysé ici, et permette la mise en place de stratégies de revitalisation plus efficaces. Les enjeux symboliques, comme la conquête d'un CAPES spécifique de provençal ou la reconnaissance du provençal comme langue de France, par ailleurs sans conséquence politique réelle dans l'état actuel de la politique linguistique française, sont à la fois trop dérisoires – et précisément de ce fait trop importants – pour faire l'objet d'une conciliation.

Ces enjeux révèlent *in fine* qui est légitime pour produire les divisions du monde social et qui a le pouvoir de distribuer le pouvoir au sein de la société provençale actuelle, même si la société provençale en question tend à être de plus en plus réduite. Il faudrait ainsi creuser la question de la répartition et de la distribution des pouvoirs dans le cadre des « langues très en danger », langues dont la pratique tend à n'être plus que symbolique. Les enjeux peuvent ici apparaître minimes, mais ils sont symptomatiques d'autres enjeux, plus largement répandus dans nos sociétés occidentales, comme la résistance aux phénomènes de globalisation économique et culturelle.

Dans ces discours, qui sont désormais entrés dans une phase de compétition d'autant plus vive que la langue finit de disparaître de ses

derniers domaines d'usages traditionnels, on a noté l'absence totale de caractérisation des locuteurs. Le trois discours analysés sont des discours visant une action en termes de macro-politique linguistique, basée sur une approche en termes de besoins qui seraient identifiés de manière objective, bien que ces politiques aient largement échoué jusqu'à présent (Canagarajah, 2006, p. 154). L'approche en termes de « langue » et non de « locuteurs » ou de « processus », a par ailleurs pour conséquence importante le gommage d'autres types de conflits sociaux, liés par exemple aux rapports de domination entre classes, et à ériger d'autres conflits là où ce sont au contraire des solidarités de classe qui prévalaient par delà les différences linguistiques (cf. Merle, 1977, pour des exemples de discours ouvriers sur les questions linguistiques en Provence).

Si cet article a pu constituer une première approche en vue d'une clarification idéologique préalable des mouvements de revitalisation linguistique en Provence, cela ne doit pas faire oublier que les discours retenus ici sont des discours militants écrits, destinés à être diffusés auprès d'un public large. Ils sont donc particulièrement marqués en termes d'idéologie, et ne sont pas nécessairement représentatifs de *tous* les comportements et discours individuels observables sur le terrain chez les locuteurs, y compris chez les néo-locuteurs, et même chez les militants. Notre travail se poursuivra donc par une analyse de discours recueillis au cours d'une enquête ethnographique visant à mieux les appréhender dans leur diversité afin de mieux comprendre les variations idéologiques des acteurs sociaux engagés dans ce processus.

Références

- Abrate, L. (2001). *Occitanie 1900-1968 : des idées et des hommes*. Puylaurens : Institut d'Etudes Occitanes.
- Ali-Khodja, A., & Boudreau, A. (2009). « Du concept de minorité à la pensée de l'exiguïté : pour une autre compréhension des phénomènes linguistiques ». *Langage et Société*, 2009 / 3(129).
- Anderson, B. (1983). *Imagined communities : reflections on the origin and spread of nationalism*. London : Verso.
- Bayle, L. (1968). *Dissertation sur l'orthographe provençale (comparée à la graphie dite occitane)*. Toulon : L'Astrado.
- Bayle, L. (1979). *Huit entretiens sur l'occitanisme et les Occitans*. Toulon : L'Astrado.
- Bayle, L. (1982). *La Provence en danger* (Second dossier occitan). Toulon : L'Astrado.
- Bec, P. (1973). *La langue occitane*. Paris : PUF.
- Blanchet, P. (2008). Compte rendu de Dominique Sumien, *La standardisation pluricentrique de l'occitan*. *Glottopol*, 11, 171-180.
- Blommaert, J., & Bulcaen, C. (2000). 'Critical Discourse Analysis'. *Annual Review of Anthropology*, 29, 447-466.
- Bourdieu, P. (1980). « L'identité et la représentation : éléments pour une réflexion critique sur l'idée de région ». *Actes de la Recherche en Sciences Sociales*, 35(1), 63-72.
- Cameron, D. (1995). *Verbal Hygiene*. London : Routledge.
- Cameron, D. (2007). 'Language endangerment and verbal hygiene: History, morality and politics'. In A. Duchêne & M. Heller (Eds.), *Discourses of Endangerment* (pp. 268-285). London & New York : Continuum.
- Canagarajah, S. (2006). 'Ethnographic Methods in language Policy'. In T. Ricento (Ed.), *Language Policy: Theory and Method* (pp. 153-169). Oxford : Blackwell.
- Fairclough, N. (1992). *Discourse and Social Change*. Cambridge : Polity.
- Fairclough, N. (2006). *Language and Globalization*. Abingdon : Routledge.
- Fédération des Associations du Comté de Nice, & Union Provençale (2003). Livre blanc pour l'avenir des langues niçoise et provençale. <http://www.comte-de-nice.org/images/stories/livreblanc.pdf>

- Field, M. C., & Kroskrity, P. V. (2009). 'Revealing Native American Language Ideologies'. In P. V. Kroskrity & M. C. Field (Eds.), *Native American Language Ideologies: Beliefs, Practices, and struggles in Indian Country* (pp. 3-30). Tucson : University of Arizona Press.
- Hale, K., Krauss, M., Watahomigie, L. J., Yamamoto, A. Y., Craig, C., Jeanne, L. M., et al. (1992). 'Endangered Languages'. *Language*, 68 (1), 1-42.
- Heller, M. (2002). *Eléments d'une sociolinguistique critique*. Paris : Didier.
- Hogg, M. A. (2005). 'Uncertainty, Social Identity and Ideology'. In S. R. Thye & E. J. Lawler (Eds.), *Social Identification in Groups*. Amsterdam : Elsevier.
- Irvine, J., & Gal, S. (2000). 'Language ideology and linguistic differentiation'. In P. V. Kroskrity (Ed.), *Regimes of language: Ideologies, politics, and identities* (pp. 35-83). Oxford : James Currey.
- Kroskrity, P. V. (2009). 'Language Renewal as Sites of Language Ideological Struggle. The Need for "Ideological Clarification"'. In J. Reyhner & L. Lockard (Eds.), *Indigenous Language Revitalization: Encouragement, Guidance & Lessons Learned* (pp. 71-83). Flagstaff (Arizona) : Northern Arizona University.
- Martel, P. (1987). « La résistible ascension de la loi du félibrige en Provence (1854-1914) ». *Graphies, idéologies linguistiques du XVI^e au XX^e siècle (Provence, Dauphine, Vallées occitanes d'Italie)* (pp. 61-73). Marseille : Obradors Occitans en Provença.
- Martel, P. (2002). *Les Cathares et l'Histoire : Le drame cathare devant ses historiens (1820-1992)*. Toulouse : Privat.
- Merle, R. (1977). *Culture occitane per avançar*. Paris : Editions sociales.
- Sumien, D. (2009). « La standardisation pluricentrique de l'occitan » : Domergue Sumien répond à Philippe Blanchet *Glottopol*. <http://www.univrouen.fr/dyalang/glottopol/telecharger/comptesrendus/reponse%20sumien.htm>
- UNESCO (2009). UNESCO Interactive Atlas of the World's Languages in Danger. 20 juillet 2009, site UNESCO : <http://www.unesco.org/culture/ich/index.php?pg=00206>
- Vernet, F. (1987). « Dix ans de querelle graphique ». *Graphies, idéologies linguistiques du XVI^e au XX^e siècle (Provence,*

Dauphine, Vallées occitanes d'Italie) (pp. 83-87). Marseille : Obradors Occitans en Provença.

Woolard, K. A. (1998). 'Language Ideology as a Field of Inquiry'. In B. B. Schieffelin, K. A. Woolard & P. V. Kroskrity (Eds.), *Language Ideologies. Practice and Theory* (pp. 3-47). Oxford & New York : Oxford University Press.

Woolard, K. A. (2004). 'Is the Past a Foreign Country? Time, Language Origins, and the Nation in Early Modern Spain'. *Journal of Linguistic Anthropology*, 14(1), 57-80.

Annexes

1) Anem Òc ! per la lenga occitana !¹⁴ (T1)

Appel pour la manifestation du 24 octobre 2009 à Carcassonne

La préservation de la diversité linguistique et culturelle dans le monde est un enjeu majeur pour les années à venir. En France, la situation légale est pourtant défavorable à cette préservation. La langue occitane, notamment, est particulièrement menacée par l'absence de politique positive pour sa transmission.

Il faut donc, sur le territoire où se parle l'occitan, créer un environnement favorable à l'enseignement de la langue et dans la langue, à la création culturelle en occitan, à la présence de la langue dans les médias électroniques, audiovisuels et écrits, à son emploi dans la vie publique et sociale.

Tout cela sera un encouragement à la transmission familiale de l'occitan.

Les médias

Nous voulons un service public de radio et de télévision en langue occitane. Nous demandons une aide spécifique pour les opérateurs privés (radios associatives, télévisions, presse) qui travaillent en faveur de la langue occitane.

La création

Le soutien à la création, c'est une politique spécifique en faveur de l'édition, du théâtre, du spectacle vivant en général, du cinéma et de la musique. La population doit pouvoir découvrir une culture trop souvent occultée.

L'enseignement

Il faut que chacun puisse choisir s'il le souhaite un enseignement sous des formes diverses :

- un enseignement en occitan (bilingue à parité horaire ou en immersion) ;
- un enseignement de la langue et de la culture occitanes ;
- un enseignement spécifique en direction des adultes

Il s'agit d'une politique d'offre généralisée.

La vie publique

L'occitan doit être présent dans la vie publique. Cela veut dire par exemple :

- le respect de la toponymie et la mise en place d'une signalisation bilingue ;
- l'encouragement à l'utilisation de la langue occitane dans les lieux publics et dans la vie sociale.

¹⁴ *allons oui ! pour la langue occitane !*

Toutes ces demandes ont pour seul but de permettre à une langue de vivre et à ceux qui le désirent de l'utiliser. Elles s'adressent à la fois aux collectivités locales occitanes (communes, communautés de communes, départements, régions) et à l'Etat. Chacun doit prendre en charge ce qui est de sa responsabilité. Il appartient à l'état d'adapter sa législation.

Coordinacion Anem Òc !

Institut d'Estudis Occitans, organisme culturel reconnu d'utilité publique

F.E.L.C.O. Fédération des Enseignants de Langue et Culture d'Oc

Calandreta, écoles occitanes associatives laïques

Òc-Bi, association de parents d'élèves pour l'enseignement bilingue public

Felibrige, mouvement de défense de la Langue d'Òc fondé en 1854 par Frédéric Mistral

2) Grande manifestation pour la langue et la Culture Provençales (T2)

Sian e saren

Grande manifestation

“Sian e saren” “nous sommes et nous serons”

Contre une pensée et une langue unique

Samedi 3 octobre 2009, de 14h00 à 18h00.

En 2008, le congrès a modifié la Constitution, désormais : les langues régionales appartiennent au patrimoine de la France (art 75.1 de la Constitution française).

Le Collectif Prouvènço considère que c'est un pas décisif vers la reconnaissance de la langue provençale comme l'une des LANGUES D'OC au sein des langues de France.

Une loi devrait voir le jour dans les mois à venir. Celle-ci listera l'ensemble des langues régionales de France. Il existe un risque important que le provençal comme l'ensemble des autres langues d'oc soient classées comme sous-dialecte de l'occitan ou d'une langue d'Oc unique.

Les provençaux sont majoritairement attachés au patrimoine linguistique et à l'identité de leur région provençale. Pour se faire entendre et faire aboutir des revendications légitimes, l'Assemblée Générale du Collectif Prouvènço, réunie à Saumane de Vaucluse le dimanche 22 mars 2009 lance un appel à manifester le 3 octobre prochain entre les villes de Beaucaire et de Tarascon.

L'assemblée Générale appelle toutes les associations, tous les provençaux, l'ensemble des pays d'Oc, à se joindre à cette manifestation pour faire de cette journée un temps fort. La reconnaissance des Langues d'Oc est plus que jamais à l'ordre du jour. La globalisation occitane ne peut conduire qu'à l'échec de la mise en oeuvre d'une politique régionaliste authentique

Nous nous rassemblerons à Beaucaire, puis nous nous dirigerons en cortège pacifique vers Tarascon où des prises de paroles et des concerts seront proposés.

Ensemble nous manifesterons pour :

*la reconnaissance de la langue provençale comme l'une des langues de France

*la prise en compte de la pluralité des langues d'Oc (béarnais-gascon, languedocien, limousin, auvergnat, provençal, cévenol et nissart)

*la création de CAPES spécifiques "Langues d'Oc" avec options provençal, béarnais, languedocien, limousin, auvergnat, nissart, cévenol...

*la création d'un Observatoire de la Langue et de la Culture provençales

*la création de véritables émissions de radios et de télévision en langue provençale, dans notre zone territoriale

* la mise en place d'une double-signalétique français/provençal étendue à l'ensemble des panneaux administratifs et de signalisation routière

Ensemble, dans la diversité nous sommes forts.

SIAN E SAREN !

Pèr lou Couleitiéu Prouvènço, Jan-Pèire Richard

3) Pourquoi un livre blanc ? (T3)

La langue provençale et la langue niçoise sont bien vivantes, à l'image de la culture qu'elles portent. La vitalité de l'édition, de la chanson, de la presse d'expression provençale ou niçoise, du théâtre ou des associations culturelles en est la démonstration.

Le succès d'audience des émissions en provençal et niçois de France 3 Méditerranée le confirme.

Malgré les handicaps qu'a pu constituer une certaine évolution de la société dans le cadre de la politique longtemps hostile de l'État, le provençal et le niçois ont bénéficié de l'attachement que leur vouent les habitants de leurs aires géographiques quelle que soit leur origine, évitant ainsi que ces langues ne deviennent résiduelles et soient en voie de disparition. Ces dernières années, une prise de conscience accrue de l'importance de nos langues s'est manifestée dans notre région par une mobilisation sans précédent des populations et associations, ainsi que des collectivités territoriales et des élus.

Il n'en reste pas moins que le provençal et le niçois sont en danger. Ils sont mentionnés comme tels dans l'*Atlas des langues en péril dans le monde*, de l'UNESCO (1996, réédition 2002). Leur transmission familiale est de moins en moins assurée, leur enseignement reste très limité et leur place dans la vie quotidienne s'efface progressivement, y compris dans le domaine privé auquel elles ont été longtemps cantonnées.

Il y a donc urgence pour nous, promoteurs de la langue provençale, de la langue niçoise et du plurilinguisme en général, à définir les actions prioritaires qu'il nous faut mener pour redonner aux langues de Provence, à côté du français, la place qu'elles méritent. Ce faisant, nous voulons sauvegarder le patrimoine qu'elles constituent et permettre à ses enfants, qu'ils y aient leurs origines ou non, de s'enrichir de la culture et de la langue du coin du monde où ils vivent.

Les langues provençale et niçoise sont en effet profondément liées à leur aire géographique, historique et culturelle. C'est de là qu'elles tiennent leur nom, et par là que passe la conscience linguistique et culturelle de leurs usagers. Elles sont indissociables de leur territoire. Elles ont des rapports privilégiés avec les autres langues du domaine linguistique d'oc.

La langue provençale, tout comme la langue niçoise, sont riches de leurs variétés, qui sont d'égale valeur hors de toute hiérarchie. La disparition de l'une de ces variétés serait une perte irremplaçable pour le patrimoine culturel de la région et, plus largement, de l'humanité.

La philosophie étant ainsi réaffirmée, rappelons le sens de la démarche engagée par le mouvement culturel provençal et niçois.

Depuis une vingtaine d'années, un important travail de recherche scientifique, de concertation associative et d'élaboration de propositions a été réalisé. Plusieurs États généraux des langues provençale et niçoise ont ainsi permis de dresser l'état des lieux, de faire le bilan des actions, et de lancer quelques pistes d'avenir.

Il paraissait important d'aller plus loin dans la démarche en précisant quelles étaient les priorités pour assurer notamment l'avenir des langues provençale et

niçoise et la richesse plurilingue de notre région. Se fixer des objectifs et se donner les moyens de les mettre en œuvre. En demandant aux pouvoirs publics de jouer leur rôle dans une véritable politique de développement des langues provençale et niçoise, mais aussi en amenant le mouvement culturel provençal et niçois à s'engager clairement sur un certain nombre d'actions qu'il estime primordiales et en y associant la société civile. Les associations sportives, les entreprises industrielles, agricoles et de services, les médias, les éditeurs, les commerçants [...] ont aussi leur place dans cet engagement pour l'avenir des langues provençale et niçoise.

Ce Livre blanc, élaboré au terme d'une riche concertation au sein du mouvement culturel provençal et niçois, sera présenté à l'occasion des premières Assises nationales des langues de France, le 4 octobre 2003 à Paris.

Mais la démarche ne s'arrêtera pas là. Elle devra se poursuivre pour s'élargir aux acteurs de la société civile, car tout un chacun, quelle que soit son origine, est concerné par le devenir des langues provençale et niçoise.

Le provençal et le niçois sont les langues de la région Provence-Alpes-Côte d'Azur. Le provençal est aussi la langue de l'Ardèche, de la Drôme et du Gard provençaux.

Elles deviendront ce que l'ensemble de la société provençale/niçoise en fera. L'heure est à la prise de conscience, à la mobilisation, à des mesures concrètes, efficaces et adaptées.

l'astrado prouvençalo - chambro sendicalo de la prèssu prouvençalo - collectif prouvènço - conseil culturel du comté de nice - association " éditions dictionnaires français-provençal " - fédération des associations du comté de nice - fédération folklorique méditerranéenne - irep-come - parlaren en vaucluso - le pays gavot - prouvençau a l'escolo - prouvènço-bus - unioun dis escrivan prouvençau - unioun prouvençalo

Francis Manzano
Université Jean Moulin Lyon 3 – CEL, EA 1663

Patrimoines langagiers méditerranéens : modèles de gestion et perspectives

1. Paradoxes méditerranéens

Les langues de la Méditerranée occidentale appartiennent principalement aux grandes familles indo-européenne (nord¹) et chamito-sémitique (sud²). Rares sont les langues n'appartenant pas à ces familles, soit en synchronie (comme le basque), soit en diachronie, comme l'étrusque, le ligurien etc. Pour cette raison, de telles langues (quand elles existent encore) n'en ont que plus d'aura, car résister sociolinguistiquement dans cette région, survivre, n'est pas une mince affaire³.

Ce phénomène d'ensemble n'est que rarement commenté. Il est pourtant révélateur d'une région du monde cul-de-sac où les deux ensembles se rencontrent et dans laquelle la pression des langues dominantes tend irrémédiablement à broyer les écarts et les reliques, tandis que les langues véhiculaires elles-mêmes se retrouvent dans un système très tendu de concurrence. Dans cette thématique et cette métaphore de la pression, il n'est pas sans intérêt de constater combien de langues internationales sont issues de ce secteur, comme expulsées vers l'extérieur dans un mouvement centrifuge ancien, latin

¹ Dans cet ensemble se met en relief la sous-famille romane, majoritaire dans l'ensemble de la région.

² Arabe et berbère en Afrique du Nord.

³ Cette remarque, si l'on sort de l'aspect proprement typologique, concerne différentes langues exemplaires et symboliquement valorisées comme le basque (déjà cité, en France et en Espagne), le berbère (au Maghreb), le breton (en France), bien que celui-ci n'intervienne pas dans la zone méditerranéenne *stricto sensu*. La solidité de ces langues est variable mais au fond, plus l'on se rapprochera (éventuellement) des derniers locuteurs, plus forte sera aussi l'aura symbolique de telles langues.

en premier lieu, puis portugais, français, castillan, italien etc.⁴ Par ailleurs une langue non romane d'importance (l'anglais) s'est glissée dans le système et elle a toutes chances de jouer un rôle de plus en plus important. Enfin, sur les rivages méridionaux, l'arabe se présente *de facto* comme le véhiculaire fondamental.

Cette propension à produire des véhiculaires au sein d'un paysage sociolinguistique très diversifié (qu'on pourrait qualifier de pôle dynamique ou du changement), est pourtant indissociable d'une très grande capacité au conservatisme (ou, à l'opposé, pôle statique). Sans doute faut-il également souligner que cette capacité au conservatisme de langues diverses ne peut être isolée d'un conservatisme plus général en matière de culture collective, d'aménagement des territoires etc., constituant une dualité régulièrement mise en avant dans le monde méditerranéen et vue (peut-être à tort, mais non sans raisons objectives) comme une lutte éternelle entre l'ancien et le nouveau. Un schisme que la littérature, le cinéma ont bien exploité. Le cas du berbère est intéressant de ce point de vue car le lien entre conservation de cette langue et conservation des schémas d'organisation sociétale est particulièrement patent, comme j'ai tenté de le montrer à partir de certains systèmes onomastiques (Manzano 2006c). On pourrait ainsi broser une représentation typique des conservatismes méditerranéens qui relie la Grèce à l'Andalousie, en passant par la Corse, la Sardaigne etc. Le cas du berbère, effectivement, mérite d'être mis en relief car cette langue a survécu au punique, au latin, à diverses langues romanes, et résiste encore à l'arabe sous nos yeux (on se situe sur une échelle qui s'exprime en millénaires). Bien entendu cela vaut aussi, à des degrés divers, pour des langues minoritaires vedettes comme le basque, le breton, le corse etc.

On ne reprendra pas dans ce cadre le tableau d'ensemble des langues régionales ou minoritaires du bassin méditerranéen occidental, d'autant que l'idée d'une typologie close des situations des langues en recul ou en réhabilitation est probablement illusoire. De

⁴ Cette dynamique d'internationalisation est parfaitement attestée pour les premières de ces langues, sous nos yeux encore. En revanche, des langues aujourd'hui réputées minoritaires ont pu historiquement démarrer un processus de diffusion, avorté par la suite. Illustre bien ce cas de figure le catalan, transporté au cours du Moyen Age vers l'Italie et la Sardaigne et dont il reste des traces historiques notables (lexicales notamment). Le catalan est par ailleurs toujours langue identitaire régionale dans le secteur d'Alguer (Alghero) en Sardaigne.

tels cadres évoluent constamment, mais on retiendra cette caractéristique fondamentale, ancienne, d'une grande diversité de langues et de variétés dialectales des langues au sein de la Méditerranée occidentale. Cette variété a pu être étudiée ici ou là sous des angles diachronique, typologique et comparatiste (Entwistle 1986, Bec 1970-1971), dialectologique (Telmon 1992, Grassi et al. 2005, Zamora Vicente 1996), statistique (Clanché 2002), politique et identitaire (Grandguillaume 1983), militant ou contre-militant (Aracil 1986, Salvador 1987, 1992) etc. On aurait le plus grand mal à en faire le tour tant les pistes sont nombreuses.

Le corpus est donc abondant et documenté, ce qui justifie précisément la nécessité de produire un essai de modélisation d'ensemble.

2. Le modèle français et le modèle espagnol, deux extrêmes

Le modèle français se trouve dans le droit fil de la construction française (langue, culture, état)⁵, qui consacre la primauté de l'unité et relègue la diversité dans l'ordre littéraire, émotionnel, identitaire. C'est certainement le modèle qui laisse le moins d'espace institutionnel aux langues régionales, si on le compare au système espagnol actuel ou encore au système italien, lesquels paraissent mieux intégrer la variation et la diversité des langues (au moins à première vue).

Peu de choses sont faites en France pour la promotion réelle et le développement des langues régionales, car en droit, dans l'esprit même de la construction nationale et républicaine française, la langue de la France est le français, ce qui détermine une attitude par définition « exceptionnelle », un écart, une parenthèse, dès qu'on se place sur le terrain des langues autochtones autres que le français⁶.

⁵ Informations nombreuses dans ce domaine. Voir pour la diachronie : Lodge (2003), et pour un recensement des langues régionales de France : Cerquiglini (2003).

⁶ Sans même parler des langues exogènes, issues de l'immigration, installées dans le paysage français et intéressant un nombre substantiel de locuteurs (par exemple : arabe ou berbère). On se souvient qu'il y a quelques années de nombreuses voix se sont élevées pour que ces langues et variétés de langues issues d'Afrique du nord prennent une place parmi les langues minoritaires de France, dès lors que de nombreux citoyens français les utilisent ordinairement. Idée principalement repoussée par le système sociolinguistique et institutionnel d'ensemble.

Le modèle espagnol, à cet égard, peut sembler beaucoup plus ouvert et favorable aux composantes ethnolinguistiques du territoire, mais en même temps le phénomène est plus récent et donc bien moins stable diachroniquement que dans le cas français. En synchronie, on pourrait néanmoins interpréter cette ouverture comme une conséquence « normale » du poids effectif des langues périphériques au sein du territoire espagnol.

La nouvelle constitution (décembre 1978), par la suite modifiée, créait juridiquement les « communautés autonomes » par le biais desquelles certaines langues régionales allaient acquérir des droits nouveaux (voir les annexes). En matière d'évolution des langues une trentaine d'années n'est pas une période suffisamment longue et l'on manque sans doute de recul, mais on peut commencer à tirer certains enseignements de ces années passées et voir apparaître quelques lignes prospectives.

L'un de ces enseignements est que l'action sur les langues, leur gestion et promotion directes, ne répondent qu'à une partie des problèmes posés par le contact des langues, en réalité on ne raisonne que sur la surface du phénomène. Il y a déjà quelque années, Lluís-Vicent Aracil, après avoir agi activement pour la promotion de la langue catalane (*lato sensu*), faisait un constat très dur, observant que les monolingues du catalan avaient disparu et que tous les « catalanoparlants »⁷ étaient désormais bilingues. Dans une métaphore de la marche il remarquait qu'après avoir eu les deux pieds en catalan... « nous avons maintenant un pied en catalan et un pied en castillan. Quand déciderons nous de lever le pied que nous avons en catalan et de poser les deux en castillan »⁸. C'est au fond cette idée de la vitrine symbolique où le catalan apparaîtrait comme promu, favorisé, tandis que sur le plan du marché économique et culturel d'ensemble le castillan poursuivrait son ascension continue et que l'hémorragie se perpétuerait en dépit de toutes les actions sur la langue⁹.

Le deuxième enseignement découle du premier : le système des

⁷ Adaptation du mot catalan.

⁸ Cité par Tudela (1986).

⁹ Et l'auteur d'évoquer alors une décision collective autoritaire de combattre la suprématie du castillan, « tout comme l'ont fait en Belgique les Flamands contre le français ».

autonomies ne serait démocratique qu'en apparence. Mettre à parité juridique une langue dominante et une langue dominée reviendrait en fait à favoriser paradoxalement la langue dominante, qui disqualifierait mieux sa concurrente car justement le cadre politique de l'autonomie soulignerait mieux les réseaux fonctionnels dans lesquels la langue en réhabilitation est en position de faiblesse, voire improductive. Mécaniquement (c'est vraisemblablement une bonne partie du raisonnement d'Aracil), une sortie possible des incohérences est alors d'abandonner la paire co-officielle et de décréter l'unicité de l'officialité dans la « langue propre ». Les développements de la catalanité active montrent bien qu'on va depuis dans ce sens.

Troisième enseignement, lié au précédent. Le système des autonomies ne prévoyait au départ que la satisfaction des revendications relatives aux langues propres (*lenguas propias*), typologiquement disjointes du castillan (à travers le triptyque fameux : catalan¹⁰, basque, galicien). Mais il est devenu mécaniquement encore, une pompe d'appel pour des formes d'affirmation identitaire très diverses, certaines anciennes et partiellement reconnues par la dialectologie romane (c'est le cas de l'asturien ou *bable*), d'autres bien plus étonnantes (cas de l'andalou¹¹). L'un des problèmes est qu'à terme toutes sortes de particularismes locaux pourraient s'engouffrer dans la brèche et l'on voit mal comment sortir de l'engrenage, tout le monde ayant à l'esprit (plus ou moins clairement) les risques politiques d'une atomisation du paysage sociolinguistique et identitaire¹².

¹⁰ Et indirectement occitan/gascon dans le cas du Val d'Aran.

¹¹ En 2006, la réforme du statut d'autonomie, sans accorder le statut de langue propre à l'andalou, introduit un article 210 (Reconnaissance et usage de la modalité linguistique andalouse) : *les moyens audiovisuels publics devront promouvoir la reconnaissance et l'usage de la spécificité linguistique andalouse dans ses différents parlars*.

¹² On pourra consulter les annexes A et B portant sur le domaine linguistique catalan, assurément l'une des vedettes du système espagnol. Le catalan se trouve engagé depuis plusieurs décennies dans une spirale de centralisation très forte (fondée sur le Nord et l'épicentre barcelonais) et des points de rupture se font sentir de plus en plus vigoureusement vers le sud (domaine valencien). On n'est pas très éloigné d'une forme de balkanisation qui revêt deux aspects fondamentaux : 1) un tri territorial inquiétant entre communes d'expression principalement castillane ou d'expression principalement valencienne, 2) une rupture symbolique et méthodologique en cours depuis plusieurs années entre valencien et catalan. Ce dernier aspect laisse pantois sur le plan dialectologique et géo-diachronique.

3. La protection des langues minoritaires dans la Méditerranée occidentale

Entre le modèle français centralisé et le modèle autonome pur tel que l'Espagne l'applique, la solution médiane pourrait être celle de la protection par l'État de ses langues minoritaires, du moins celles dont il veut bien reconnaître l'existence, même si la distance typologique n'est pas assurée par rapport au castillan¹³. Cette solution semble apparaître en Espagne pour l'asturien, l'aragonais (voir les textes des annexes C et D), ainsi que pour l'andalou, comme on vient de le voir. Tout est loin d'être clair, car il s'agit, dans les grandes lignes, d'un engagement de principe, rarement assorti de mesures vraiment concrètes et contraignantes pour l'État. En revanche, la mission des académies de langues régionales et associations diverses est grandement facilitée, du moins en bonne partie légitimée par les statuts de protection définis.

Dans le cas de l'Espagne on voit d'ailleurs assez bien que le système est susceptible d'évolutions. Par exemple, depuis plusieurs années sont encouragés les travaux de recensement et/ou de normalisation des langues régionales, ainsi que les expériences d'enseignement. Cela s'est vu dans le cas de l'aragonais, avec la création de « départements des langues d'Aragon » dans les établissements d'enseignement secondaire (2000), même si l'expérience est semble-t-il restée sans effet statutaire. Toujours en Aragon, une intéressante *loi sur le patrimoine culturel* a été adoptée en 1999, qui reconnaît l'existence de l'aragonais comme langue minoritaire et pourrait préparer une éventuelle classification comme « langue propre ». Le fait qu'une loi sur le patrimoine culturel ait vu le jour est très intéressant, car l'on commence à voir ainsi que des objets culturels, des langues et les systèmes identitaires qui leur sont associés peuvent devenir des objets du droit social collectif. Certains articles (donnés en annexe D) ne manquent pas de déranger nos habitudes pluriséculaires d'indifférence (voir par exemple la création d'un étonnant *devoir de conservation*), mais cela est probablement l'un des signes d'une évolution significative des mentalités en matière

¹³ C'est le nom le plus couramment attribué dans la péninsule à la variété véhiculaire nationale et d'État (*castellano, castellà* etc.), même si, d'un point de vue dialectologique historique, cette appellation rappelle plutôt l'origine régionale/de la langue en question.

d'écologie des langues dans l'ouest européen. Cette orientation aura certainement des conséquences.

3-1. Le cas italien

En matière de protection des langues, l'Italie semble s'être orientée vers la voie médiane qui vient d'être évoquée.

La constitution de 1948 (article 6) prévoyait déjà :

La République protège par des mesures convenables les minorités linguistiques¹⁴.

En 1999, la *loi de protection des minorités linguistiques historiques* énonce encore :

Article 1

1) La langue officielle de la République est l'italien.

2) La République, qui valorise le patrimoine linguistique et culturel de la langue italienne, fait la promotion et la valorisation des langues et des cultures protégées par la présente loi.

Article 2

En vertu de l'article 6 de la Constitution et en harmonie avec les principes généraux établis par les organisations européennes et internationales, la République protège la langue et la culture des populations albanaise, catalane, germanique, grecque, slovène et croate, et de celles qui parlent le français, le franco-provençal, le frioulan, le ladin, l'occitan et le sarde.

Les textes parviennent ainsi à maintenir le principe d'unicité de l'officialité (ce qui rapproche de la France) tout en prévoyant indirectement, mais statutairement, la différence ethnolinguistique (ce qui éloigne immédiatement de la France et rapprocherait plutôt de l'Espagne contemporaine).

Différents facteurs peuvent permettre d'expliquer cette ligne de conduite. Parmi les éléments d'intelligibilité bien connus il y a, fondamentalement, le fait que l'unification moderne de l'Italie est différente des unifications plus anciennes et autoritaires

¹⁴ Bien entendu, l'idée de mesures « convenables » peut être critiquée pour son flou, mais elle indique bien en même temps qu'on se situe dès le début du processus de protection des langues minoritaires sur une voie réellement intermédiaire.

(politiquement et linguistiquement) de la France et de l'Espagne. Jusqu'au bout l'ambiance est davantage à la tolérance ici que dans les deux autres cas, ce qui vaut autant pour le pôle minorant que pour le pôle minoré. Car on perd un peu trop souvent de vue que les tensions sociolinguistiques sont l'affaire d'une communication d'ensemble et donc de deux ou plusieurs partenaires. La responsabilité de l'État lorsqu'il est typologiquement centralisateur est certes grande, mais il faut souligner que les minorités réagissant intellectuellement sur des modes inculqués par l'histoire de l'État (normalisation, sélection, autoritarisme etc.), au bout du compte on ne sort guère et on n'a guère de chances de sortir d'une situation électrique. La France et l'Espagne le démontrent fort clairement, chacun des pays à sa manière. La France qui ne cède pas fondamentalement sur le dossier, comme l'Espagne, qui paraît avoir cédé, sans que pour autant le climat ne cesse d'être violent. C'est même souvent le contraire.

Remarquons que la langue centrale (que nous appelons italien) semble beaucoup mieux diffusée ici qu'en Espagne, et même qu'en France (les locuteurs usuels de l'italien représenteraient 87 %) ¹⁵, mais en même temps les données montrent que cette progression continue ne s'est pas faite par une déconstruction négative, un rejet du tissu régional. Au contraire, dans la ligne de ce qu'ont montré les dialectologues romanistes au cours du XX^e siècle, une forme de complémentarité naturelle, non remise en cause, s'établit entre langue centrale et langues régionales, voire locales (ce qui serait apparemment plus près de la réalité compte tenu de l'atomisation dialectale de l'Italie). Bien souvent, les langues d'Italie peuvent être vues comme des « dialectes » de l'italien (au bout d'une sorte de *continuum*) et il est avéré que l'on utilise de manière très significative le « dialecte » en famille et dans les contextes proprement locaux, tant au nord qu'au sud (plus de 50% dans les Abruzzes ou en Sardaigne, plus de 70% en Sicile ou en Vénétie), quand les langues sont bien romanes et plus ou moins proches de l'italien central ¹⁶.

Il existe en outre différentes langues n'appartenant pas au gallo-italien ou à l'italo-roman ¹⁷ (alémanique, francoprovençal, français, occitan ou provençal etc.). C'est ce qui explique qu'existe ici

¹⁵ Données tirées du site de Jacques Leclerc (<http://www.tlfq.ulaval.ca/AXL>).

¹⁶ Pour les langues d'Italie, on peut se référer à Grassi et al. (2005, 2006) et Telmon (1990, 1992).

¹⁷ En réalité, cela concerne également le sarde, langue romane particulière typologiquement, non réductible à l'italo-roman.

également un régime d'autonomie pour cinq régions du pays : Vallée d'Aoste, Trentin/Haut-Adige, Frioul/Vénétie Julienne, Sicile, Sardaigne. Mais cette autonomie n'a pas grand-chose à voir avec l'autonomie telle qu'elle fonctionne en Espagne. Comme le montrent quelques extraits relatifs au régime spécial de la Vallée d'Aoste qu'on peut prendre en exemple (annexe E), on retrouve assez nettement le critère de protection et de prééminence de l'italien (notamment sur le terrain judiciaire) même si le texte évoque une parité des langues (et non une co-officialité comme en Espagne). La propension patrimoniale de cette loi apparaît aussi dans la reconnaissance parallèle de la minorité *walser* (vallée du Lys)¹⁸, où seulement trois communes de langue germanique ont en théorie les mêmes droits que la partie « française » et francoprovençale.

Comme on peut le constater également, il existe des articles relatifs à la maîtrise attendue du français dans l'administration, aux applications scolaires etc., mais leur efficacité semble en réalité de plus en plus faible. Dans l'administration, en dépit d'une connaissance de principe du français par les fonctionnaires, l'italien occupe un espace toujours croissant ainsi que dans l'enseignement où son poids est de plus en plus fort à mesure qu'on progresse dans les cursus. Présent dans les premiers cycles (accompagné partiellement par le francoprovençal), le français voit ensuite sa part réelle se réduire dans l'enseignement secondaire et l'enseignement supérieur.

Le cas de la Vallée d'Aoste montre bien que les différentes mesures énoncées par les textes n'ont pas de caractère coercitif, ce que dénoncent bien entendu différents militants. Il y a là une grande différence avec le système espagnol, même si ce dernier (comme nous l'avons compris) amène son autre lot de problèmes. L'impression d'ensemble est alors celle d'un cadre protectionniste ou autonome mou, fortement théorique, et qui laisse l'essentiel des actions à mener aux communautés présumées être protégées par la loi. C'est pourquoi certains observateurs jugent très mal la gestion italienne des différences linguistiques, y voyant une sorte de laisser-faire généralisé sous couvert de protection patrimoniale. Les lois de protection des minorités déguiseraient en fait la progression sociolinguistique du seul italien. Ajoutons que les *règles en matière de protection des*

¹⁸ Une loi régionale d'août 1998 a précisé encore plus ces aspects relatifs à la protection patrimoniale.

minorités linguistiques historiques (décembre 1999), définissent une demande nécessaire de 15 % de électeurs d'une circonscription administrative (Conseil provincial¹⁹) pour qu'une minorité fasse l'objet de mesures légales de protection. A ce compte là, pour nous en tenir à l'exemple de la Vallée d'Aoste, le francoprovençal ne serait bientôt plus une minorité.

Un dernier vice du système doit être absolument souligné. Il n'est pas propre au système italien, mais bien plus général. C'est celui qui consiste à proposer pour des langues et identités régionales, voire très locales (cf. le cas du walser) menacées (et même très menacées²⁰), des solutions par une langue typologiquement apparentée mais qui n'est pas en réalité la langue régionale : français dans le cas du francoprovençal, allemand dans le cas du walser.

Généralement ces langues sont à l'évidence elles-mêmes des langues dominantes et destructrices dans leurs propres sphères d'action (en France²¹, en Suisse, en Allemagne etc.) et, par ailleurs, leur caractère fonctionnel et leur connaissance réelle sont localement très fragiles. Ainsi le français a-t-il été langue de superstructure et surtout langue de l'écrit, du culte etc. dans les Alpes italiennes de l'ouest (vallées vaudoises de langue d'oc, Vallée d'Aoste). Le définir comme une langue paritaire de l'italien (avec les réserves déjà faites) c'est d'une certaine manière condamner la véritable langue régionale vivante et encore parfois implantée dans les écarts ruraux²². Tout se passe au fond comme si les grandes langues d'expansions de l'Europe occidentale ne se reconnaîtraient qu'entre elles et procéderaient jusqu'au bout à des partages et à des re-découpages territoriaux avantageux, y compris dans leur gestion des faits régionaux. Pour revenir à la France, très coutumière de cette pratique, c'est ainsi que de jeunes Flamands du département du Nord sont aujourd'hui amenés

¹⁹ Les électeurs doivent être résidents des municipalités et c'est le Conseil provincial qui procède aux consultations. Le texte prévoit en alternative la demande d'au moins 1/3 des conseillers municipaux.

²⁰ Moins d'un millier de personnes peuvent être considérées comme des locuteurs du walser (estimation actuelle : 600 personnes), avec le flou que cela implique. Quand au francoprovençal il serait également en régression et l'on entrerait en ce moment dans une phase cruciale, où 15 % des Valdôtains seulement l'utiliseraient (surtout dans les zones écartées).

²¹ Est-il bien nécessaire de souligner ici que le français a pratiquement évacué de tous les réseaux fonctionnels et statutaires le francoprovençal de France ?

²² La plupart des études révèlent la progression très rapide de l'italien en zone urbaine et la relativement bonne tenue du francoprovençal dans les communes rurales.

à apprendre le néerlandais, de jeunes Alsaciens l'allemand etc²³, contribuant ainsi involontairement au recul du substrat dialectal originel de leur propre identité régionale, déjà mal en point.

L'Europe ne serait-elle pas, tout simplement, en train de redéfinir des macro-zones véhiculaires sous couvert d'intérêt pour ses minorités ?

3-2. Et le Maghreb ?

Bien qu'il existe dans ce dossier des langues régionales du Maghreb quelques points communs avec l'Europe, pour comprendre correctement ce qui se passe en Afrique du Nord il faut prendre un peu de distance par rapport à ce que nous avons écrit jusqu'à présent. Non seulement le Maghreb nous renvoie à des sections linguistiques fondamentalement non indo-européennes (bien que les langues indo-européennes, romanes en premier lieu, soient régulièrement constitutives de la donne), mais il nous renvoie aussi à des processus d'intégration politique, identitaire et langagière originaux, assez différents dans le détail des processus européens. Je ne parlerai pas directement ici des implications de la langue française au Maghreb (ce qui nous entraînerait trop loin²⁴), mais ceci nous amène par ailleurs à comprendre qu'il faut raisonner sur l'ensemble du paysage linguistique. L'idée que j'essaie pour ma part d'affiner depuis plusieurs années est celle d'un système tripolaire d'Afrique du Nord qui, structurellement, tend à bloquer pour l'instant le paysage linguistique du Maghreb et procure indirectement certaines forces à l'affirmation berbère (Manzano 1996)²⁵.

Au Maghreb, en termes de nombre de langues, la situation paraît nettement plus simple à première vue qu'elle ne l'est en Europe. Deux langues maghrébines se partagent l'espace, l'arabe (langue officielle du Maroc, de l'Algérie et de la Tunisie) et le berbère, présent surtout au Maroc et en Algérie²⁶, en phase d'extinction en Tunisie (Manzano

²³ En l'occurrence le flamand de France (*Fransch Vlaemsch* ou *Vlaemsch van Frankryk*), composante du flamand occidental (*West Vlaemsch*) et l'alsacien (alémanique).

²⁴ Et plus globalement des langues romanes et de leurs influences au Maghreb.

²⁵ Le lecteur qui voudra approfondir pourra se reporter à différents travaux fondateurs de Gilbert Grandguillaume (notamment 1983), ainsi qu'à différents travaux personnels dont certains sont rappelés dans la bibliographie.

²⁶ Par ailleurs présent dans différents pays du Sahel africain (Mali, Niger). En Tunisie, le berbère

1997a). C'est le résultat lointain de la superposition vraisemblable de deux vagues chamito-sémitiques dans la moitié septentrionale du continent africain. Du point de vue qui nous concerne, il n'est pas indifférent de constater que le berbère est la langue la plus anciennement implantée dans cette région²⁷. C'est donc, dans l'état de nos connaissances, la seule langue autochtone en profondeur. L'arabe, qui par ailleurs a seul le statut de langue d'État, ne s'est diffusé ici qu'à partir du VII^e siècle. Jusqu'au bout cet argument rejaillit sur les relations langagières et identitaires au Maghreb, fondant évidemment différentes revendications en Algérie et au Maroc, bien que les groupes de la région se rejoignent dans une cohésion supérieure de type islamique (à distinguer rigoureusement de l'arabo-islamisme).

L'une des choses qui sautent aux yeux, quand on compare avec l'Europe, c'est donc la simplicité apparente de la confrontation, l'apparence binaire et massive du système, très éloignée des atomisations dans les contacts entre les langues européennes. Il n'y aurait qu'une seule langue minoritaire ou minorisée, ces adjectifs ne convenant pas toujours comme on va le voir. En tout cas le berbère n'est guère une langue « régionale » puisqu'on le trouve à peu près partout, ou alors il faudrait inventer le terme plus approprié de langue « pluri-régionale »²⁸. Car il se présente sous forme d'îlots souvent séparés les uns des autres, avec aggravation du phénomène durant la seconde moitié du XX^e siècle.

L'éclatement géolinguistique du berbère est en effet, aujourd'hui du moins, l'un des facteurs les plus handicapants pour cette langue. Toutes les cartes actuelles de répartition montrent bien que la continuité territoriale de cette langue fait de plus en plus problème et qu'en même temps le berbère se rencontre en des endroits divers.

peut en effet être considéré dans une phase terminale.

²⁷ Sous sa forme ancienne du libyque, le rameau berbère est l'un des plus anciennement attestés dans cette région du monde. L'Histoire le montre coexistant chronologiquement avec le grec ancien et avec le latin, bien entendu avec d'autres langues aujourd'hui disparues, comme l'étrusque, le gaulois etc. Il a échangé avec certaines de ces langues, latin en premier lieu (Manzano 1999b), ce qui est encore une preuve de cette antiquité et fonde donc bien une situation parfaitement exceptionnelle dans le concert des langues actuelles de la Méditerranée.

²⁸ Les langues régionales d'Europe évoquées (France, Espagne, Italie) sont généralement compactes dans leurs répartitions et, pour cette raison même fortement exposées à une pression frontale des véhiculaires.

On constate en effet que les seules zones encore relativement compactes et continues se trouvent dans l'Atlas marocain (*tachelhit* = Sous, Haut-Atlas, Anti-Atlas ; *tamazight* = Moyen-Atlas) et en zone saharienne (*tamacheq* = pays Touareg). Le système actuel du berbère se présente donc comme une peau de léopard, avec net effilochement sous forme d'îlots plus ou moins vastes d'ouest en est.

Cet affaissement géodiachronique d'ensemble peut être appréhendé comme le résultat principal²⁹ de la propagation historique et géographique de l'arabe d'est en ouest. Au bout du compte résistent linguistiquement des secteurs montagnards (Atlas, Rif, Kabylie, Aurès), semi-désertiques ou désertiques (Sous, Adrar, Hoggar, sud tunisien) des îles (Djerba) ou, plus généralement, des secteurs relativement excentrés. Il convient d'ajouter que cette vision géographique et topographique du conservatisme berbère ne suffirait pas à elle seule à tout expliquer. Sans doute les zones berbères correspondent-elles à des espaces retirés, marginalisés. C'est l'idée ancienne des « espaces refuges » au Maghreb (montagnes et déserts). Mais ce sont surtout des zones d'économie rurale très ancienne (archéotypiques en quelque sorte) fondées sur des organisations sociales méditerranéennes³⁰, le plus souvent contradictoires des modèles d'aménagement venus par le bas, l'arabe, le français, ou par l'État. Ces organisations, telles qu'elles ont été lumineusement étudiées par des chercheurs comme Germaine Chantreaux (Kabylie), Germaine Tillion (Aurès) ou Jacques Berque (Atlas marocain)³¹, sont

²⁹ Bien entendu, le recul du berbère n'est pas uniquement dû à la propagation de l'arabe. Les langues romanes ont, au moins indirectement, joué un rôle (notamment le français). Il n'en est pas moins vrai que la confrontation diachronique *arabe* vs. *berbère* reste fondamentalement la matrice du système sociolinguistique d'Afrique du Nord.

³⁰ En ce sens que les contraintes domaniales méditerranéennes (climatiques et biogéographiques, hydrologiques, topographiques, agro-pédologiques etc.) ont longuement et fondamentalement pesé sur la genèse de l'ethnie berbère. Que celle-ci provienne d'une fragmentation protohistorique d'une famille anthropologique moyen-orientale (le segment « chamitique » ou « hamitique »), peu importe au fond, si ce n'est bien entendu que par ce biais une très lointaine typologie commune avec les langues sémitiques puisse apparaître. Le plus important pour comprendre ce qui s'est passé au Maghreb de longue date (voir plus loin si l'on étend la réflexion à l'Afrique romaine) et ce qui s'y passe encore en matière d'identités et de langues, est de bien comprendre que le monde berbère s'est développé depuis plusieurs millénaires dans un environnement eu-méditerranéen typique, générant ainsi un espace langagier et culturel soudé qui, probablement, reste à ce jour l'un des plus remarquables conservatoires d'états anciens en Méditerranée, toutes régions confondues.

³¹ Les travaux de ces différents observateurs s'égrènent durant la première moitié du XX^e siècle, durant une période où les sociétés berbères se présentaient comme fortement homogènes et relativement peu pénétrées par les koinés langagières ou culturelles. Cette vague importante de travaux se fondait elle-même sur une tradition d'ethnographie française et de linguistique du monde

de ce fait des organisations conservatrices et typées sociolinguistiquement, et pour ces raisons sans doute d'une très grande fragilité.

Le berbère est-il une langue minoritaire ? Oui et non. C'est oui véritablement en Tunisie, où il est passé au dessous du seuil légal de 1 % des locuteurs (Manzano 1999a). Non au Maroc où l'on estime généralement qu'entre 1/3 et la moitié de la population parle usuellement le berbère. Pas vraiment non plus en Algérie où un bon quart de la population (peut-être plus³²) l'utilise régulièrement.

Si le berbère est de moins en moins une langue minoritaire quand on se déplace géographiquement vers l'ouest, il reste partout une langue clairement minorisée et il a surtout été, dans l'Histoire moderne du Maghreb une langue déniée, objet d'un véritable tabou jusqu'à très récemment, ce qui traduit une tension identitaire d'envergure. Cette tension identitaire a des origines à la fois anciennes (autochtonie berbère vs. expansion arabe) et modernes (passé colonial). Elle est également liée à la conception même des unités nationales au Maghreb, sub-continent qui hésite structurellement et de longue date entre :

1) Une définition arabo-islamique de la nation (la *'umma*), qui peut pactiser (toujours provisoirement, même si le provisoire est long) avec la différence, mais ne saurait l'admettre et la prévoir de manière consubstantielle dans son organisation. Les solutions extrêmes, qui illustrent cette logique, n'ont pas manqué de se manifester depuis le XIX^e siècle. Elles rejoignent généralement le panarabisme et prévoient ou programment à terme la disparition des variations dialectales de l'arabe et surtout l'élimination du berbère.

2) Une définition européenne, romane et française. Celle-ci permet au moins que l'on admette l'existence de la différence, qu'on veuille

berbère qui s'était développée de la seconde moitié du XIX^e au tout début du XX^e siècles, des études célèbres de Hanoteau et Letourneux (notamment 1873) à celles de Charles de Foucauld par exemple.

³² Les enquêtes et sondages en de telles matières sont déjà sujets à caution en Europe. Mais c'est plus grave encore au Maghreb où le principe même d'enquête sur les pratiques langagières et les identités provoque immédiatement une série de réflexes sociaux de rejet et de fermeture. Bien entendu, de vraies enquêtes d'envergure sur les composantes identitaires sont pratiquement exclues de fait. On doit donc se résoudre le plus souvent à des estimations qui, par ailleurs, ne suffisent pas à remplacer une connaissance pratique du terrain.

bien la dire, même si la gamme des attitudes est assez large (faut-il reconnaître cette différence, lui permettre de s'exprimer ? etc.), comme on l'a vu pour l'Europe. Dans le cas précis du Maghreb, colonisé par la France, il faut souligner que la France a fait en faveur du berbère ce qu'elle n'a guère fait pour ses propres langues régionales, créant par anticipation une forme de droit d'antériorité (patrimoniale et fonctionnelle) de la langue et de la culture berbères. On comprend en fait que pour les politiques et une bonne part de l'intelligentsia de l'époque (fin XIX^e, début XX^e), un projet plus ou moins nettement exprimé consistait à désolidariser à terme les composantes berbérophones d'un bloc maghrébin globalement hostile à la présence française (Manzano 2007).

Pour finir de caractériser le berbère, un autre phénomène révèle bien le profond décalage par rapport à l'Europe. Des monolingues régionaux pouvaient exister jusqu'au milieu du XX^e siècle en France, en Espagne et en Italie, mais en Europe la page semble bien tournée désormais ou sur le point de l'être définitivement (même si l'on peut encore discuter sur certains cas). Au Maghreb en revanche ce statut est toujours possible dans différentes régions berbérophones. Non seulement il concerne de très vieilles personnes, mais il peut également impliquer des locuteurs d'âges divers dans des régions encore très isolées, des secteurs ruraux sédentaires voire semi-nomades ou nomades où la scolarisation et l'intrusion de l'État ont réellement très peu d'effets, et où la langue de communication ordinaire reste fondamentalement le berbère. Bref une situation écolinguistique et sociale que l'Europe ne connaît probablement plus depuis plusieurs décennies. À défaut de vrai monolinguisme on a souvent affaire à un bilinguisme très déséquilibré mais en faveur de la langue locale, non en faveur de l'arabe. Sur ce point aussi le Maghreb n'est pas l'Europe³³.

Si l'on envisage la question de l'écrit, il est certain que la langue berbère est principalement orale. Si on l'a écrite c'est autrefois, ou encore à époque moderne, mais semble-t-il sur des marges (pays

³³ On peut donc considérer que dans les zones où il est activement parlé (Algérie et plus clairement encore Maroc), le berbère n'est pas à proprement parler une langue minoritaire, mais bien une langue « majoritaire », en termes de masse d'usages quotidiens. Ce cas montre bien qu'il faut parfois distinguer langue minoritaire et langue minorisée.

Touareg assurément³⁴). Henri Duveyrier (*Bulletin de la société de géographie*, 1863)³⁵ mentionnait en ces termes le maintien d'une écriture berbère :

Tandis que dans tous les États barbaresques une femme sachant lire et écrire est une exception très rare, presque toutes les femmes touâreg lisent et écrivent le berbère, et quelques unes lisent et écrivent aussi l'arabe.

Le système *tifinagh* (ou encore *tifinaghe*)³⁶, dont il s'agit ici, a été réactivé durant les quarante dernières années, et l'on peut en somme le qualifier de néo-tifinagh (Aghali-Zakara & Drouin, 1997 : 109). Il prendrait vraisemblablement sa source dans le libyque ancien et, plus loin en amont, dans le système orthographique du punique (phénicien)³⁷. L'utilisation de l'orthographe berbère était encore des plus problématiques dans les années 1970 à 1980, voire objet de franche répression. La situation s'est progressivement améliorée depuis, mais le passif est lourd et du temps sera sans doute nécessaire pour parvenir à une situation véritablement pacifiée. Il est à remarquer que l'écriture berbère a été choisie institutionnellement au Maroc, où l'IRCAM tend à la diffuser depuis sa création³⁸.

On peut constater que le tifinagh bénéficie d'engouement sur la toile (WEB), à travers différents sites notamment conduits par des Berbères expatriés en Europe. On rejoint ainsi une forme de « vénérabilité » et de mystère, cette profondeur méditerranéenne déjà évoquée, d'autant plus logique que la langue se perd dans la nuit des temps. Mais en même temps, un schisme méthodologique est en cours

³⁴ Sur ce thème, on pourra consulter Aghali-Zakara Mohamed & Drouin, Jeannine « Ecritures libyco-berbères. Vingt-cinq siècles d'histoire », dans Zali, Anne & Berthier, Annie (1997 : 99-111).

³⁵ Henri Duveyrier avait pu explorer le Hoggar pendant deux années, à partir de 1859. Il en tira l'ouvrage *Exploration du Sahara, les Touâreg du Nord* (Challamel, 1864).

³⁶ Les deux graphies coexistent. L'écriture à la française (avec -e muet terminal) est la plus récente des deux. Elle touche d'autres mots-clés de l'ethnolinguistique berbère, comme *amazigh/amazighe* par exemple.

³⁷ On doit ici rappeler que l'alphabet phénicien (développé à la fin du II^e millénaire av. J.-C.) fut une véritable plaque tournante, ancêtre des graphies de la plupart des langues chamito-sémitiques mais aussi indo-européennes.

³⁸ Institut Royal de la Culture Amazighe, mis en place au début des années 2000. Sous l'égide de cet institut notamment (dirigé par le sociolinguiste Ahmed Boukous), le berbère est enseigné dans un certain nombre d'écoles primaires du Maroc, mais pour l'instant l'extension aux collèges et lycées ne s'est pas faite.

depuis plusieurs années. On le résumera rapidement en disant que deux hypothèses se présentent assez nettement.

Dans la première, le système tfinagh serait étendu et deviendrait la norme fondamentale du berbère. Cela serait à première vue très cohérent, car on affirmerait ainsi la différence par rapport à l'arabe mais aussi aux langues européennes. Mais le point faible pointe aussitôt, car on risquerait bien évidemment un enfermement dans un système internationalement illisible. Seuls des autochtones, de plus formés par des apprentissages préalables (ce qui suppose en outre une réelle réussite de la scolarisation), domineraient un tel système, mais en quelque sorte en circuit fermé. Il est probable à ce sujet qu'au delà des motivations légitimes de curiosité, peu d'Européens (et même de Maghrébins) seraient amenés à l'acquiescer³⁹.

C'est pourquoi existe depuis assez longtemps une deuxième hypothèse, consistant à entretenir une ou des graphies de type international (sur un mode majoritairement européen et roman) avec les aménagements nécessaires pour le berbère. Cette voie, qui couperait sans doute de l'« authenticité » ou de la « spécificité » berbère, présenterait bien entendu l'avantage de rendre le berbère lisible directement et d'apprentissage plus facile dans l'espace méditerranéen et européen qui touche directement au Maghreb.

Dans les faits, on voit bien pour finir que les deux voies ou hypothèses sont amenées à coexister et s'interpénétrer, notamment au sein des sites internet déjà évoqués. Cette indécision globale révèle bien l'interférence d'enjeux identitaires, historiques mais aussi fonctionnels et économiques.

Ceci nous amène directement à la question de la réhabilitation du berbère. Est-il d'ailleurs normal de parler de réhabilitation d'une langue qui concerne des millions de personnes dans leur vie quotidienne aujourd'hui encore⁴⁰ ? Disons simplement qu'après une phase de mutisme ou de négationnisme (qui a duré jusqu'aux années

³⁹ Le critère économique reste essentiel, qui peut bousculer les réticences à aborder des systèmes orthographiques exotiques, ce que l'on voit bien pour le chinois aujourd'hui. Mais on comprend bien que ce critère ne puisse jouer pour le berbère, dans l'état actuel des relations internationales politiques et économiques.

⁴⁰ Se confirme ici une différence saisissante avec les situations européennes. Là où il est parlé, le berbère est vraiment parlé dans la plupart des situations et ce sont l'arabe (voire le français) qui ont le plus grand mal à se faire une place dans le réseau fonctionnel.

1980-90), des mouvements très tendus (« Printemps berbère » au Maroc, diverses insurrections kabyles en Algérie dans les années 90⁴¹ etc.) ont eu pour conséquences quelques actes (d'ailleurs très critiqués) concernant la reconnaissance de la langue berbère dans les deux pays principalement concernés. Un *Haut Commissariat à l'Amazighité* existe en Algérie depuis quelques années, ainsi qu'un *Institut Royal de la Culture Amazighe* (IRCAM) au Maroc, déjà signalé.

L'enseignement du berbère est désormais possible dans le primaire au Maroc, mais les expériences sont en fait trop récentes pour qu'on puisse en tirer un premier bilan distancié. Ce que l'on peut constater en revanche, c'est la relative incohérence des canaux d'affirmation et de reconnaissance du berbère qui se mettent en place en ce moment à travers le Maghreb, du fait des deux hypothèses graphiques plus haut évoquées.

Ainsi, les autorités marocaines encouragent-elles explicitement l'utilisation du tfinagh, qui devient la norme de valorisation, d'enseignement, et tendent-elles aussi vers un berbère (marocain) standard (même si les variations régionales sont prévues dans le système⁴²). Les premiers manuels publiés l'ont été intégralement dans ce code écrit, dans la logique définie par l'IRCAM. En revanche, en Algérie on s'orienterait plutôt vers des solutions plus empiriques reprenant fréquemment les caractères européens ou « internationaux », mais en même temps on refuse l'idée de norme centrale, puisque chaque grand groupe géographique et linguistique (*Kabylie*, *Aurès* etc.) tend en ce moment à sa propre définition du parler et de l'écrit directs.

C'est dire que l'Algérie et le Maroc (à travers les autorités mais aussi les groupes de pression favorables au berbère) s'orientent vers des solutions bien différentes pour ne pas dire à terme opposées⁴³.

⁴¹ Pour réagir en 1991 notamment à la loi portant généralisation de la langue arabe dans l'administration.

⁴² La plupart des instructions et manuels prévoient les trois zones dialectales du berbère reconnues au Maroc : nord rifain, centre (Moyen-Atlas, Atlas septentrional) et sud (Atlas, Anti-Atlas).

⁴³ Comme on l'a vu, l'une des causes de cette divergence réside dans l'idée que l'orthographe spécifique du berbère handicaperait la diffusion (internationale notamment) de cette langue, tandis que le recours aux alphabets de type roman faciliteraient lisibilité, diffusion et apprentissage. Pour

Comme on l'a déjà remarqué, le recul du berbère se fait par fractionnement, les unités s'isolant de plus en plus les unes des autres. Ceci est particulièrement net en Algérie et peut-être la volonté d'aligner la réhabilitation du berbère sur cette structure d'évolution géolinguistique est-elle en partie cohérente. Le risque est évidemment de creuser les écarts par une forme de valorisation des différences qui d'une certaine manière pourrait précipiter la division et l'affaiblissement du berbère par rapport aux deux langues théoriquement « homogènes » disponibles en Afrique du nord : arabe et français⁴⁴.

La solution marocaine a le mérite de poser le berbère comme une langue également homogène, et le modèle peut donc tendre à minimiser les différences dialectales. D'autre part, en attribuant au berbère son propre système graphique, la solution amplifie l'originalité typologique et symbolique de cette langue dans la Méditerranée occidentale ; mais en même temps elle risque fort d'enfermer le berbère dans une illisibilité fondamentale⁴⁵ qui localisera forcément cette langue comme une langue ethnique, une langue conservatoire opaque, phénomène ennuyeux dans l'optique d'une ouverture économique et culturelle à même de favoriser certaines formes de reprise de cette langue. Sur ce plan, malgré ses nombreux locuteurs, le berbère est frappé des mêmes faiblesses que plusieurs langues minorisées de la Romania déjà évoquées. Beaucoup d'acteurs du marché comprendraient mal la nécessité pour faire des affaires de passer par le berbère, l'occitan, le catalan, le galicien etc. quand on dispose pour cela de langues véhiculaires rodées comme l'arabe, le français, l'espagnol, le portugais, voire l'anglais. Il y a probablement des degrés entre ces langues minorisées, mais le mal est

certaines autres probablement, même l'orthographe arabe pourrait peut-être servir le berbère, mais on toucherait alors à des problématiques identitaires très complexes.

⁴⁴ On se réfère ici à une représentation de l'homogénéité. Le français est vu au Maghreb comme une langue d'une pièce, fortement policée, normalisée. On voit généralement mieux les défauts de l'arabe (opposition classique/dialectaux etc.) tout en reconnaissant une forme d'unité générale de cette langue face au berbère. Quant à ce dernier, il est vu la plupart du temps comme un ensemble de parlures marginales, paysannes, montagnardes etc. c'est-à-dire à peu près l'inverse de ce que l'on appelle communément langue.

⁴⁵ On a déjà souligné qu'au plan international les chances de rayonnement et de diffusion seront alors très faibles. Ajoutons que l'immense majorité des nationaux (déjà réticents pour d'autres raisons) n'apprendraient vraisemblablement pas à lire le berbère et pourraient ainsi être confortés dans l'idée perverse que le berbère n'est plus une langue de son époque, qu'il peut (ou doit) donc disparaître.

fondamentalement le même.

Le chemin est donc très difficile à trouver, parsemé d'embûches. Aucun des États concernés ne semble prêt dans l'immédiat à admettre l'idée que le berbère puisse recevoir un quelconque caractère vraiment officiel (à l'image de ce qu'on trouve en Espagne, voisin et modèle courant vers lequel se tournent de plus en plus souvent les regards maghrébins). On a vu plus haut le pourquoi de la négation automatique de cette possibilité au Maghreb. Mais en même temps, le Maghreb politique et économique se rapproche constamment de l'Europe, tandis que l'Europe elle-même finit concrètement au Maghreb⁴⁶. Si différents pays et/ou groupes sociaux du Maghreb forment depuis quelques années des projets d'intégration à la communauté européenne, nul doute que la question des identités reviendra à l'occasion d'un rapprochement de ce type, amenant les attitudes en matière de gestion du contact des langues à se redéfinir mécaniquement dans le sens probable d'une qualification au moins patrimoniale plus profonde et franche du berbère (version minimaliste), voire dans le sens de l'accession du berbère à une forme (même atténuée) de reconnaissance ou de co-officialité (version maximaliste). Le voisinage direct de l'Espagne et de l'Italie pourrait peser dans ce sens et remettre en cause l'immobilisme du Maghreb en cette matière. Il faut en effet comprendre que l'interaction entre les différentes rives de la Méditerranée est forte, économiquement, culturellement et linguistiquement⁴⁷.

Que cette interaction n'ait pas d'incidence immédiate est parfaitement compréhensible. Qu'il n'y en ait pas du tout sur le long terme est en revanche hautement improbable.

⁴⁶ La remarque est symbolique mais aussi concrète. Rappelons que l'Espagne est présente sur la rive marocaine dans ses territoires de Melilla et de Ceuta (et d'autres, plus menus encore), qui mettent l'espagnol directement en contact avec l'arabe et le berbère, depuis très longtemps. Les Iles Canaries, d'autre part, sont primitivement des îles berbères (guanche), bien que la langue berbère n'y soit plus qu'un lointain souvenir. Comme on le sait, l'actualité migratoire met souvent ces territoires de souveraineté espagnole à la une de l'information.

⁴⁷ Le rapprochement, sous des formes diverses, a commencé. Des structures catalanes jouent un rôle de premier plan dans ce domaine. Dans cet ordre d'idée, on peut noter qu'en 2006 était signée une convention entre le CMA (*Congrès mondial amazigh*, basé en France) et l'Unesco Catalunya (Unescocat), qui mettait en avant les « nations amazighe et catalane ». On voit ainsi cultures et langues régionales du bassin méditerranéen se rapprocher (non sans manipulations collectives) pour poser en ligne de mire des concepts refusés par l'État central espagnol ou les États maghrébins (nations amazighe, nation catalane).

4. L'avenir des langues régionales, minoritaires ou minorisées de la Méditerranée occidentale

4-1. Quels risques de régression ?

Une idée répandue dans les opinions publiques est que plusieurs de ces langues feraient déjà partie du passé, qu'elles seraient d'autant plus vouées à la régression qu'elles n'ont plus de place réelle dans la vie moderne. C'est leur inadaptation fondamentale au changement sociétal qui serait source de leur perte. Qu'on s'en réjouisse ou qu'on s'en plaigne, le processus est donné, vu, comme une mécanique depuis longtemps enclenchée et désormais irréversible. Et comme par ailleurs la plupart des régions de la Méditerranée occidentale sont vues comme archaïques économiquement (voir plus haut : 1. *Paradoxes méditerranéens*), nul doute que dans l'esprit du plus grand nombre la remise à niveau économique et culturelle de ces régions fragiles supposera de passer par certaines langues véhiculaires, perçues quant à elles comme de vraies langues de la modernité.

Pour ce qui concerne la Méditerranée occidentale il y a, on l'aura compris, plusieurs façons de répondre à la première partie de la question. Si l'on raisonne sur des données numériques acceptables, la réponse la plus évidente est que la régression se constate à peu près partout sur le terrain⁴⁸. Une régression du nombre des locuteurs, à première vue plus accentuée là où l'autoritarisme est le plus fort. Et l'on pense tout particulièrement à la France où les graphes d'abaissement sont peu discutables (Clanché 2002), bien que certaines surprises soient pourtant possibles dans les vitesses de régression (Manzano 2006a).

Cette régression s'accompagne dans la plupart des cas d'une restriction croissante des fonctions et des zones socioculturelles dans lesquelles on peut parler les langues en question : il s'agit le plus

⁴⁸ Le problème général des langues régionales de l'ouest européen, dans l'ensemble, est celui d'un recul des langues minoritaires originelles sur le terrain (les « badumes » de Jean Le Dù). Bien que ce phénomène tende à être masqué par des processus de reprise statutaire ou symbolique (médias, scolarisation etc.) qui se situent à un niveau différent. Malheureusement, en général, ces phénomènes partiels de reprise ne parviennent pas à compenser la perte de locuteurs réels et originels sur le terrain, la langue continue globalement de reculer. De tels problèmes se perçoivent très bien pour des langues comme l'occitan ou le breton, un peu moins pour des langues vedettes comme le catalan, mais en réalité le problème s'y constate aussi avec la régression des variétés locales (natives) sur le terrain, notamment sous la pression de la norme catalane « centrale ».

souvent d'une restriction de la langue utilisée dans des réseaux sociaux de plus en plus réduits et confidentiels, en même temps qu'elle est utilisée par des locuteurs de plus en plus âgés. Et c'est ce qui explique d'ailleurs que certaines langues résistent mieux que d'autres, dès lors que les sociétés qui les utilisent restent plus ancrées dans des modes de faire-valoir traditionnel. Du coup elles peuvent alors présenter une pyramide plus équilibrée de locuteurs et des secteurs d'utilisation bien plus diversifiés. J'ai souligné ce phénomène à plusieurs reprises ces dernières années en m'appuyant sur deux domaines géographiques bien tranchés que j'avais eu l'occasion d'étudier : le domaine gallésant et le domaine occitan. Le gallo (malgré les militants) est vu comme typologiquement proche du français, davantage un patois local/régional qu'une langue⁴⁹. L'occitan, au contraire, colle plus facilement au *status* de langue, ne serait-ce que pour des raisons diachroniques et ethnolinguistiques. Néanmoins, sous nos yeux, c'est la première de ces langues qui semble s'en tirer le mieux, bien connue et pratiquée dans des campagnes de Haute-Bretagne restées fort conservatrices quant au faire-valoir. Dans de telles conditions le gallo a toujours sa place, fonctionnellement, là où le français ne pourrait que difficilement le remplacer.

C'est en bonne partie pour le même genre de raisons que l'occitan recule relativement plus vite et profondément que le gallo en Haute-Bretagne. Pour prendre un exemple assez concret, les mutations successives de l'espace languedocien au cours du XX^e siècle, ont donné peu de chances d'adaptation à une langue primitivement liée à des sociétés méditerranéennes (ou sub-méditerranéennes dans les régions semi-montagnardes) fondées sur une association de la polyculture traditionnelle (blé, vigne, olivier, arboriculture) et de l'élevage. En quelques décennies, ce legs méditerranéen a été mis à mal (pour ne pas dire pulvérisé) par le développement de la monoculture viticole, et l'occitan (en dépit des efforts des uns et des autres) s'est constamment affaibli au profit du français régional, mieux adapté aux changements en cours et à l'orientation vers une économie viticole de masse⁵⁰.

⁴⁹ Ceux qui connaissent ce domaine géographique savent aussi que la plupart des locuteurs, aujourd'hui encore, qualifient le gallo de « français écorché », « français déformé » etc.

⁵⁰ Il faut bien comprendre que dans un tel processus, des pans entiers (notamment pour le lexique usuel, les locutions, les lexies [recettes, proverbes, dictons etc.], le discours) deviennent sans objet

Sorti bien fragilisé de cette aventure du XX^e siècle, on peut légitimement craindre le pire pour l'occitan du Languedoc au XXI^e siècle. Après les nombreuses difficultés issues de la monoculture languedocienne et leurs conséquences (exodes ruraux, perte/appauvrissement de la langue natale), d'autres se manifestent, qui ne seront pas sans incidences. D'une part la viticulture elle-même recule sous nos yeux (campagnes européennes d'arrachages) mais on ne retournera plus pour autant aux sociétés languedociennes méditerranéennes d'autrefois, en osmose éprouvée avec l'occitan. D'autre part, cette région, pour les mêmes raisons d'ensemble, tantôt se dépeuple encore, tantôt se stabilise ou progresse démographiquement, mais en incorporant des populations d'origine extérieure, soit venues du sud (Maghreb notamment), soit venues du nord de l'Europe (Manzano 2005a, 2006b). De moins en moins l'occitan n'aura de chances d'être langue de réunion, de socialisation éventuelle dans ce nouveau processus d'immigration, ce qui très normalement ne fera qu'accroître toujours plus le poids de la langue française dans la mécanique⁵¹.

Les actions collectives militantes peuvent-elles influencer sur de telles mécaniques ? En partie oui, mais encore une fois de manière plutôt superficielle et symbolique. Evidemment, de telles actions renforcent la visibilité de la langue en question dans l'espace public, mais le risque est connu, c'est celui de masquer (le plus souvent involontairement) le recul effectif de cette langue. En outre, la production militante de normes directrices (compréhensible sur un plan politique et pédagogique) touche en réalité à l'originalité même des langues sur le terrain que l'on fragilise ainsi encore plus, souvent sans en avoir conscience. Le plus tragique est ainsi qu'un système très ancien de dialecte peut disparaître en quelques décennies sous l'action d'une langue dite « nationale » autant que sous l'action d'une langue régionale normativisée (en l'occurrence le catalan dans le cas traité

immédiat, là où convient de mieux en mieux le français. Mais au-delà de cet aspect « fonctionnel », c'est de proche en proche le besoin et l'envie même de la langue d'origine qui disparaissent, d'autant que durant de telles phases la modernité implique presque automatiquement le recours au français. L'occitan régresse alors par le biais d'une sorte de maladie sociale, une propension au mutisme, qui touchent les individus et de proche en proche l'ensemble du groupe.

⁵¹ Il n'y a pas si longtemps, les immigrés pouvaient encore s'intégrer (au moins partiellement) par le biais de l'occitan. Tel fut le cas de différents immigrés espagnols à l'issue de la guerre civile espagnole, notamment (mais pas seulement) du fait d'origines linguistiques proches (par exemple, cas des Catalans, Valenciens etc.).

par Perea 2007). Les communautés militantes devraient s'inquiéter de tels phénomènes, les comprendre, les intégrer, en vue d'actions plus adaptées sur les langues qu'elles entendent protéger.

4. 2. Fonctions des langues régionales.

En fin de compte, la résistance potentielle d'une langue est principalement reliée aux fonctions que peut assurer cette langue. Or nous voyons que dans la plupart des cas, les langues de grande diffusion (qui sont aussi des langues d'État dans la macro-région observée), parviennent à supplanter la plupart des langues minorisées dans une série de fonctions qui leur sont en principe réservées : sociabilité, mécanismes ethniques locaux, agro-pastoralité, famille. Quand la langue dominante s'impose dans de tels secteurs identitaires, la situation devient inquiétante car la langue minorisée est véritablement menacée. Au bout du compte la langue locale devient de plus en plus une langue qu'on « cite », qui traverse par fragments la langue dominante installée, elle ne sert donc plus directement mais le recours (au moins partiel) à cette langue reste un temps une forme de réflexe ou de signum identitaire.

C'est donc en dernier lieu la résistance sociétale qui arme ou désarme la résistance linguistique. Le développement précédent relatif au gallo et à l'occitan le montre bien.

Enfin, une même langue peut relever des deux pôles (résistance/disparition). Cela se voit très bien dans le cas du berbère de Tunisie, qui n'a plus que quelques fonctions terminales (il est en cela nettement comparable à l'occitan, à Djerba par exemple) et court le risque de disparaître à court terme. En revanche, le même berbère (mais sous d'autres formes géolinguistiques) est encore très fonctionnel dans l'Atlas marocain dès lors qu'il convient parfaitement à l'organisation et au fonctionnement de groupes ruraux peu touchés par les influences venues du « centre » marocain.

4. 3. Relativité des minorités.

Dans le même ordre d'idée, il est intéressant d'observer que la notion de minorité est devenue parfaitement relative, pas toujours déterminée par des conditions objectives, mais plutôt par les opportunités de découpage géographique et politique ou les

mécaniques sociolinguistiques des États concernés. Ainsi le mirandais, dialecte « asturo-léonais » est-il officialisé depuis 1999 comme langue d'une minorité (pourtant réduite géographiquement et numériquement⁵²) au Portugal, alors que l'asturien d'Espagne, incontestablement parlé par beaucoup plus de locuteurs et sur un territoire plus important⁵³, ne parvient pas à une reconnaissance pleine⁵⁴. Il est vrai que le Portugal s'apparente plus à un pays monolingue (contrairement à l'Espagne) et qu'il y a peu de minorités à protéger. Cela compte sans doute.

On peut de même évoquer les cas intéressants du Val d'Aran en Espagne, de la Vallée d'Aoste et de la communauté « walser » en Italie.

Dans le premier cas, la haute vallée de la Garonne, de langue gasconne, voit à nouveau quelques milliers de personnes disposer d'une langue co-officielle dans le cadre de l'autonomie de la Généralité de Catalogne (dont fait partie le Val d'Aran). Ainsi, le gascon qui ne dispose d'aucun statut particulier en France⁵⁵, jouit d'une reconnaissance et d'une fonctionnalité institutionnelle réelle⁵⁶. On ne peut trop dire ce qu'il en résultera dans l'avenir. Mais une chose est certaine : l'aranais n'est en somme plus (ou n'est plus vu comme) un « dialecte » du gascon, une somme de patois locaux. Il remonte immédiatement dans la hiérarchie des représentations.

Dans le deuxième cas, la Vallée d'Aoste jouit d'un statut d'autonomie qui permet au francoprovençal et au français (avec les réserves plus haut évoquées) de bénéficier de certaines protections⁵⁷. Mais le cas est encore plus frappant pour la langue germanique de la

⁵² Quelques villages au Nord-est du Portugal (moins de 500 km²) et quelques milliers de locuteurs, semble-t-il en régression.

⁵³ La principauté des Asturies s'étend sur 10.000 km² et compte une population de 600.000 habitants.

⁵⁴ L'annexe C (article 4) révèle bien qu'en fournissant un cadre politique autonome aux régions concernées, on ne reconnaît pas pour autant la langue comme moyen d'expression officiel. Les thèmes affichés de la promotion, de la protection ou du volontariat (relativement à l'asturien ou bable) confirment bien ce point de vue.

⁵⁵ En France, le gascon se trouve même secondarisé par rapport à la norme principale de l'occitan. Le cas du Val d'Aran n'en est donc que plus lumineux.

⁵⁶ Voir directement le site du Conseil Général : <http://www.conselharan.org>. Cette officialisation a notamment donné lieu à une loi adoptée par le parlement de Catalogne le 22 septembre 2010.

⁵⁷ On remarquera toutefois un flou certain dans les articles (annexe E, articles 39 et 40), lesquels oscillent entre une forme de radicalité et une marge d'adaptabilité manifeste dans les rapports statutaires entre français et langue de la république.

communauté Walser⁵⁸, toujours utilisée dans la haute vallée du Lys, elle-même constitutive de l'autonomie valdotaine. Quelques centaines de personnes parlent encore cette langue dans ce qui constitue *de facto* l'emplacement le plus méridional des langues germaniques en Europe. Le statut sociolinguistique de ce groupe n'est pas très affirmé et ces parlers subsistent dans un contexte plurilingue où interviennent aussi le francoprovençal et le piémontais, sans compter les langues normativisées comme l'italien, l'allemand ou le français. On distingue localement des parlers qualifiés de *tisch* et de *toïtschu*, plus ou moins éloignés de l'allemand standard. Mais ici comme ailleurs, dans une situation que beaucoup considèrent comme presque terminale, les sursauts culturels peuvent favoriser presque automatiquement le déport du système vers l'allemand standard, à même de favoriser une fixation (même insatisfaisante) de la langue locale (scolarisation, affichage etc.)⁵⁹. On peut en outre constater que dans un tel contexte terminal les langues minoritaires en arrivent à s'encaster les unes dans les autres. Par exemple, à Gaby, l'une des quatre communes⁶⁰ :

*Le groupe francoprovençal utilise le « patois », le piémontais, le français ou l'italien. Le toïtschu est parlé par ceux qui, tout en vivant à Gaby, sont originaires du groupe allemand. Dans l'interaction entre les deux groupes, la langue utilisée est le patois dans la variété de Gaby, surtout chez les personnes d'âge moyen ou chez les anciens, tandis que chez les plus jeunes c'est le piémontais et l'italien que l'on utilise.*⁶¹

Assurément le walser ne représente plus grand chose en terme de poids réel, nettement moins que l'aranais ou même que le mirandais évoqués un peu plus haut⁶². Sur des bases seulement statistiques ou sociolinguistiques, un tel groupe court le risque d'être emporté à terme

⁵⁸ Cette communauté prolonge en territoire politique italien un ensemble principalement réparti de la Suisse (Valais, d'où l'appellation germanique Walser) à l'Autriche. Les locuteurs sont répartis dans quatre communes d'une vallée elle-même peu accessible.

⁵⁹ Voir annexe E, article 40-bis.

⁶⁰ Les communes sont, en remontant la vallée : Issime, Gaby, Gressoney Saint-Jean, Gressoney La Trinité. Comme on le comprend dans ces lignes, et sans pouvoir entrer dans le détail, l'aire francoprovençale fait une incursion dans la commune de Gaby. En réalité, c'est bien d'un petit segment de la frontière entre langues romanes et langues germaniques qu'il s'agit ici.

⁶¹ Traduction libre d'après Musso (2007 : 29, note 12).

⁶² Des enquêtes assez récentes (notamment celles de la Fondation Chanoux, Aoste) permettent de présumer qu'entre 15 et 20% des habitants des communes des deux Gressoney et d'Issime auraient le germanique local comme langue maternelle. Si l'on rapporte les données à l'ensemble de la communauté autonome de la Vallée d'Aoste, on tombe à 0,03%. Ceci ne peut que souligner la fragilité probable du groupe mais en même temps son caractère exceptionnel sur le plan identitaire.

bref, même dans l'hypothèse où l'allemand fixerait pour quelques temps encore un sursaut germanique dans ce secteur (dans ce cas, au bout du compte, c'est bien l'allemand qui chasserait le walser). Mais c'est aussi à un tel stade de gravité historique, que d'autres éléments prennent le relais semble-t-il. On le voit bien dans le cas de cette communauté très caractéristique du grand système alpin par commodité parfois qualifié de tyrolien. Le paysage, l'écologie, l'habitat, les traditions locales peuvent venir au secours de la langue, pourtant mal en point. Car c'est effectivement à ce stade que l'on prend en général conscience avec acuité d'un tissu ethno-écologique dont la langue peut devenir du coup le fer de lance. Mais n'est-ce pas alors trop tard ?

4.4. Action sur les langues minoritaires.

Une idée très compréhensible et honorable dans son principe est que pour sauver les langues minoritaires, il faut les solidifier par le biais de normes centralisatrices, les amener concrètement à l'École, les officialiser etc., bref les armer d'un faisceau de propriétés leur permettant de rivaliser avec les langues dominantes.

On voit pourtant que dans différents cas expérimentaux volontaristes depuis quelques décennies (cf. Espagne, Italie), cela ne paraît pas suffire. Il faut donc bien admettre que l'action politique directe sur les langues est limitée par la mécanique fonctionnelle et économique du paysage sociolinguistique. Par exemple il semblerait que le castillan continue de progresser en tant que langue de l'économie et du pouvoir alors même que différents signaux le montrent directement menacé par les langues propres (dans les provinces autonomes du moins). Si l'on remonte au nord, on comprend de mieux en mieux qu'il ne suffit pas de créer des mots occitans pour l'informatique, le commerce, l'industrie etc. pour que l'occitan parvienne à concurrencer le français (ou l'anglais) dans de tels domaines. On remarque d'ailleurs que cette volonté passe souvent assez mal auprès des jeunes apprenants de la langue régionale, qui peuvent juger cette volonté plutôt ridicule (Verny 2002).

En revanche, des langues pour lesquelles la production de normes « centrales » est faible ou peu avancée, semblent assez souvent *de facto* en meilleure posture que des langues pour lesquelles

l'engagement militant est très fort et ancien⁶³. Ceci rejoint l'idée plus haut évoquée que toute normalisation centralisante et autoritaire porte en elle les germes d'une réduction des variétés dialectales de la langue présumée être aidée, ce que l'on a déjà évoqué pour le catalan, modèle du genre (4.1, Perea 2007). Au point qu'on peut se demander si le remède ne risque pas bien souvent de faire disparaître en même temps la maladie et le malade.

5. Les sirènes de la mondialisation

Depuis quelques années, on parle beaucoup de mondialisation, de globalisation. Il n'est pas question de revenir en détail sur de tels concepts en constante réfection, et surtout très discutables car présentés comme des évidences à la communauté intellectuelle. Toutefois, si l'on admet que la globalisation économique puisse se définir comme l'extension du modèle capitaliste au reste du monde, on peut aussi comprendre que cette extension puisse éventuellement créer les bases d'une homogénéisation économique et culturelle de surface. Car il s'agit bien, pour l'instant du moins, d'un phénomène de surface, et surtout d'un phénomène par rapport auquel nous manquons de recul sensé. Trop de gens, des linguistes notamment, donnent un crédit exagéré à ce genre de considérations. L'anglicisation à peu près intégrale du monde, liée à la fameuse mondialisation, était vue par certains comme inévitable il y a quelques années seulement. On remarquera que, très versatile, comme les cours de la bourse, le discours peut changer en quelques années. Aujourd'hui c'est le chinois que tous les enfants du monde devraient apprendre pour survivre au siècle suivant. La prochaine panne du système-monde de l'Orient (Chine, Inde) pourrait nous ramener au portugais (ou « brésilien ») si le Brésil se confirmait sur le devant de la scène et, pourquoi pas un jour, au catalan si cette langue poursuivait son ascension économique et politique au sein de la communauté européenne.

On pourrait faire de ce genre d'épouvantails et de mirages collectifs un catalogue très amusant, si l'on ne perdait pas de vue par ailleurs que ce climat perpétuel d'annonces à l'emporte-pièces séduit en

⁶³ Ayant réfléchi durant plusieurs années sur cette question à propos de la Bretagne romane je me permets de renvoyer globalement le lecteur qui souhaiterait approfondir à quelques travaux cités en bibliographie (Manzano 1996a, 1996b, 1997, 2003a, 2003b, 2006a).

permanence des intellectuels, des universitaires, des politiques, des décideurs.

L'histoire ne cesse pourtant de donner des leçons, jamais retenues. On rappellera que certaines élites françaises du XVI^e siècles étaient persuadées par différentes Cassandre de l'époque que leur langue allait être bientôt remplacée par l'italien, tant celui-ci influait sur la cour du roi. Plus tard les élites de différents pays d'Europe tenaient pour argent comptant l'universalité prochaine de la langue française, nouveau latin de l'Europe, sur la base des catalogues de poncifs de Bouhours (XVII^e) ou de Rivarol (XVIII^e). La vision ainsi développée d'une expansion européenne puis planétaire du français, sans être totalement infondée (car le français s'est effectivement répandu depuis l'époque classique⁶⁴), péchait par la chimère qui lui fut souvent associée ensuite : celle d'une diffusion polyvalente en tous lieux, en toutes activités, en tous contextes, en tous milieux sociaux⁶⁵. Or c'est ici que le bât blesse : une langue, quelle qu'elle soit, ne peut servir à l'expression de tous les besoins d'une pluralité de sociétés hétérogènes historiquement, culturellement, sociologiquement, et variant elles-mêmes suivant les lieux, les activités, les contextes, les milieux sociaux.

Ce phénomène, imparable en dehors des rêveries littéraires, vaut évidemment pour les langues d'hier, d'aujourd'hui et de demain. Il vaut aussi pour les langues véhiculaires comme d'ailleurs pour les langues régionales. C'est véritablement un phénomène de sociolinguistique générale fondamentalement contradictoire des simplifications abusives qui circulent en matière d'impacts langagiers de la « mondialisation ».

Retenons toutefois l'hypothèse d'une matrice économique (et par la suite vaguement culturelle) qui impliquerait en effet, presque d'office, une harmonisation des systèmes de communication et de lecture du

⁶⁴ Outre les supports littéraires et culturels de la diffusion, on rappellera ici que le français et nombre de langues européennes ont directement bénéficié des colonisations hors d'Europe puis des évolutions politiques post-coloniales jusqu'à nous, lesquelles ont généralement conforté la position des langues véhiculaires issues de la colonisation.

⁶⁵ Les fondateurs du français classique élitaire (dans la lignée de Vaugelas et de l'Académie), ne souhaitaient pas une telle expansion du français, langue d'élite exclusivement conçue pour d'autres élites en France et à travers l'Europe. Ce n'est qu'aux XVIII^e et XIX^e siècles (dans la suite de la période révolutionnaire notamment) qu'allait se développer l'idéologie d'une langue potentiellement capable de passer en tous lieux et tous milieux, langue de l'universalité, de la libération et de la liberté etc.

monde, tendant notamment à réduire le nombre des langues d'usage : dans un tel paradigme on considère en effet que trop de langues nuiraient à une communication efficace, c'est un des fondements du raisonnement mondialiste⁶⁶.

Derrière ce système présenté comme éminemment concurrentiel (= le marché), on retrouvera forcément la question de la primauté de certaines langues et des compressions symboliques qui s'en suivent : l'anglais (ou l'anglo-américain) est-il indiscutablement la langue internationale, ou le devient-il sous nos yeux ? quelles places pour la francophonie ? l'hispanophonie ? la lusophonie ? l'arabophonie ? etc. qui intéressent directement notre secteur d'observation. On voit que le problème dépasse immédiatement la seule question des langues minoritaires ou régionales, mais que la solution pour ces dernières ne peut être trouvée que dans le cadre d'une analyse systémique.

Certains linguistes tentent de modéliser par avance le système international de contact et de dynamique des langues qui semble se mettre en place sous nos yeux. Louis-Jean Calvet, par exemple évoque depuis quelques années un système dit « gravitationnel »⁶⁷. Bien qu'on ne puisse proposer de loi générale, cet auteur pense qu'un modèle trifonctionnel d'offre tend (et tendra probablement de plus en plus) à proposer aux locuteurs :

1) une langue internationale, utilisable dans les relations avec l'extérieur. L'anglais est la langue vedette sur ce point, bien qu'il ne soit pas à proprement parler la seule langue internationale pour l'instant. [cercle 1]

2) une langue nationale ou plus justement une langue de l'État, dès lors que la nation se définit très différemment à travers les cultures.

⁶⁶ La perversité du raisonnement est usuellement démontrée par les mécanismes majoritaires de formation ou de recrutement dans les sociétés d'Europe occidentale, ces mécanismes tendant à focaliser sur une langue principale, l'anglais. Que l'anglais soit langue internationale est incontestable, le problème n'est pas là. En revanche le concevoir comme langue d'interaction internationale principale ou exclusive fausse les réalités socio- et écolinguistiques. Ainsi est entretenue une imagerie de fait contraire aux besoins plurilingues et pluri-culturels des sociétés en présence. On peut toutefois considérer qu'une prise de conscience de cette anomalie fondamentale a débuté et qu'elle s'affirmera plus encore durant les décennies qui viennent. Pour être cohérente et efficiente, il devient évident que la réflexion devra concerner (avec des degrés différents) les langues internationales, les langues nationales mais aussi les langues régionales.

⁶⁷ Voir par exemple, sur son site personnel (<http://perso.wanadoo.fr/Louis-Jean.Calvet>) : « Globalización, lenguas y políticas lingüísticas » (Santiago du Chili, 2005).

C'est la langue normalisée, standardisée, qui permet généralement de s'insérer dans la vie publique du pays. [cercle 2]

3) une langue grégaire, qui peut être une variante régionalisée de la langue de l'État (comme le français régional de Provence en France ou l'arabe dialectal au Maroc (et plus généralement au Maghreb). Mais il peut s'agir d'une langue distincte de la langue d'État, appartenant ou pas à la même famille que cette dernière. Tels seront les cas du catalan en Espagne, du berbère en Afrique du nord ou du basque en France et en Espagne, avec des degrés d'éloignement typologique évidents. [cercle 3]

Ce modèle suppose des cercles de communication en même temps différents et complémentaires, appelant des registres, variétés de langues, ou langues différentes, utilisés en fonction du contexte et adaptés à ce contexte. En cela, cette ventilation est plus conforme aux réalités sociolinguistiques ou écolinguistiques que la vulgate mondialiste plus haut évoquée.

De 1 vers 3 on va vers un maximum de familiarité et d'adaptation à l'environnement local, ce que la plupart des dialectologues savent depuis longtemps : il y a des choses qu'on peut dire avec la langue grégaire (cercle 3) et qu'on dira nettement moins bien (voire pas du tout) avec la langue du cercle 2. Quant à la langue du cercle 1, que permet-elle d'exprimer au juste, quand on y regarde de plus près : des choses techniques bien entendu, relativement nettes et précises (encore faut-il alors dominer un certain nombre de technocetes de la langue internationale), mais surtout une succession de banalités car toute globalisation rabote automatiquement, impitoyablement, les innombrables différences identitaires portées par les langues du cercle 2 et du cercle 3. Justement parce que la langue internationale ne peut absolument pas les assumer.

S'il y a donc un argument puissant pour éviter de paniquer trop vite devant les effets possibles d'une supposée mondialisation langagière, c'est bien celui-ci : sauf à imaginer une économie standardisée et une culture internationale unique (ce qui est fondamentalement utopique, non souhaitable de toute façon, et que les sociétés concrètes se chargeront bien de faire voler en éclats), les gens vivant heureusement et pour longtemps dans des lieux précis, anciens et pétris d'histoire,

ils trouveront toujours le moyen de réaffirmer leur identité régionale et sociale, y compris de plus en plus à travers les langues du cercle 2. Si l'on regarde par exemple ce qui se passe dans le Golfe du Lion (Manzano 2005a, 2006b) on peut penser que la langue d'oc va disparaître ou a déjà disparu en plus d'un endroit, non sans transmettre des données indispensables des identités originelles au français régional cette région. On peut voir aussi que de nouvelles minorités s'installent, de principe plus urbain (d'origine maghrébine par exemple) en s'intégrant progressivement elles-mêmes par le biais du français régional, mais peu ou pas du tout par le biais des langues régionales⁶⁸. La question de savoir si l'anglo-américain pourrait remplacer à terme les langues « nationales » ou les langues « minoritaires » n'a guère d'intérêt. Elle est surtout absurde. Pour ne considérer que l'espace méditerranéen dont il est ici question on ne voit pas comment et pourquoi (fonctionnellement) cela se produirait, même s'il existe quelques indices concordants de descente des langues germaniques vers la Méditerranée (Manzano 2006b). L'hypothèse de remplacements d'envergure n'est pas plus sérieuse aujourd'hui qu'hier, car elle suppose toujours l'illusion d'une culture ou acculturation communes, ce qui n'est pas pour demain.

A ces réserves près, quelle pourrait être la place des langues minoritaires dans l'espace méditerranéen ?

Tout porte à croire que quelques langues bien placées dans l'histoire profonde de cette région, domineraient plus encore à terme la Méditerranée occidentale. Le français et l'espagnol (castillan) sans doute, véritables langues de contact entre rive nord et rive sud (laissant sans doute une place croissante à l'anglais), mais assurant aussi le truchement entre la Méditerranée et le reste du monde. Ces langues domineraient probablement l'arabe d'une part (langue identitaire et véhiculaire d'une grande partie du Maghreb), ainsi que l'italien (dimension internationale moindre).

Quel serait alors l'intérêt concret du maintien de langues minoritaires, notamment de faible extension géographique,

⁶⁸ Un peu plus au sud, la communauté gitane du Roussillon s'intègre quant à elle en catalan et en français régional. Mais les racines historiques du processus sont ibériques pour cette communauté, ce qui amène à effectuer une distinction fondamentale par rapport à l'intégration linguistique des Maghrébins. Certains de ces phénomènes ont été récemment étudiés par Escudero (2004).

relativement isolées typologiquement et numériquement réduites ? Il est certain que leur valeur de langues de communication nationale ou internationale ne pourra être mise en avant dans un schéma d'ensemble, car de ce point de vue elles sont en position d'inégalité avec les précédentes et les choses ont peu de chances de changer fondamentalement dans l'avenir. Vraisemblablement seraient-elles alors promises à disparition, marginalisation, ou converties en « langues-réserves », patrimoniales, à protéger etc. Le mouvement, comme on l'a vu, a déjà commencé et les actions vigoureuses de promotion ne paraissent pas en mesure de freiner cette inclination générale. On suivra donc avec d'autant plus d'intérêt l'évolution du dossier des langues régionales d'Espagne en réaffirmation, ou celui du berbère au Maghreb.

Il reste une grande inconnue, et de là pourrait venir le salut. Comme je l'écrivais il y a quelques années (2000 : 76, note 50) : « Il y a cinquante ans, la disparition d'une plante, d'un animal, d'une langue, dans le monde ou chez nous heurtait peu le Français moyen. Dans nos résultats d'enquêtes revient aujourd'hui, de plus en plus fort, l'écho qu'une langue régionale qui n'est plus parlée c'est l'ensemble des traditions, des connaissances, bref l'âme du pays qui s'en va ».

Le propos vaut au delà de l'espace français. Incontestablement, un profond mouvement écologique émerge depuis quelques décennies et les discours simplistes sur la mondialisation ne peuvent que le renforcer. Ce mouvement est de moins en moins mondain et artificiel, car il amène de plus en plus de peuples méditerranéens à prendre conscience du fait que la recherche obsédante d'unicité et de standardisation en quelque matière que ce soit est très mauvaise conseillère. Il faut des langues « ponts », mais celles-ci ne permettent pas tout. Il ne s'agit pas, en sens inverse, de prôner un éclatement perpétuel des objets. Dans ce raisonnement, la préservation de la diversité passera de plus en plus au premier plan, de manière non anecdotique mais dynamique et créative, dans une région du monde par ailleurs très fragile sur le plan de la biodiversité. Si ce mouvement se poursuit et se renforce, sans doute pourra-t-on alors trouver de vraies solutions pour que les identités régionales ne soient plus vécues comme des handicaps ou des scories d'un monde ancien, mais au contraire comme les éléments constitutifs d'un avenir commun.

Références bibliographiques

- Aracil, Lluís Vicent. 1986. *Papers de sociolingüística*. Barcelona, Magrana.
- Bec, Pierre. 1970-1971. *Manuel pratique de philologie romane*. Paris, Picard [2 volumes].
- Bec, Pierre. 1973. *La langue occitane*. Paris : Presses Universitaires de France [1^{ère} éd. 1963].
- Cerquiglini, Bernard (dir.). 2003. *Les langues de France*. Paris, Presses Universitaires de France.
- Clanché, François. 2002. « Langues régionales, langues étrangères : de l'héritage à la pratique ». Paris, INSEE Première, n° 830, février.
- Entwistle, William J. 1982. *Las lenguas de España : Castellano, Catalán, Vasco i Gallego-portugués*. Madrid, Istmo, 4ème éd.
- Escudero, Jean-Paul. 2004. *Les Gitans catalans et leur langue : une étude réalisée à Perpignan*. Ed. De la Tour Gile.
- Grandguillaume, Gilbert. 1983. *Arabisation et politique linguistique au Maghreb*. Maisonneuve & Larose.
- Grassi Corrado, Sobrero Alberto, Telmon Tullio. 2005. *Fondamenti di dialettologia italiana*. Laterza.
- Grassi Corrado, Sobrero Alberto, Telmon Tullio. 2006. *Introduzione alla dialettologia italiana*. Laterza.
- Leray, Christian & Manzano, Francis (dir.). 2003. *Langues en contact : Canada, Bretagne. Cahiers de Sociolinguistique n° 7*. Presses Universitaires de Rennes.
- Leclerc, Jacques. Site internet : *L'aménagement linguistique dans le monde* (<http://www.tlfq.ulaval.ca/AXL/>)
- Lodge, Anthony. 1997. *Le Français. Histoire d'un dialecte devenu langue*. Paris, Fayard.
- Manzano, Francis. 1996. « Sur les mécanismes du paysage sociolinguistique et identitaire d'Afrique du Nord ». *Revue Langage et Société*, Paris, n° 75, mars.
- Manzano, Francis (dir.). 1996a. *Langues et parlers de l'Ouest : pratiques langagières en Bretagne et Normandie. Cahiers de sociolinguistique n° 1*, Presses Universitaires de Rennes.
- Manzano (F.). 1996b. « Sur le statut sociolinguistique du gallo : une identité en question ». In Manzano (F.). 1996a.
- Manzano, Francis (dir.). 1997. *Vitalité des parlers de l'Ouest et du*

- Canada francophone. Cahiers de sociolinguistique* n° 2-3. Presses Universitaires de Rennes.
- Manzano, Francis. 1997b. « Le gallo à la fin du XX^e siècle ». in Manzano (F.). 1997.
- Manzano, Francis. 1999. *Langues du Maghreb et du sud méditerranéen*. Cahiers de sociolinguistique n° 4. Presses Universitaires de Rennes.
- Manzano, Francis. 1999a. « La fin du berbère à Djerba (Tunisie) ». In Manzano (F.). 1999.
- Manzano, Francis. 1999b. « Eléments d'archéologie du paysage sociolinguistique d'Afrique du Nord : quelques considérations sur les contours proto-romans avant arabisation (Tunisie, Algérie orientale) ». In Manzano (F.). 1999.
- Manzano, Francis. 2000. « De la dynamique du français, langue d'État et de pouvoir. Quelques rappels sur la diffusion entre le XV^e et le XX^e siècle. ». In Dubois (C.), Kasbarian (J.-M.) & Quéffélec (A.). 2000. *L'expansion du français dans les Suds (XV^e-XX^e siècles)*. Hommage à Daniel Bagionni. Aix en Provence, Publications de l'Université de Provence.
- Manzano, Francis. 2003a. « Sur le contact français-gallo. Observations diachroniques, sociolinguistiques et anthropologiques ». in Leray (Ch.) & Manzano (F.). 2003.
- Manzano, Francis. 2003b. « Diglossie, contacts et conflits de langues... à l'épreuve de trois domaines géo-linguistiques : Haute Bretagne, Sud occitano-roman, Maghreb ». In Blanchet (Ph.) & Robillard (D. de). 2003. *Langues, contacts, complexité. Perspectives théoriques en sociolinguistique*. Cahiers de sociolinguistique, n° 8. Presses Universitaires de Rennes.
- Manzano, Francis. 2004. « Pratiques et représentations linguistiques à la marge sud du territoire français (Languedoc, Roussillon) ». In : Moreau (M.-L.). 2004. *Langues de frontières et frontières de langues*. Revue électronique *Glottopol*, n°4, juillet.
- Manzano, Francis. 2004b. « Situation and uses of Occitan in Languedoc ». In *The Sociolinguistics of Southern 'Occitan' France, revisited. The International Journal of the Sociology of Language*, n° 169 (dir. Ph. Blanchet & H. Schiffman). Prof. Joshua Fishman, General Editor. 2004.
- Manzano, Francis. 2005. « Le Maghreb, laboratoire de la francophonie ? ». In A.-O. Alami & C. Valat (dir.), *La francophonie*

arabe : pour une approche de la littérature arabe francophone. Revue *Horizons maghrébins*, n° 52/2005. Toulouse, Presses Universitaires du Mirail.

Manzano, Francis. 2005a. « Langues et identités du Golfe du Lion et de la France méditerranéenne ». Revue *La France Latine*, n° 141. Rennes.

Manzano, Francis. 2006a. « Les langues régionales de France sont-elles égales dans le recul ? Eléments de réflexion et de programmation pour une approche anthropologique, écologique et systémique des langues de France ». Th. Bulot, V. Fillol, C. Moïse (dir.). Revue électronique *Marges linguistiques*, janvier.

Manzano, Francis. 2006b. « Considérations sur les langues minoritaires du Golfe du Lion et de la France méditerranéenne. Langues d'origine et vernaculaires dans le cadre d'une prospective du système général ». Revue *La France latine*, n° 143. Rennes.

Manzano, Francis. 2006c. « Berbères, berbérité. Noms, territoires, identités. Considérations anthropologiques, lexicographiques et onomastiques en suivant Germaine Tillion (2000) et quelques autres ». In Manzano (F.) dir. *Noms propres, dynamiques identitaires et sociolinguistiques. Cahiers de sociolinguistique* n° 11. Presses Universitaires de Rennes.

Manzano, Francis. 2007. « Sur l'implantation du français au Maghreb : systémique et fractures identitaires au tournant des XIX^e et XX^e siècles ». Revue *Le français en Afrique*, Nice, n° 22.

Musso, Michele (2007). « Labili confini », dans Ronco & Musso (2007). Pages 25-31.

Perea, Maria-Pilar. 2007. "Dialectalization or the Death of a Dialect: The Case of the Catalan Subdialect Spoken in the Costa Brava". *Dialectologia et Geolinguistica*. 15/2007.

Ronco, Imelda & Musso, Michele. 2007. *Eiscchemgseiti. Les dictons van a voart*. Associazione culturale Augusta. Tipografia valdostana.

Salvador, Gregorio. 1987. *Lengua española y lenguas de España*. Ariel Lingüística.

Salvador, Gregorio. 1992. *Política lingüística y sentido común*. Madrid, Istmo.

Telmon, Tullio. 1990. *Guida allo studio degli italiani regionali*. Edizioni dell'Orso.

Telmon, Tullio. 1992. *Le minoranze linguistiche in Italia*. Edizioni dell'Orso.

- Tudela, Joan (ed.). 1986. *El futur del català. Una radiografia sociolingüística*. Barcelona, Publicacions El Món.
- Verny, Marie-Jeanne. 2002. « Une enquête en lycée : images et représentations de l'occitan ». In Boyer (H.) & Gardy (Ph.). 2002. *Dix siècles d'usages et d'images de l'occitan*. Paris, L'Harmattan.
- Zali, Anne & Berthier, Annie. 1997. *L'aventure des écritures/Naissances*. Bibliothèque nationale de France.
- Zamora Vicente, Alonso. 1996. *Dialectologia española*. 2^{ème} éd. Madrid, Gredos.

Annexes

A -Textes relatifs à la Catalogne (provinces de Barcelona, Gerona, Lérida, Tarragona)

Statut d'autonomie de 1979

Article 3

- 1) La langue propre à la Catalogne est le catalan.
- 2) Le catalan est la langue officielle de la Catalogne, comme l'est aussi le castillan qui est la langue officielle de tout l'État espagnol.
- 3) La Generalitat doit garantir l'usage normal et officiel des deux langues et adopter les mesures nécessaires pour en assurer la connaissance et créer les conditions qui permettent d'atteindre leur pleine égalité en ce qui concerne les droits et les devoirs des citoyens de la Catalogne.
- 4) Le parler aranais doit être enseigné et faire l'objet d'une protection et d'un respect particuliers.

Statut d'autonomie de 2006 (annule les précédents)

[La langue propre et les langues officielles]

- 1) La langue propre de la Catalogne est le catalan. Comme tel, le catalan est la langue d'usage normale et prioritaire des administrations publiques et le moyen de communication public en Catalogne, ainsi que la langue normalement utilisée comme véhicule et moyen d'apprentissage dans l'enseignement.
- 2) Le catalan est la langue officielle de la Catalogne, aussi bien que le castillan qui est la langue officielle de l'État espagnol. Tous les individus ont le droit d'utiliser les deux langues officielles et les citoyens de la Catalogne ont le droit et le devoir de les connaître. Les pouvoirs publics de la Catalogne doivent fixer les mesures nécessaires pour faciliter l'exercice de ces droits et l'accomplissement de ce devoir. Conformément aux dispositions de l'article 32, il ne peut exister de discrimination pour l'usage de l'une ou l'autre langue.
- 3) La Generalitat et l'Etat doivent entreprendre les actions nécessaires pour la reconnaissance du caractère officiel du catalan dans l'Union européenne et la présence et l'usage du catalan au sein des organismes internationaux ainsi que dans les traités internationaux à contenu culturel ou linguistique.
- 4) La Generalitat doit promouvoir la communication et la coopération avec les autres Communautés et les autres territoires qui partagent ce patrimoine linguistique avec la Catalogne. À cette fin, la Generalitat et l'État, comme il convient, peuvent souscrire à des conventions, traités et autres mécanismes de collaboration pour la promotion et la diffusion extérieure du catalan.
- 5) La langue occitane, appelée aranais dans le val d'Aran, est la langue propre

et officielle de ce territoire ainsi que la langue officielle en Catalogne, conformément aux dispositions du présent Statut et des lois de normalisation linguistique.

B-Textes relatifs à la communauté valencienne (provinces de Castellón de la Plana, Valencia et Alicante)

Statut d'autonomie (1982)

Article 7

- 1) Les deux langues officielles de la Communauté autonome sont le valencien et le castillan.
- 2) Tous ont le droit de les connaître et de les utiliser.

Loi sur l'usage et l'enseignement du valencien (1983)

Article 1

- 1) La présente loi a pour objectif général d'appliquer et de développer les dispositions de l'article 7 du Statut d'autonomie qui réglemente l'usage normal et officiel du valencien dans toutes les sphères de la vie en commun, ainsi que son enseignement.
- 2) Sur cette base, les buts et les objectifs spécifiques de la présente loi sont les suivants :
 - a) Que le droit de tous les citoyens de connaître et de faire usage du valencien constitue un fait.
 - b) Protéger la récupération du valencien et garantir son usage normal et officiel.
 - c) Réglementer les critères d'application du valencien dans l'administration, dans les moyens de communication sociale et dans l'enseignement.
 - d) Délimiter les territoires sur lesquels prédomine l'usage du valencien et du castillan.
 - e) Garantir, selon des principes de progression et de volontariat, la connaissance et l'usage du valencien sur tout le territoire de la Communauté.

Statut d'autonomie (2006)

Article 6

- 1) La langue propre de la Communauté valencienne est le valencien.
- 2) La langue valencienne est officielle dans la Communauté valencienne, tout comme le castillan, qui est la langue officielle de l'État. Tous ont le droit de les connaître et de les utiliser, et de recevoir son enseignement dans la langue valencienne.
- 3) La Generalitat garantit l'usage normal et officiel des deux langues, et adopte les mesures nécessaires pour assurer leur connaissance.

C-Textes relatifs aux Asturies

Statut d'autonomie (1982)

Article 4

- 1) Le bable bénéficiera de protection. Son emploi et sa présence dans les médias et au sein du système scolaire seront encouragés, en respectant dans tous les cas les variétés locales et le caractère volontaire de son apprentissage.
- 2) Une loi de la Principauté réglera la protection, l'usage et la promotion de l'asturien.

Article 10

- 1) La principauté des Asturies a compétence exclusive dans les domaines mentionnés ci-après, sans préjudice de ce qui est établi aux articles 140 et 149 de la Constitution: [...]
- 2) la promotion et la protection de l'asturien sous ses diverses formes qui, en tant que variétés linguistiques, sont pratiquées dans le territoire de la principauté des Asturies.

D-Textes relatifs à l'Aragon

Loi du 10 mars 1999, relative au patrimoine culturel aragonais

Article 1 (Objet)

La présente loi a pour objectif la protection, la conservation, la progression, la recherche, la diffusion, la promotion, le développement et la formation afin de transmettre aux générations futures le patrimoine culturel aragonais et les biens qui le composent, quel que soit son régime juridique et fonctionnel, en garantissant son usage en tant qu'apport social et facteur de développement durable pour l'Aragon.

Article 2 (Patrimoine culturel aragonais)

Le patrimoine culturel aragonais est composé de tous les biens matériels et immatériels en relation avec l'histoire et la culture d'Aragon, qui présentent un intérêt anthropologique, anthropique, historique, artistique, architectonique, mobilier, archéologique, paléontologique, ethnologique, scientifique, linguistique, documentaire, cinématographique, bibliographique ou technique, ou qui ont été découverts dans la mesure où ils se trouvent en surface comme dans le sous-sol ou sous la surface des eaux.

[...]

Article 4 (Langues minoritaires)

L'aragonais et le catalan, langues minoritaires d'Aragon, dans le cadre duquel sont comprises les diverses spécificités linguistiques, constituent une richesse

culturelle propre et doivent être particulièrement protégés par l'Administration.

Article 5 (Droit d'usage)

Toutes les personnes ont le droit de bénéficier du patrimoine culturel aragonais, conformément aux dispositions prévues dans la réglementation en vigueur.

Article 6 (Devoir de conservation)

- 1) Toutes les personnes ont le devoir de conserver le patrimoine culturel aragonais, en l'utilisant rationnellement et en adoptant des mesures préventives, de défense et de récupération nécessaires pour garantir son usage par les générations futures.
- 2) Dans tous les cas, les personnes qui ont connaissance d'une situation de danger ou de destruction, consommée ou imminente, ou de la détérioration d'un bien du patrimoine culturel aragonais devront, dans les plus brefs délais possibles, les porter à la connaissance de la Mairie correspondante, du Département responsable du patrimoine culturel ou des forces et des corps de sécurité, lesquels vérifieront l'objet de la dénonciation et agiront conformément au droit. L'Administration de la Communauté autonome portera à la connaissance du requérant les actions entreprises.
- 3) Les associations culturelles aragonaises enregistrées légalement pourront collaborer avec l'Administration dans les tâches indiquées aux points précédents.

E-Textes relatifs à la Vallée d'Aoste

**Statut spécial de la Vallée d'Aoste - Constitution de la Région
Loi constitutionnelle de 1948, mise à jour en 2001.**

Article 1

La Vallée d'Aoste est constituée en Région autonome, dotée de la personnalité juridique, dans le cadre de l'unité politique de la République italienne, une et indivisible, sur la base des principes de la Constitution et selon le présent Statut. Le territoire de la Vallée d'Aoste comprend les circonscriptions des communes qui en font partie à la date de l'entrée en vigueur de la présente loi. La Région a pour chef-lieu Aoste.

Langue et organisation des écoles

Article 39

- 1) Dans les écoles de n'importe quel ordre ou degré qui dépendent de la Région, un nombre d'heures égal à celui qui est consacré à l'enseignement de l'italien est réservé, chaque semaine, à l'enseignement du français.
- 2) L'enseignement de quelques matières peut être dispensé en français.

Article 40

- 1) L'enseignement des différentes matières est organisé selon les dispositions et les programmes en vigueur dans l'État, moyennant des adaptations opportunes aux nécessités locales.
- 2) Ces adaptations, ainsi que la liste des matières pouvant être enseignées en français, sont approuvées et rendues exécutoires, après consultation de Commissions mixtes composées de représentants du Ministère de l'Instruction publique, de représentants du Conseil de la Vallée et de représentants du corps enseignant.

Article 40 - bis

- 1) Les populations de langue allemande des communes de la Vallée du Lys indiquées par loi régionale ont droit à la sauvegarde de leurs caractéristiques et de leurs traditions linguistiques et culturelles.
- 2) Aux populations visées au premier alinéa est assuré l'enseignement de la langue allemande dans les écoles au moyen des adaptations nécessaires aux besoins locaux.

**EXTRAIT DES PUBLICATIONS EN VENTE AU CENTRE
D'ÉTUDES LINGUISTIQUES**

(CEL, 18 rue Chevreul, 69362 Lyon cedex 07)

Les publications suivantes (antérieures à 2011) sont toujours disponibles et peuvent être commandées par courriel :

cel@univ-lyon3.fr

[Prix port payé]

Série « dialectologie »

- Le croissant : le concept et le mot. Contribution à l'histoire de la dialectologie française au XIX^e siècle*, par Guylaine Brun-Trigaud, 1990. (53,36 € TTC).
- Dialectologie et Littérature du domaine d'oïl occidental. Sur le thème des enquêtes dialectales*. Actes du IV^e colloque tenu à Poitiers, du 6 au 8 mars 1990, publiés par Brigitte Horiot, 1995. (22,87 € TTC).
- Français de France et Français du Canada. Les parlers de l'Ouest de la France, du Québec et de l'Acadie*, sous la direction de Pierre Gauthier et Thomas Lavoie, 1995. (42,69 € TTC).
- L'onomastique au carrefour des sciences humaines*. Actes du XI^e colloque de la Société française d'Onomastique, Université Lyon 3 (10-13 octobre 2001), édités par Brigitte Horiot et Corina Veleanu, 2004. (21,10 € TTC).
- La dialectologie hier et aujourd'hui (1906-2006)*. Actes du colloque international tenu à l'Université Lyon 3 (7, 8 et 9 décembre 2006). Textes rassemblés et édités par Brigitte Horiot, 2009. (25 € TTC).

Série « Langues en contact »

- Le français, ailleurs et toujours : place et fonctions du français dans les autres langues*. Actes du colloque international tenu à l'Université Lyon 3 (9 et 10 mai 2005). Textes rassemblés et édités par Brigitte Horiot, 2008. (22 € TTC).

Série « démodialectologie »

- Des données empiriques à l'objet d'étude : la constitution d'un corpus de démodialectologie (français populaire d'Abidjan)*, par Mariette Meunier-Crespo, 2 tomes, 1994. (42,69 € TTC).

Série « histoire de la linguistique »

- Introduction à l'étude du « Dictionnaire étymologique ou origines de la langue française » de Gilles Ménage (1694). Les étymologies de Ménage : science et fantaisie*, par Isabelle Leroy-Turcan, 1991. (21,95 € TTC).
- Noëls en français et en dialectes du XVI^e siècle*, édités par Marina Fey, 2008. (22 € TTC).

Série « didactique »

- Il signor Bianchi in ospedale. Méthode audio-orale d'italien pour le personnel hospitalier*, par Jean-Louis Chauzit, 1991. (12,20 € TTC).
- Le fonctionnement de la démocratie espagnole actuelle (textes, documents et terminologie)*, par Mariette Meunier-Crespo, 1998. (22, 87 € TTC).

Série « germanique ancien »

- Althochdeutsch Syntax und Semantik. Akten des Lyonner Kolloquiums zur Syntax und Semantik des Althochdeutschen (1-3 März 1990)*, publiés par Yvon Desportes, 1992. (30,49 € TTC).
- Les préverbes et la prévervation en allemand au IX^e siècle*, par Yvon Desportes, 4 tomes, 1998. (213, 43 € TTC).

Série « langues de spécialité »

- Le subjonctif futur dans la langue juridique espagnole actuelle*, par Mariette Meunier-Crespo, 1998. (21,34 € TTC).

Hors série, mélanges, collectifs

- Mélanges offerts au professeur Lothar Wolf « Je parle, donc je suis... de quelque part »*. Textes rassemblés et édités par Brigitte Horiot, Elmar Schafroth, Marie-Rose Simoni-Aurembou, 2005. (36,92 € TTC).

Centre d'Etudes Linguistiques **[Linguistique, dialectologie]**

Le Centre d'Etudes Linguistiques (EA 1663) est une équipe d'accueil de l'Université de Lyon (Université Jean Moulin – Lyon 3, Faculté des langues), dont les locaux se trouvent au 18, rue Chevreul – 5^{ème} étage (Lyon, 7^{ème} arrondissement).

L'équipe regroupe des chercheurs romanistes, anglicistes ou germanistes autour des thématiques de la linguistique comparative, de la dialectologie, de la sociolinguistique et de la pragmatique.

Les zones géographiques couvertes par le CEL sont celles de l'Europe de l'Ouest et de la Méditerranée occidentale (langues en contact). Les domaines d'expertise du groupe concernent principalement les langues et cultures régionales de France et de l'Europe romane, l'anthropologie linguistique, la dialectologie et la sociolinguistique (73^{ème} section du CNU et 7^{ème} section du CNU).

Le Centre d'Etudes Linguistiques est à ce titre support d'un master recherche spécifique « Linguistique et dialectologie », orienté vers la formation de jeunes chercheurs en sociolinguistique et dialectologie – programme sur la page :

http://facdeslangues.univ-lyon3.fr/medias/fichier/master-lce-linguistique-et-dialectologie-05-04-2011_1302260055589.pdf

Il est également laboratoire d'accueil et de direction pour des recherches doctorales en « Sciences du langage » et « Cultures et langues régionales » :

<http://facdeslangues.univ-lyon3.fr/formation/diplomes-nationaux/doctorat/doctorat-linguistique-linguistique-generale-langues-slaves-langues-romanes-didactique-des-langues-64398.kjsp?RH=LAN-RECHcel>

Les correspondances doivent être adressées au professeur directeur de l'équipe, et directeur des publications du CEL :

cel@univ-lyon3.fr
francis.manzano@univ-lyon3.fr

**Imprimé par le Service Édition
de l'Université Jean Moulin Lyon 3**

Le volume *Unité et diversité de la linguistique* comporte des interventions effectuées soit par des chercheurs du Centre d'Etudes Linguistiques, lors du séminaire « Les mercredis du CEL » programmé en 2008-2009, soit par différents invités au cours de rencontres organisées par les chercheurs romanistes ou anglicistes de l'équipe d'accueil EA 1663.

Les thématiques abordées sont celles de l'anthropologie linguistique, de la pragmatique, de la dialectologie et de la sociolinguistique, principalement appliquées aux corpus contemporains (français, anglais) et aux langues régionales en régression.

La résonance sociale de cet ensemble de textes est donc particulièrement nette.

Pour autant c'est bien sur fond d'unité de la linguistique qu'un temps de la recherche universitaire est ici évoqué, au moment où la linguistique lyonnaise amorce de nouvelles mutations collectives.

Contributions de :

Peter Stockwell (University of Nottingham), Luu Dinh Phuc, Mariette Meunier-Crespo (CEL), Denis Jamet (CEL), François Maniez (Université Lyon 2), Dan MacIntyre, Brian Walker (University of Huddersfield), Laure Gardelle (CEL, ENS de Lyon), Veronika Koller (Lancaster University), Martina Pitz (CEL), Manuel Jobert (CEL), Brigitte Horiot (CEL), James Costa (ICAR, INRP), Francis Manzano (CEL).