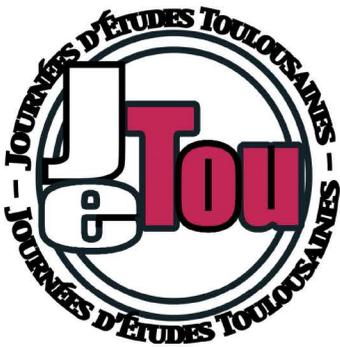


# Variation et variabilité dans les Sciences du Langage : analyser, mesurer, contextualiser

## Variation and variability in the Language Sciences: analysing, measuring, contextualising



Actes de la 4<sup>ème</sup> édition des JéTou  
16 et 17 mai 2013, Toulouse, France

Proceedings of the 4<sup>th</sup> edition of JéTou  
May 16<sup>th</sup> and 17<sup>th</sup> 2013, Toulouse, France



**JéTou 2013**  
**Variation et variabilité dans les**  
**Sciences du Langage :**  
**analyser, mesurer, contextualiser**

Actes des Journées d'Études Toulousaines 2013  
16 – 17 mai 2013  
Université de Toulouse II – Le Mirail



# Table des matières/ Table of contents

## Organisation des JéTou/ JéTou organisation

Appel à communications.....	6
Call for papers.....	9
Remerciements/ Acknowledgements.....	12
Comité scientifique/ Scientific committee.....	13
Comité d'organisation/ Organisation committee.....	15
Annuaire des participants/ Directory of participants.....	16

## Actes des JéTou/ Proceedings of JéTou

### - Conférencière invitée/ Conference by guest speaker

<b>Storying Research : Terrain(s), postures, interprétations autour du/ des plurilinguisme(s).....</b>	<b>20</b>
Danièle MOORE	

### - Session 1 : Français Langue Etrangère/ French as a Foreign Language

<b>« M e n j a z o v u t Nati. I love ma famille, t s h o v r e b a, et j'espère que c'est mutuel » : variations langagières multiples au sein d'un atelier d'écriture plurielle.....</b>	<b>21</b>
Noëlle MATHIS	

<b>Une nécessaire contextualisation : le cas d'une étude de la stratégie enseignante d'étayage, en classe de FLE, en milieu universitaire chinois.....</b>	<b>32</b>
Charlotte VALLAT	

<b>Cinéma et variabilité de l'expérience émotionnelle des apprenants de FLE.....</b>	<b>43</b>
Claire DEL OLMO	

## - Session posters 1/Poster Session 1

Ditransitive complementation in medical research articles.....53  
Sabiha CHOURA

Vers une taxonomie phonétique des sibilantes non voisées dans les  
variétés de l'espagnol.....59  
Anne-Sophie DE CLERCQ

Comparaison de types d'erreurs orthographiques en FLM et FLE : une  
revue d'articles.....64  
Katayoon KATOOZIAN

« Non les gens (ils) sont pas dangereux » : variation et redoublement  
du sujet.....70  
Géraldine ZUMWALD KÜSTER

## - Session 2 : Morphologie et Syntaxe/Morphology and Syntax

Comment traiter la variation dans la communication par SMS ? Le cas  
de l'interrogation totale.....76  
Alexander GURYEV

*Je saicherais pas te dire* : erreurs et variations flexionnelles dans la  
structure des paradigmes verbaux français.....88  
Xavier BACH & Louise Esher

Neoclassical compounds and language registers.....98  
Marine LASSERRE

La variation morphosyntaxique du français radiophonique algérien..109  
Nabil SADI

## - Conférencière invitée/Conference by guest speaker

La variation et les corpus : une entrée méthodologique, avec l'exemple  
du style.....121  
Françoise GADET

### **- Session 3 : Phonétique et Phonologie/Phonetics and Phonology**

**Methodological considerations in the study of sociophonetic variation in an underdocumented minority language: Somali Bantu Kizigua as a case study.....129**  
Holman TSE

**L'alternance entre [β] et [w] dans les terminaisons de l'imparfait de l'indicatif en catalan dans la région de La Selva.....140**  
Eva BOSCH I ROURA

**La variation entre l'hiatus et les diphtongues croissantes en portugais brésilien.....151**  
Vanessa MEIRELES

### **- Session posters 2/Poster Session 2**

**Etude de la variation écholalique chez un enfant autiste.....160**  
Carmen Lucia FLOREZ PULIDO

**Analyse psychopathologique comparée des usages du signe dans la névrose et la psychose.....167**  
Arthur MARY

**« L'économiquement correct » : une source d'ambiguïtés.....173**  
Loredana RUCCELLA

### **- Session 4 : Analyse du discours/Discourse analysis**

***Ethos* des locuteurs dans des émissions économiques radiophoniques : variables et variations.....180**  
Nathalie GERBER

### **- Session 5 : Langue des signes/Sign Language**

**Modélisation du référencement en langue des signes : Analyse des variantes gestuelles.....190**  
Monia BEN MLOUKA

## Appel à communications

Les doctorants des laboratoires de Sciences du Langage de Toulouse :

- CLLE–ERSS (Équipe de Recherche en Syntaxe et Sémantique)
- Octogone–Lordat (Centre Interdisciplinaire des Sciences du Langage et de la Cognition)
- LARA–CPST (Centre Pluridisciplinaire de Sémiotique Textuelle)

organisent la quatrième édition des JéTou (Journées d'études Toulousaines).

Ces journées ont pour but de rassembler des doctorants et jeunes docteurs (jusqu'à 3 ans après la soutenance de la thèse) issus des différentes disciplines des Sciences du Langage autour d'une thématique ouverte et transversale.

Cette édition 2013 sera consacrée à une réflexion sur le thème : « Variation et variabilité dans les Sciences du Langage : analyser, mesurer, contextualiser ».

Les notions de variation et de variabilité sont au cœur de nombreux domaines liés aux Sciences du Langage. Le linguiste peut aussi bien être amené à les interroger en tant qu'objet de recherche qu'au travers des problèmes méthodologiques qu'elles posent.

La variation comme objet d'étude offre ainsi un terrain de réflexion particulièrement riche et ouvert. Elle permet par exemple d'apporter un éclairage spécifique, et souvent fondamental, à nombre de problématiques à travers les travaux de contextualisation historique, théorique ou sociologique qu'elle peut impliquer. Par ailleurs, le positionnement face à une norme ou un standard, ou l'étude des systèmes et des structures qu'ils imposent représentent des pistes de recherche qui émergent autour du concept de variabilité. Avec le récent développement de la linguistique de corpus qui permet la description, l'analyse et la comparaison des systèmes linguistiques à grande échelle, les questions de changement, d'évolution, d'instabilité ou d'incertitude apparaissent également comme des enjeux scientifiques majeurs.

À un niveau plus méthodologique, c'est la manière de traiter variation et variabilité qui soulève des questionnements essentiels. On pourra par exemple s'interroger sur la sélection, voire la construction d'outils de mesure ou de cadres d'analyse pertinents, qu'ils soient qualitatifs ou quantitatifs. De nombreuses démarches entrent donc en résonance avec des problématiques liées à la variabilité : codages et transcriptions, traitement de corpus, entretiens et enquêtes de terrain, outils statistiques, logiciels et programmes, sélection et exploitation de variables...

Ainsi, dans l'optique de favoriser l'appropriation des thèmes de ce colloque par des doctorants et jeunes chercheurs issus de disciplines variées, les communications pourront porter sur l'ensemble des domaines suivants :

- **Phonétique, phonologie** : La notion de variation est évidemment pertinente en dialectologie à travers la comparaison des différentes variétés des langues à travers le monde. La variation intralinguistique quant à elle amène à analyser et à comparer les réalisations phonétiques au sein d'un même groupe d'individus en fonction de paramètres de type sociologique, phonotactique, intonatif, acoustique, etc. Ces analyses permettent de décrire les dynamiques internes aux langues et ainsi leur évolution.
- **Morphologie** : La variation s'étudie tant en morphologie dérivationnelle (inter- et intra-paradigme : variation de thèmes et de désinences, supplétion...) qu'en morphologie constructionnelle (outputs de mots construits ou schèmes de construction par exemple).
- **Sémantique** : La linguistique de corpus qui s'intéresse aux structures qui font sens dans la langue (du niveau lexical au niveau discursif) permet d'étudier les phénomènes de variation (entre termes, concepts, marqueurs ou structures syntaxiques, sémantiques et discursives) qui apparaissent par exemple dans les oppositions discours oral vs. écrit ou langue générale vs. de spécialité. La variation peut aussi être approchée en diachronie, au niveau du lexique (néologie, glissement de sens...) et à tous les niveaux de description linguistique.
- **Psycholinguistique ou neuropsycholinguistique** : L'étude de différents types de variabilité offre dans ces domaines la possibilité d'approcher les mécanismes de traitement du langage. On peut par exemple observer des différences de performances langagières chez un même sujet (variabilité inter-tâche) ou des différences d'activation cérébrale inter-sujets pour une même tâche (variabilité du substrat cérébral impliqué). Le profil linguistique du locuteur peut également constituer une donnée à prendre en compte pour expliquer des différences de performances dans la parole saine ou pathologique.
- **Didactique des Langues Étrangères** : La contextualisation de la situation didactique au niveau sociolinguistique, culturel, social, économique et politique permet l'adaptation à un niveau global et conduit à une multitude de "paysages didactiques". En outre, en didactique du FLE par exemple, la problématique de la norme à enseigner est toujours au cœur des réflexions.
- **Traitement Automatique des Langues (TAL)** : La prise en compte de la variabilité linguistique nécessite la mise en œuvre de diverses techniques

et ressources applicables à différents niveaux de traitement. Des applications comme la recherche d'information ou le résumé automatique sont ainsi directement confrontées au problème de la variation intervenant aussi bien aux niveaux morphologique et syntaxique que sémantique et amènent à l'utilisation de stratégies telles que la catégorisation ou l'exploitation d'ontologies.

- **Sémiotique, sémantique textuelle, analyse du discours** : Les disciplines du sens sont constamment confrontées à des questions de variabilité qui touchent à différents niveaux de grain de l'analyse textuelle. Au niveau le plus global, la contextualisation de tout acte de communication (conditions de production, construction discursive, conditions d'interprétation) peut constituer un travail capital pour informer une analyse. À un palier inférieur, les travaux d'inspiration bakhtinienne montrent à quel point la circulation discursive peut charger un même fragment textuel d'une multitude de sens et de voix. Enfin, les variations s'observent également au niveau de la microsémantique : tropes, diachronie, ou influence des genres poussent à interroger la stabilité des sèmes et des sèmes en fonction de l'incidence du contexte.
- **Sociolinguistique** : Des concepts comme ceux de frontières linguistiques, de communautés ou réseaux linguistiques, ou encore d'isoglosses montrent que la variation et la variabilité peuvent constituer un véritable objet d'étude en sociolinguistique, que ce soit au niveau des usages (discursifs, lexicaux, grammaticaux) ou des prononciations (réalisations phonétiques). Ces deux notions amènent également à mettre à l'épreuve certains modèles et protocoles méthodologiques tel que l'héritage labovien par exemple pour la sociolinguistique de l'anglais, voire pour la linguistique de corpus de manière plus générale.

Cette liste est évidemment non-exhaustive et toute proposition de communication se rapportant aux questions de la variation et de la variabilité dans les Sciences du Langage sera examinée.

Ces JéTou 2013 proposent ainsi une thématique particulièrement stimulante et ouverte que chaque courant des Sciences du Langage pourra s'approprier à sa manière. Ces journées seront alors l'occasion pour tous ceux qui le souhaitent de s'interroger, de débattre, et de confronter leurs travaux et leurs réflexions dans un cadre transdisciplinaire toujours propice à des échanges constructifs.

## Call for papers

The doctoral students in the Language Sciences laboratories at Toulouse University, France:

- CLLE–ERSS (Équipe de Recherche en Syntaxe et Sémantique)
- Laboratoire Octogone–Lordat (Centre Interdisciplinaire des Sciences du Langage et de la Cognition)
- LARA–CPST (Centre Pluridisciplinaire de Sémiotique Textuelle)

are organizing the fourth edition of *JéTou* (Journées d’Etudes Toulousaines), an international conference aiming at gathering doctoral students and young researchers (who have defended their dissertation within the past three years) together, from the different disciplines of the Language Sciences, on an open and multidisciplinary theme.

This 2013 edition will be devoted to a reflexion on the following theme: “Variation and Variability in the Language Sciences: analysing, measuring, contextualising”.

The notions of variation and variability are central to various studies in the Language Sciences. The linguist can investigate these notions as a research object or focus on the methodological issues these same notions raise.

Variation as a research object offers a wide range of possible reflexions. It can bring specific and fundamental perspective to the studies in the Language Sciences, through theoretical, historical or sociological contextualisation for instance. Besides, finding one’s positioning in relation to a norm, a standard or studying the constraints and structures that a norm, a standard can impose, are as many research directions linked to the concept of variability. With the recent development of corpus linguistics, which allows for the description, analysis and comparison of the linguistic systems of the different languages on a large scale, the issues of change, evolution, instability or uncertainty seem to be of crucial scientific importance.

On a more methodological level, taking variation and variability into account seems to be one of the essential issues for the Language Sciences. The linguist can ask himself whether he is selecting and using the best measuring tools or the accurate analytical frameworks (whether they are qualitative or quantitative). Many approaches and many protocols try to take the notion of variability into account, for example through codings and transcriptions, fieldwork investigations and interviews, statistical tools, softwares and computer programming, selection of variables...

Therefore, in order to encourage researchers in the different fields of the Language Sciences to make the theme of this 2013 edition of the *JéTou* their own, the contributions can focus on all the following issues:

- **Phonetics/phonology:** The notion of variation is relevant in phonetics and phonology through the comparison of the different varieties of language around the world. As far as inter and intra speaker variation is concerned, studying variation leads to comparing phonological and phonetic realizations within one specific group or several groups of individuals, based on sociological parameters, phonotactics, intonational contours, acoustics, etc. These potential analyses allow for the description of language dynamics and language evolution.
- **Morphology:** Variation can be studied both on the basis of derivational morphological phenomena (inter and intra paradigm: variation in stems and exponents, suppletion) as well as on the basis of constructional morphological phenomena (outputs of word-formation processes or constructional patterns...).
- **Semantics:** Corpus linguistics, which focuses on the study of the structures that have meaning in language (from the lexical level to the discursive level) enables researchers to study phenomena involving variation and variability (between words, concepts, markers or syntactic, semantic and discursive structures) that can emerge from the oral vs. written discourse dichotomy or the general language vs. language for specific purposes (LSP) dichotomy. Variation can also be tackled diachronically, as far as the lexicon is concerned for instance (neologisms, shifts in meaning...) or as far as any level of linguistic description is concerned.
- **Psycholinguistics/neuro-psycholinguistics:** Studying variability in these disciplines can often mean focusing on the language processing mechanisms. For example, researchers can observe differences in language performances in one subject (inter-task variability) or differences in cerebral activation in several subjects for the same task (variability in the cerebral substrate involved). The speakers' linguistic profiles can also be relevant data to take into account in order to explain differences in language performance between healthy speakers and speakers suffering from a pathological condition.
- **Language pedagogy:** Contextualising the didactic situation at the sociolinguistic, cultural, social, economic and political levels allows for a general adaptation to that specific situation and leads to a potential multiplicity of "didactic landscapes". Besides, as far as language pedagogy for the teaching

of the French language is concerned, a major issue remains the linguistic norm that should be taught and how to define it.

- **Natural language processing:** Taking variability into account in natural language processing and computational linguistics means using different techniques and resources that can be applied on distinct levels of processing. Such applications as information retrieval or automatic summarization have to deal with variation, whether it is found at the morphological, syntactic or semantic level, and therefore bring about the use of strategies known as categorisation or ontology engineering, among others.
- **Semiotics, textual semantics, discourse analysis:** These disciplines constantly have to deal with the notion of variability, which is linked to different levels of textual analysis. At the most comprehensive level, contextualising any communication act (conditions of production, discursive construction, conditions of interpretation) is of paramount importance in order to formulate and improve an analysis. At a lower level now, research inspired by Bakhtin's works has shown to what extent the circulation of discourses can load a multitude of meanings and voices onto any textual fragment. Finally, variation can also be observed at a microsemantic level: tropes, diachronic impact or the influence of genres compel the researcher to question the stability of lexical meaning according to the incidence of context.
- **Sociolinguistics:** Concepts such as linguistic boundaries, linguistic networks or communities as well as isoglosses show that variation and variability are central notions in sociolinguistics and often research topics for sociolinguists, as far as general usage is concerned (whether it is discursive, lexical, grammatical) or pronunciation (phonetic realizations). Both notions of variation and variability can also question some theoretical models or some methodological protocols, for instance William Labov's work, as far as the study of the English language is concerned, or even as far as corpus linguistics in general is concerned.

Obviously, this list does not include all the research possibilities following from the theme of the conference, which is why every submission that will address the notions of variation and variability in the Language Sciences will be read and reviewed.

JéTou 2013 offers a challenging theme that every field of the Language Sciences can make their own. This conference will give the opportunity, to those who wish to come, to ask questions, discuss and compare their work, their methods, their reflexions in an interdisciplinary context, which favours positive exchanges and constructive debates.

## Remerciements/Acknowledgements

Le comité d'organisation des JéTou remercie chaleureusement Françoise Gadet et Danièle Moore d'avoir accepté notre invitation, l'ensemble des membres du comité scientifique pour leurs relectures, ainsi que tous ceux qui ont apporté leur contribution à l'organisation de ces journées.

The organisation committee for JéTou wants to express its deepest gratitude to Françoise Gadet and Danièle Moore for accepting our invitation and being our two guest speakers, to all the members of the scientific committee for reading and correcting all the submissions and the articles for our conference, as well as to all those who have helped and contributed to the organisation of this event.

Partenaires des JéTou 2013/JéTou 2013 Sponsors:

- Laboratoire CLLE-ERSS et axes de recherche MorPhoLex, S'caladis et TAL,
- Unité de recherche interdisciplinaire Octogone,
- Département de Sciences du Langage de l'Université Toulouse II – Le Mirail,
- UFR de Langues, Littératures et Civilisations Étrangères de l'Université Toulouse II – Le Mirail,
- Université Toulouse II – Le Mirail,
- Conseil Régional de Midi-Pyrénées
- Agence Universitaire de la Francophonie,
- Mairie de Toulouse.



## Comité scientifique/Scientific committee

Mark ARONOFF (Stony Brook University, USA)  
Corine ASTÉSANO (Octogone–Lordat, Université de Toulouse, France)  
Michel AURNAGUE (CLLE–ERSS, Université de Toulouse, France)  
Michel BALLABRIGA (CPST, Université de Toulouse, France)  
Michel BILLIÈRES (Octogone–Lordat, Université de Toulouse, France)  
Gilles BOYÉ (CLLE–ERSSAB, Université de Bordeaux, France)  
Myriam BRAS (CLLE–ERSS, Université de Toulouse, France)  
Inès BRULARD (CLLE–ERSS, Université de Toulouse, France)  
Anne CONDAMINES (CLLE–ERSS, Université de Toulouse, France)  
Marie-Ange DAT (CREN, Université de Nantes, France)  
Jacques DURAND (CLLE–ERSS, Université de Toulouse, France)  
Carine DUTEIL-MOUGEL (ENSIL, Université de Limoges, France)  
Karine DUVIGNAU (CLLE–ERSS, Université de Toulouse, France)  
Cécile FABRE (CLLE–ERSS, Université de Toulouse, France)  
Françoise GADET (MoDyCo, Université Paris X, France)  
Pascal GAILLARD (Octogone–Lordat, Université de Toulouse, France)  
Claudine GARCIA-DEBANC (CLLE–ERSS, Université de Toulouse, France)  
Bruno GAUME (CLLE–ERSS, Université de Toulouse, France)  
Hélène GIRAUDO (CLLE–ERSS, Université de Toulouse, France)  
Nabil HATHOUT (CLLE–ERSS, Université de Toulouse, France)  
Lydia-Mai HO-DAC (CLLE–ERSS, Université de Toulouse, France)  
Daniel HUBER (CLLE–ERSS, Université de Toulouse, France)  
Marie-Paule JACQUES (LIDILEM, Université de Grenoble, France)  
Gerd JENDRASCHEK (Regensburg University, Germany)  
Mélanie JUCLA (Octogone–Lordat, Université de Toulouse, France)  
Barbara KÖPKE (Octogone–Lordat, Université de Toulouse, France)  
Anna KUPŚĆ (CLLE–ERSSaB, Université de Bordeaux, France)  
Laurence LABRUNE (CLLE–ERSSaB, Université de Bordeaux, France)  
Dany LAUR (CPST, Université de Toulouse, France)  
Anne LE DRAOULEC (CLLE–ERSS, Université de Toulouse, France)  
Michelle LECOLLE (CELTED, Université de Metz, France)  
Elise LEROY (CETIM, Université de Toulouse et Websourd, France)  
Vanda MARIJANOVIC (Octogone–Lordat, Université de Toulouse, France)  
Régis MISSIRE (CPST, Université de Toulouse, France)  
Fabio MONTERMINI (CLLE–ERSS, Université de Toulouse, France)  
Danièle MOORE (Simon Fraser University, Canada)  
Steven MOORE (CLLE–ERSS, Université de Toulouse, France)

Véronique MORICEAU (LIMSI, Paris, France)  
Patrick MPONDO-DICKA (CPST, Université de Toulouse, France)  
Philippe MULLER (IRIT, Université de Toulouse, France)  
Jean-Luc NESPOULOUS (Octogone–Lordat, Université de Toulouse, France)  
Musanji NGALASSO-MWATHA (CLLE–ERSS, Université de Toulouse, France)  
Marie-Paule PÉRY-WOODLEY (CLARE–CELFA, Université de Bordeaux, France)  
Aurélie PICTON (Université de Genève, Suisse)  
Laurent PRÉVOT (LPL, Université d’Aix-en-Provence, France)  
Josette REBEYROLLE (CLLE–ERSS, Université de Toulouse, France)  
Nathalie ROSSI-GENSANE (CLLE–ERSS, Université de Toulouse, France)  
Inès SADDOUR (Octogone–Lordat, Université de Toulouse, France)  
Halima SAHRAOUI (Octogone–Lordat, Université de Toulouse, France)  
Patrick SAUZET (CLLE–ERSS, Université de Toulouse, France)  
Catherine SCHNEDECKER (LILPA, Université de Strasbourg, France)  
Nathalie SPANGHERO-GAILLARD (Octogone–Lordat, Université de Toulouse,  
France)  
Ludovic TANGUY (CLLE–ERSS, Université de Toulouse, France)  
Jean-Michel TARRIER (CLLE–ERSS, Université de Toulouse, France)  
Olga THÉOPHANOUS (Octogone–Lordat, Université de Toulouse, France)  
Pascale VERGELY (SPH, Université de Bordeaux, France)  
Alessandro ZINNA (CPST, Université de Toulouse, France)

## Comité d'organisation/Organisation committee

**Caroline ATALLAH**

caroline.atallah@univ-tlse2.fr

CLLE-ERSS, Université Toulouse 2 – Le Mirail, France & IRIT, Université Toulouse 3 – Paul Sabatier, France

**Guillaume CARBOU**

guillaume.carbou@univ-tlse2.fr

CPST, Université Toulouse 2 – Le Mirail, France & LERASS, Université Toulouse 3 – Paul Sabatier, France

**Marie-Mandarine COLLE-QUESADA**

mcolle@univ-tlse2.fr

U.R.I Octogone-Lordat, Université Toulouse 2 – Le Mirail, France

**Claire DEL OLMO**

claire.delolmo@hotmail.fr

U.R.I Octogone-Lordat, Université Toulouse 2 – Le Mirail, France

**Marine LASSERRE**

marine.lasserre@univ-tlse2.fr

CLLE-ERSS, Université Toulouse 2 – Le Mirail, France

**Simon LEVA**

simon.leva@univ-tlse2.fr

CLLE-ERSS, Université Toulouse 2 – Le Mirail, France & IRIT, Université Toulouse 3 – Paul Sabatier, France

**Émilie MASSA**

emilie.massa@univ-tlse2.fr

U.R.I Octogone-Lordat, Université Toulouse 2 – Le Mirail, France

**Cécile VIOLLAIN**

cecile.viollain@univ-tlse2.fr

CLLE-ERSS, Université Toulouse 2 – Le Mirail, France

## Annuaire des participants/Directory of participants

### **Xavier BACH**

xavier.bach@oxon.org

Research Centre for Romance Linguistics, University of Oxford, United Kingdom

Thèse sous la direction de Mary Dalrymple et Martin Maiden, en préparation

*Morphologie, syntaxe, classes*

### **Monia BEN MLOUKA**

mlouka@irit.fr

Institut de Recherche en Informatique de Toulouse (IRIT), Université Toulouse III, France

Thèse sous la direction de Patrice Dalle, en préparation

*Analyse de la langue des signes*

### **Eva BOSCH I ROURA**

eva.bosch.roura@gmail.com

Departament de Filologia Catalana, Universitat de Barcelona, Spain

Thèse sous la direction de Maria-Rosa Lloret, en préparation

*Catalan, variation géolectale, phonologie*

### **Sabiha CHOURA**

sabihachoura@yahoo.fr

Research Unit in Discourse Analysis (GRAD), Faculty of Arts and Humanities of Sfax, Tunisia

Thèse sous la direction de Akilla Sellami-Baklouti

*Syntax, semantics, systemic functional linguistics*

### **Anne-Sophie DE CLERCQ**

anne-sophie.declercq@umons.ac.be

IRSTL, Université de Mons, Belgique

### **Claire DEL OLMO**

claire.delolmo@hotmail.fr

U.R.I Octogone-Lordat, Université Toulouse 2 – Le Mirail, France

Thèse sous la direction de Michel Billière, en préparation

*Français Langue Étrangère, cinéma, émotion*

**Louise ESHER**

University of Nottingham & Research Centre for Romance Linguistics, University of Oxford, United Kingdom

*Morphologie, linguistique historique, romanistique*

**Carmen Lucia FLOREZ PULIDO**

kappulletto@gmail.com

Université Paris-Descartes – école doctorale 180 & Laboratoire MoDyCo, Université Paris Ouest Nanterre la Défense, France

Thèse sous la direction de Laurent Danon-Boileau, en préparation

*La communication chez les enfants autistes*

**Françoise GADET**

MoDyCo & Université Paris Ouest Nanterre La Défense, France

Conférencière invitée

*Sociolinguistique, variation en français, style, contacts de langue, variétés « non légitimes » du français (français parlé, variétés non standards et non centrales du français, effets du contact des langues)*

**Nathalie GERBER**

nathalie.gerber@gmail.com

CREM, Université de Lorraine, France

Thèse sous la direction de Laurent Perrin, en préparation

*Enonciation, ethos, médias*

**Alexander GURYEV**

alexander.guryev@unine.ch

Institut des Sciences du langage et de la communication, Chaire de linguistique française, Université de Neuchâtel ED 268, Suisse & Université de Paris III Sorbonne Nouvelle, France

Thèse sous la direction de Marie-José Béguelin et Florence Lefeuvre, en préparation depuis mars 2012

*Structures interrogatives et variation syntaxique dans la communication par SMS*

**Katayoon KATOOZIAN**

kathy.katoozian@laposte.net

DILTEC (Didactique des langues, des textes et des cultures), Université Sorbonne Nouvelle – Paris 3, France

Thèse sous la direction de Danièle Manesse, en préparation

*Orthographe française, erreur, typologie*

**Marine LASSERRE**

marine.lasserre@univ-tlse2.fr

CLLE-ERSS, Université Toulouse 2 – Le Mirail, France

Thèse sous la direction de Fabio Montermini, en préparation depuis octobre 2011

*Morphologie constructionnelle, composition, lexique*

**Arthur MARY**

a.mary@gmx.com

Chercheur associé LIRCES, Université de Nice – Sophia Antipolis, France & ATER

auprès du LCPI, Université de Toulouse 2 – Le Mirail, France

Thèse sous la direction de Jacques Cabassut, soutenue le 19 janvier 2013

*Psychanalyse, discours sectaire, délire à deux*

**Noëlle MATHIS**

mathisnoelle@yahoo.fr

Université d'Avignon et des Pays de Vaucluse, France & Simon Fraser University, Canada

Thèse sous la direction de Claudine Moïse et Danièle Moore, en préparation

*Identités, plurilinguisme, littératies*

**Vanessa MEIRELES**

vanessameireless@gmail.com

Université Paris 8 UMR 7023 Structures Formelles du Langage, France

Thèse sous la direction de Joaquim Brandão de Carvalho et Jean-Louis Aroui, en préparation

*Sandhi vocalique, métrique et diphtongaison en portugais*

**Danièle MOORE**

Simon Fraser University, Burnaby, Canada

Conférencière invitée

*Sociolinguistique, didactique des langues, acquisition du langage, étude des plurilinguismes sociaux, analyse des interactions de classe, développement plurilittératié, acquisition des langues tierces dans différents contextes de contacts de langues et de cultures*

**Pierre-Antoine NAVARETTE**

pa\_navarette@yahoo.fr

CeReS, Université de Limoges, France

Thèse sous la direction de Jacques Fontanille et Nicolas Couégnas, soutenue le 12 janvier 2013

*Sémiotique de l'espace*

**Loredana RUCCELLA**

loryruc@hotmail.com

MoDyCo, Université de Paris X – Nanterre La Défense, France

Thèse sous la direction de J.F. Jeandillou et C. Marello, en préparation

*Euphémisme et politiquement correct*

**Nabil SADI**

sadinabil1@yahoo.fr

Laboratoire LAILEMM, Université de Bejaia, Algérie

Thèse sous la direction de Gudrun Ledegen et Abdenour Arezki, soutenue le 11 avril 2012

*Français parlé des médias*

**Holman TSE**

hbt3@pitt.edu

University of Pittsburgh, USA

Thèse sous la direction de Scott Kiesling, en préparation

*Sociolinguistics, variation, change*

**Charlotte VALLAT**

charlotte.vallat@hotmail.fr

U.R.I Octogone-Lordat, Université Toulouse 2 – Le Mirail, France

Thèse sous la direction de Michel Billière et Michel Grandaty, soutenue le 11 juin 2012

*Didactique de l'oral, stratégies enseignantes (étayage), contexte chinois*

**Géraldine ZUMWALD KÜSTER**

geraldine.zumwaldkuester@unifr.ch

Université de Fribourg, Suisse

Thèse sous la direction de Alain Berrendonner, en préparation

*Les pronoms clitiques et leur implication dans des phénomènes de variation*

## **Storying Research : Terrain(s), postures, interprétations autour du/des plurilinguisme(s)**

Danièle Moore<sup>1</sup>

(1) Simon Fraser University (Vancouver, Canada)

**Résumé** À partir de l'exemple d'une recherche de terrain en milieu autochtone canadien, dans un contexte de revitalisation des langues et cultures locales, cette conférence se propose de discuter comment l'étude des dynamiques complexes des plurilinguismes et de réappropriation des langues invite à réinterroger nos manières de construire la démarche de recherche, nos méthodologies, nos postures interactionnelles, nos interprétations, nos conceptualisations de l'individu-participant, de l'espace, du temps, tout comme nos modes de transmission du savoir académique (Wilson, 2008).

Nous discuterons l'idée que la narration de recherche participe d'un « *storywork* » au sens d'Archibald (2001), qui implique des formes particulières de co-engagement dans la recherche de terrain et ses mises en sens (Wilson, 2008) et questionne tout à la fois l'investissement et la responsabilisation du chercheur (Heller, 2002) et les effets socioculturels de sa recherche (Hendry & Fitznor, 2012). C'est alors sous l'angle du questionnement de la posture altéritaire du chercheur (de Robillard, 2008) que nous discuterons les tensions entre variation et contextualisation (des représentations et des pratiques).

### **Pistes bibliographiques**

ARCHIBALD J.-A. (Q'um Q'um Xiiem). (2008). *Indigenous Storywork, Educating the Heart, Mind, Body, and Spirit*. Vancouver : UBC Press.

DE ROBILLARD D. (2008). *Perspectives alterlinguistiques, volume 2 – Ornithorynque*. Paris : L'Harmattan.

HELLER M. (2002). *Éléments d'une sociolinguistique critique*. Paris : Didier.

HENDRY J. & FITZ NOR L. (2012) (eds). *Anthropologists, Indigenous Scholars and the Research Endeavour. Seeking Bridges Towards Mutual Respect*. New York : Routledge.

WILSON S. (2008). *Research is ceremony. Indigenous Research Methods*. Halifax and Winnipeg : Fernwood Publishing.

**« M e n j a z o v u t N a t i . I l o v e m a f a m i l l e , t s h o v r e b a , e t j ' e s p è r e q u e c ' e s t m u t u e l » : v a r i a t i o n s l a n g a g i è r e s m u l t i p l e s a u s e i n d ' u n a t e l i e r d ' é c r i t u r e p l u r i e l l e**

Noëlle Mathis<sup>1</sup>

(1) Université d'Avignon et des Pays de Vaucluse & Simon Fraser University  
mathisnoelle@yahoo.fr

**Résumé** Dans cet article, j'analyse les discours et représentations d'apprenants sur leurs pratiques de littératies plurilingues dans le contexte singulier d'un atelier d'écriture plurielle proposé dans une classe de FLE à l'université. Je tente notamment de mettre en exergue les (dis)continuités qu'ils instaurent dans l'atelier afin de comprendre comment ils s'autorisent (ou non) à exploiter les ressources multiples de leur répertoire plurilittératié et à construire leurs identités plurilingues.

**Abstract** In this article, I analyze learners' discourses on and representations of their multilingual literacy practices in a writing workshop taking place in a French as a Second Language university class. My goal is to highlight the (dis)continuity learners establish in this specific context as a way of understanding how they authorize themselves (or not) to use the multiple resources from their biliterate repertoire and to build plurilingual identities.

**Mots-clés :** Plurilinguisme, Variations, Identités, Littératies, Atelier d'écriture

**Keywords:** Plurilingualism, Variation, Identities, Literacy, Writing Workshop

## 1 Introduction

Qu'est-ce qui pourrait être plus facile à dire que quelques mots sur soi... ? C'est évident ! Mais, quand je commence à écrire: "je suis...", mon Babylon me fait perdre tous les mots. Les quatre personnes à l'intérieur de moi commencent à parler toutes ensemble dans une langue différente. Elles se bousculent en prétendant atteindre la page la première. So, finalement je dirais : M e n j a z o v u t Nati. I love ma famille, t s h o v r e b a, et j'espère que c'est mutuel. <sup>1</sup> Nati (« biographie », atelier 2009)

Mon territoire, mon pays a un circomstance (sic)<sup>2</sup> déterminante, sa frontière, la mer. Cette extension longue, d'eau infinie qui peut être par quelqu'un une prison mais qui est aussi et toujours une invitation au voyage, au départ.

- Vous êtes d'où?
- Majorque
- C'est où ça?

Bon, Il n'y a pas mal de personnes que ne connaissent pas du tout cette île et ceux qui la connaissent ils n'ont pas souvent une image très raffinée. Ce sont plutôt des clichés comme la paëlla, la corrida, etc. Alina (extrait de « Mon pays », atelier 2010)

Ces deux extraits écrits par des apprenants adultes plurilingues font partie de textes autobiographiques recueillis lors d'un atelier d'écriture plurielle que je propose en tant qu'enseignante et chercheuse dans ma classe de FLE dans un centre universitaire d'études françaises. Caractérisés par l'hybridité, le mélange de types et genres discursifs, l'alternance codique, l'enchâssement et le dialogisme, ils questionnent le contexte de production dans lequel ils ont été écrits. C'est pourquoi, mon étude se centrera particulièrement sur le contexte singulier que représente l'atelier d'écriture à travers l'analyse des discours des apprenants sur leurs pratiques de littératies plurilingues. Les entretiens menés avec les apprenants à propos de leurs rapports à l'écrit en fonction de ce contexte permettront de voir comment ils s'autorisent (ou non) à exploiter les ressources de leur répertoire plurilittératié, à avoir un discours sur celles-ci et à manifester leurs identités plurilingues. Sur le plan de la réflexion didactique se posera la question de la meilleure prise en compte par les enseignants et les chercheurs des contextes d'appropriation plurilittératiés, notamment dans le cadre de l'enseignement universitaire. A l'instar de Bemporad et Moore (sous presse), je tente de saisir comment les apprenants font des liens entre leurs expériences de langues, leurs pratiques de littératies et leur construction identitaire tout en apprenant la langue cible. Je fais l'hypothèse que le contexte d'appropriation ouvert à des pratiques de littératies se construisant sur des continua de possibles aide les apprenants à (re)construire leurs identités plurilingues.

---

<sup>1</sup> Ce texte, la biographie de Nati (un pseudonyme), est présenté tel qu'elle l'a proposé dans une version corrigée. Les espaces entre les lettres des expressions M e n j a z o v u t (je m'appelle) et t s h o v r e b a (mot qui fait référence à l'histoire et à la rédemption dans l'histoire médiévale géorgienne) sont ceux choisis par Nati.

<sup>2</sup> Le texte de l'atelier 2010 est présenté en version non corrigée par l'enseignante, non seulement parce qu'il est disponible, mais aussi pour être au plus proche de l'interlangue de l'apprenante.

*« M e n j a z o v u t Nati. I love ma famille, t s h o v r e b a, et j'espère que c'est mutuel » : variations langagières multiples au sein d'un atelier d'écriture plurielle*

## **2 Cadre théorique**

### **2.1 Littératie(s) et pratiques de littératies**

Le terme de littératie(s) se comprend comme étant « l'ensemble des activités humaines qui impliquent l'usage de l'écriture en réception et en production » (Jaffré, 2004 : 31). Je l'emploie au pluriel pour mettre l'emphase particulière sur le fait que la littératie est variable selon les contextes sociaux, les groupes culturels et les individus (Street, 2000). De plus, je considère les littératies comme des pratiques sociales (Barton, Hamilton 2010 ; Cope, Kalantzis, 2000) construites sur des bases épistémologiques qui considèrent que les manières dont les individus approchent et pratiquent l'écrit (et la lecture) sont elles-mêmes basées sur des conceptions en liens étroits avec les savoirs, l'identité et la posture des individus impliqués dans ces pratiques. Par extension, les pratiques de littératie(s) (Dagenais, 2012 ; Dagenais, Moore, 2008) désignent à la fois les pratiques sociales et les conceptions liées à la lecture et à l'écriture, surtout en lien avec la notion de pouvoir. C'est pourquoi, j'emploie le terme « pratiques de littératies » car il permet de mettre en lien les pratiques des apprenants, leurs discours sur leurs pratiques dans le contexte d'appropriation singulier que représente l'atelier d'écriture plurielle et la manifestation de leurs identités.

L'atelier d'écriture plurielle se situe au cœur d'une didactique du plurilinguisme (Moore, 2006) qui favorise l'émergence, la valorisation et l'utilisation d'autres langues que la langue cible, le FLE, et ce dans un but, à la fois, d'apprentissage de la langue et de développement d'une compétence plurilingue et pluriculturelle (Coste, Moore, Zarate, 1997 / 2009). Ces langues et leurs pratiques mixtes et hétérogènes constituent le répertoire plurilingue des apprenants et font partie intégrante des interactions de classe. Par ailleurs, les apprenants mobilisent les ressources variées de leur répertoire lors de pratiques de littératies plurilingues (Martin-Jones, Jones, 2000). Ces auteures proposent, en effet, que les pratiques d'écriture et de lecture, l'utilisation de certains genres, styles et types de textes varient en fonction du contexte singulier et situé dans lequel elles ont lieu. Le locuteur-scripteur plurilingue aura la possibilité dans des circonstances favorables de s'emparer de matériaux de langue(s), de les façonner, de les transformer et de les recréer de manière complexe, unique et personnelle. Dagenais et Moore (2008), quant à elles, proposent la notion de « répertoire plurilittératié », un néologisme construit sur le modèle de répertoire plurilingue et en lien avec la compétence plurilingue et pluriculturelle (Coste, Moore, Zarate, 1997/2009). Cette notion prend en compte les façons complexes et multiples qu'une personne plurilingue possède pour donner du sens à travers ses pratiques langagières. En lien plus particulièrement avec les pratiques d'écriture, et en s'inspirant de Hornberger (1999) dont le modèle de continua de bilittératie est évoqué plus loin, Moore (2006 : 118) propose le néologisme de pratiques plurilittératiées pour « définir toute instance au cours de laquelle la communication se construit dans plus de deux langues dans et autour de l'écrit ». Elle considère ainsi que le répertoire des apprenants plurilingues est composé de compétences dans plusieurs littératies, et ce dans plus d'une langue. Moore (2012 : 63) précise que le terme plurilittératié indique que les langues et les formes de littératies se développent les unes par rapport aux autres en relation aux contextes dans et à travers lesquels elles peuvent émerger tout en contribuant à l'apprentissage et à la construction identitaire.

### **2.2 Pratiques plurilittératiées, construction identitaire et contextes**

Je cherche à faire le lien entre des pratiques plurilittératiées et la construction identitaire d'apprenants adultes plurilingues. Quand je fais référence au concept complexe de l'identité,

j'adopte la notion d'identités plurilingues de Moore et Brohy (2013) qu'elles définissent de la manière suivante : « les identités plurilingues et pluriculturelles sont une catégorie de l'identité (individuelle et/ou collective). Elles s'expriment au travers de l'usage que fait un locuteur de ses langues, et au travers de ses discours sur celles-ci. Elles sont marquées par l'instabilité et l'ambivalence. Elles se révèlent de manière différente selon les choix des locuteurs à l'intérieur d'un ensemble de possibles, choix qui dépendent, entre autres, des trajectoires de vie individuelle, des catégories sociales et linguistiques à disposition et des interprétations que donne l'individu aux circonstances locales dans lesquelles il est amené à négocier la différence et à signaler ses affiliations » (Moore, Brohy, 2013 : 297). Plus particulièrement, je tente d'établir le lien entre pratiques de littératies et identités en fonction du contexte d'appropriation en me référant à Porquier et Py (2004). Dans cette étude, le *macro*-contexte correspondrait aux nombreux niveaux que sont l'idéologie monolingue véhiculée par la société française, les recommandations du Cadre Européen Commun de Références pour les Langues (CECRL), et l'interprétation de ces textes par le centre de langue (et ses enseignants) dans lequel l'atelier d'écriture a lieu. Le *micro*-contexte, quant à lui, serait l'atelier, dans lequel les pratiques de littératies sont proposées en fonction de mes propres représentations et celles des apprenants de ce que constituent la(es) langue(s) et les littératies, en lien ou non à d'autres langues et d'autres pratiques de littératies. Dialogisme, polyphonie, intertextualité, emprunt, transferts de matériel linguistique étant des composants essentiels des textes de base étudiés en classe, ils se projettent dans les interactions et se retrouvent dans les textes des participants à l'atelier. Ces derniers, le *micro-micro*-contexte, sont constitués des énoncés des apprenants que je considère comme étant fondamentaux dans cette étude car ils constituent le pivot pour l'analyse de la construction *in situ* des identités.

De nombreux sociolinguistes s'intéressant aux questions d'éducation (voir Moore 2006) ont souligné les (dis)continuités et les interdépendances entre les *micro* et *macro*-contextes. Bemporad (2012) propose que ces (dis)continuités conditionnent les comportements, les pratiques et les identités des locuteurs. Dans cette étude, je propose d'examiner les discours tenus par les locuteurs sur leurs pratiques de littératies : ils établissent un lien (ou non) avec les langues de leur répertoire et les autres pratiques et contextes d'écriture que ceux proposés en atelier. Considérés comme des acteurs sociaux et créatifs, aptes à répondre à leurs contextes sociaux et situés, ils posent des « actes d'identité » (LePage, Tabouret-Keller, 1985). Par ces « actes », à travers leurs pratiques, ils révèlent leurs adhésions et allégeances (ou non) à certaines normes langagières et, ce faisant, expriment leur voix singulière.

### 2.3 L'atelier d'écriture plurielle

L'atelier se caractérise par certaines (dis)continuité(s) entre les contextes *micro* et *macro* qui se manifestent par les tensions suivantes : d'un point de vue *micro*, lors d'interactions en classe, je propose aux participants d'utiliser les langues de leur répertoire plurilingue alors que nous sommes dans un cours de FLE, dans lequel nombreux sont les étudiants qui estimerait que la langue à utiliser ne devrait être que la langue cible, le français. De plus, je leur propose d'écrire des textes autobiographiques en lien avec leurs langues, leur trajectoire de mobilité, leur rapport à l'altérité, dans un but de placer leur expérience personnelle au cœur du projet éducatif, ce qui pourrait être considéré comme un obstacle pour certains apprenants, du fait que ces activités langagières se situent dans un cadre universitaire. Enfin, et surtout en ce qui concerne les récits écrits, je permets, et j'encourage même, que l'écriture en français soit accompagnée d'autres langues qui font partie du répertoire plurilingue des apprenants. Ces activités langagières entrent en tension avec la norme, la représentation circulante. D'un point de vue *macro*, l'atelier d'écriture se déroule dans un centre

« *M e n j a z o v u t Nati. I love ma famille, t s h o v r e b a, et j'espère que c'est mutuel* » : variations langagières multiples au sein d'un atelier d'écriture plurielle

universitaire de langue française dans lequel la didactique du plurilinguisme n'est encouragée ni implicitement ni explicitement. De plus, les textes proposés dans le cadre de production écrite consistent pour la plupart en des textes considérés comme étant plus formels (lettres, essais argumentatifs).

### 3 Participants et méthodologie

Les participants à l'atelier sont des adultes plurilingues provenant de pays divers tels que la Géorgie, l'Espagne, les Etats-Unis. Ils sont inscrits dans un programme intensif de quinze heures hebdomadaires sur une durée de quatorze semaines avec l'objectif de perfectionner leur français dans un but personnel et / ou professionnel. L'atelier d'écriture représente trois heures hebdomadaires de ce programme. L'un a eu lieu en 2009, l'autre en 2010. Toutes les séances ont été enregistrées et transcrites. Les données analysées toutefois, dans le cadre de cet article, sont issues d'entrevues collectives effectuées en fin de trimestre avec les étudiants suite à leur participation à l'atelier. Pendant ces entretiens semi-dirigés, les apprenants ont un discours sur leurs pratiques de littératies en lien (ou non) avec leurs langues et sur le contexte d'appropriation. Les entretiens sont accompagnés de textes produits par les apprenants lors des ateliers d'écriture.

Dans cet article, à l'instar de Bemporad et Moore (sous presse), j'interprète les discours des apprenants en m'inspirant du modèle de « continua de bilittératie » de Hornberger (1989). Ce modèle s'articule autour de continua inter-reliés et imbriqués qui interagissent sur l'apprentissage et les pratiques de littératies des apprenants plurilingues. Cette notion de continuum présente également dans les travaux francophones (Lüdi, Py, 2003 ; Porquier, Py, 2004 ; Moore, 2006) consiste à souligner une continuité entre des pôles qui sembleraient être opposés mais qui, dans la pratique, sont inextricablement imbriqués. Hornberger (1989) et Hornberger et Skilton-Sylvester (2010) identifient notamment des pôles d'interrelations dynamiques et contextuellement évolutives : entre L1 et L2 (ou Ln), entre l'écrit et l'oral, entre les contextes *micro* et *macro*, entre des communautés bilingues et monolingues, entre des contenus contextualisés et décontextualisés (qui ne font pas référence aux contextes dans et grâce auxquels ils ont été produits). Deux approches d'analyse seront utilisées, l'une thématique puisqu'il s'agit pour les apprenants de réfléchir à leurs pratiques de plurilittératies dans un contexte particulier. L'autre porte sur la forme même de la langue, c'est à dire la manière dont ils parlent de leurs pratiques de plurilittératies en lien (ou non) avec leurs langues et dans d'autres contextes. Dans les deux cas, l'analyse sera faite en fonction des pôles d'interrelations précédemment cités et en mettant en exergue des « actes d'identités » (LePage et Tabouret-Keller, 1985), et notamment d'identités plurilingues.

## 4 Le discours des apprenants

### 4.1 Nati : « les langues que l'on parle c'est nous »

Le premier cas concerne une apprenante géorgienne, Nati<sup>3</sup>, plurilingue en géorgien, russe, anglais et français, qui exprime son rapport aux langues dans l'atelier :

moi je pense que cet atelier d'écriture c'est très spécial parce nous on a écrit on a parlé nous par rapport à nos langues mais pas les langues par rapport à nous . les langues qu'on parle c'est nous comment on est dans le monde des langues et comme on est tous

---

<sup>3</sup> Tous les prénoms donnés sont des pseudonymes

comme soit deux langues soit trois langues soit polyglotte c'est très intéressant . c'est c'est expérience unique je veux dire parce nous par exemple si tu habites en Russie et si tu parles russe uniquement tu n'aurais pas besoin d'écrire des choses comme ça (entretien 2009)

Nati exprime que l'atelier d'écriture lui a procuré un contexte d'expression singulier (*c'est très spécial*). Elle en cite la cause : des pratiques de l'oral et de l'écrit qu'elle met sur le même pied d'égalité (*on a écrit on a parlé*). Elle utilise la forme plurielle du pronom personnel sujet (*nous*) pour parler de son rapport aux langues. Le pronom *nous* s'interprète de plusieurs manières. Il peut s'agir d'elle-même faisant partie du groupe d'apprenants ; d'elle-même et des apprenants avec lesquels elle a effectué l'entretien ; d'elle-même en tant que Géorgienne ayant habité en Russie, car plus loin elle énonce *parce que nous par exemple si tu habites en Russie*. Dans tous les cas, ce *nous* représente une hétérogénéité énonciative qui exige de l'interlocuteur une participation empathique à la production de sens. L'atelier lui aurait permis de se mettre, en tant que sujet plurilingue, au cœur du projet d'écriture et non l'inverse : *nous par rapport à nos langues mais pas les langues par rapport à nous*. Cette construction syntaxique singulière forme une sorte de miroir en plaçant d'abord le pronom personnel *nous* (A) en début de proposition suivi de la locution de comparaison *par rapport à* (B) et du syntagme *les langues* (C) puis établit une opposition en utilisant le connecteur *mais* suivi de la négation *pas* puis *les langues* (C) *par rapport à* (B) *nous* (A), soit une construction de type A + B + C mais pas C + B + A. De ce fait, dans l'usage de langue, Nati instaure un pôle d'interaction dont les extrémités seraient d'un côté l'individu plurilingue et de l'autre les situations de multilinguisme<sup>4</sup> et utilise ce pôle pour identifier le contexte de l'atelier comme étant centré sur l'individu bi-/plurilingue. De toute évidence, Nati s'identifie aux langues qu'elle personnifie (*les langues qu'on parle c'est nous / comment on est dans le monde des langues*), et ce en parlant d'elle-même en référence à un individu inscrit dans un collectif (utilisation importante de *nous*, *on*, *nos*). Ce *nous* se transforme en *tous* qui, ici je l'interprète comme étant les apprenants plurilingues de l'atelier d'écriture qu'elle qualifie comme étant *deux langues* ou *trois langues* et *polyglotte*. On notera l'étrange que procure les termes *deux langues* ou *trois langues* qu'elle utilise comme s'il s'agissait d'adjectifs, à la place de « bilingues » et « trilingues », et ce, notamment en apposition avec le terme *polyglotte*. Enfin, Nati évoque un autre contexte géographique (*la Russie*) et le locuteur monolingue (*si tu parles russe uniquement*) en mentionnant que l'écriture de récits concernant son rapport aux langues n'y aurait pas de raison d'être. De ce fait, elle met en exergue une autre forme de (dis)continuité entre un contexte *micro* plurilingue (*l'atelier d'écriture*) et un contexte *macro* qu'elle considère monolingue (*la Russie*), l'un centré sur les individus plurilingues (*polyglotte*) et l'autre sur des individus monolingues (*si tu parles russe uniquement*).

En résumé, dans ce récit, Nati instaure des continua à plusieurs niveaux que ce soit entre l'écrit et l'oral, entre des contextes *micro* et *macro*, entre des individus bilingues et monolingues. Elle introduit également une dialectique entre soi et l'autre, un soi qui se construit par rapport à un collectif. Ses identités plurilingues se manifestent dans un rapport aux langues possible dans l'espace singulier de l'atelier d'écriture. Hétérogènes, elles s'expriment de manière dialogique en prenant en considération des espaces linguistiques et géographiques divers. On retrouve également ces attributs d'hétérogénéité dans son texte

<sup>4</sup> Notons que pour Moore (2006), le terme *multilinguisme* caractérise une société ou une communauté dans laquelle plusieurs langues coexistent alors que le *plurilinguisme* se rapporte à l'individu en tant que locuteur, acteur social, apprenant et médiateur de plusieurs langues.

« *M e n j a z o v u t N a t i . I l o v e m a f a m i l l e , t s h o v r e b a , e t j ' e s p è r e q u e c ' e s t m u t u e l* » : variations langagières multiples au sein d'un atelier d'écriture plurielle

présenté en introduction : personnification des langues, utilisation de l'alternance codique, enchevêtrement des langues.

## 4.2 Monica: « je suis comme elle »

Monica est une apprenante catalane plurilingue (catalan, castillan, anglais, français) d'une trentaine d'années. Jeune docteure en biologie, elle suit ce cours afin d'augmenter ses chances de trouver un emploi dans un centre de recherche en France. A la question de savoir si l'atelier d'écriture lui a permis de découvrir quelque chose sur elle-même, Monica répond :

oui oui oui parce que avec les lectures et l'écriture les deux choses euh on a un avis un reflex de nous de nous de nous-mêmes on dit ça ? [mm] et parce que quand comme comme a dit T. c'est de l'expérience des autres que peut-être tu je n'avais pensé jamais que que c'est ma c'est cela est ma situation je n'avais pas pensé jamais et après de lire et de voir et tu dis oui oui oui c'est comme moi [c'est comme moi] je suis comme je suis comme elle et tu peux réfléchir (entretien 2010)

Monica exprime clairement son approbation avec l'usage de trois *oui*. Elle précise que ce sont les activités de lecture et d'écriture qui ont contribué à cette découverte (*avec les lectures et l'écriture*). Elle insiste qu'il s'agit bien des deux activités (*les deux choses*) et établit de la sorte une continuité entre des activités de littérature qui concernent à la fois la lecture et l'écriture. On notera qu'elle fait d'abord référence au groupe en utilisant le pronom « nous » (*un reflex de nous*) en faisant l'usage d'une bifocalisation (Bange, 1992) : elle réfléchit à l'expérience du groupe tout en ayant une activité métalinguistique au cours de laquelle elle vérifie l'utilisation de l'adverbe « même » (*de nous de nous-mêmes on dit ça ?*). L'expérience d'autrui lui permet de découvrir des éléments d'elle-même à travers une réflexivité poussée qui se manifeste par plusieurs boucles réflexives (Authier-Revuz, 1995) : *comme a dit T. c'est de l'expérience des autres / c'est ma situation / c'est comme moi*. Ces boucles réflexives ont pour fonction de commenter ses propres dires en introduisant une notion de dialogisme. De plus, cela semble être une surprise pour Monica : elle répète à deux reprises *je n'avais pensé jamais / je n'avais pas pensé jamais* tout en essayant de s'auto-corriger sur l'utilisation de la particule négative *jamais* (autre bifocalisation). Par ailleurs, elle utilise le deuxième pronom personnel dans son processus de réflexivité pour établir de la distance par rapport à elle-même : *tu dis oui oui oui c'est comme moi tu peux réfléchir*. De ce fait, elle utilise une notion de miroir (*un reflex* qui signifie ici un reflet, *réfléchir* qui ici est à prendre dans les deux sens du terme, celui du reflet et celui de la réflexion, *je suis comme elle* je ressemble à l'autre). Cette utilisation peut s'interpréter comme une dialectique entre soi et l'autre, un continuum entre le « je » singulier (et pluriel) et le « tu » qui permet au « je » d'être. Ces deux procédés, la bifocalisation et l'utilisation de boucles énonciatives, à travers une activité métalinguistique intense, semblent essentiels à Monica pour parler de son rapport aux langues et pour exprimer ainsi son identité plurilingue.

Lorsque je lui demande de préciser le texte auquel elle fait référence, Monica répond :

oui je pense que c'est le premier texte de Nancy Huston que il exprime différents euh anecdotes sur sa vie et sa vie des langues et quand il quand il mélange les langues quand il pense avec je pense que c'est les gros mots par exemple qu'il peut dire elle peut dire des gros mots en français mais n'est pas en anglais et pour moi et pour la majorité de nous je pense que c'est c'est vrai tu peux dire des gros mots par exemple en français mais que tu ne diras jamais dans ta langue maternelle (entretien 2010)

L'apprenante exprime sans aucune hésitation la source et l'auteure du texte auquel elle fait référence<sup>5</sup>. Elle décrit graduellement le contenu du texte en forme de zoom en allant du plus large au plus précis (*le premier texte, différents anecdotes sur sa vie, sa vie des langues, il mélange les langues, les gros mots, les gros mots en français, pas en anglais*), ce qui montre que son activité de réflexivité se manifeste au cours de la narration d'un événement important pour elle. On constate que l'événement langagier proposé par Huston, à savoir le mélange des langues et des registres de langue, rejoint directement l'expérience de l'apprenante : *et pour moi et pour la majorité de nous je pense que c'est c'est vrai*. Monica fait référence à elle-même (*pour moi*) ainsi qu'aux membres du groupe (*pour la majorité de nous*) avec beaucoup de fluidité. Sa construction de soi se fait dans la confrontation à l'altérité, dans le balancement entre « je », « nous » et « tu ». De manière très claire, elle exprime une expérience commune : *tu peux dire des gros mots par exemple en français mais que tu ne diras jamais dans ta langue maternelle*. Le passage au pronom personnel *tu* à valeur générique lui permet d'établir une distance réflexive en se donnant la possibilité de s'insérer elle en tant que locutrice plurilingue dans l'énoncé. L'événement langagier est de toute importance et il s'agit de l'utilisation d'un certain registre -ici, la vulgarité- dans la langue cible mais pas dans la langue maternelle.

En résumé, de ces deux extraits, on notera que Monica instaure une continuité entre lecture et écriture, et entre les usages multiples des langues qu'utilise l'auteur et les siens à l'oral. La dialectique entre soi et l'autre établie lui permet d'envisager la variation linguistique notamment en ce qui concerne le registre familier voire vulgaire entre deux langues de son répertoire plurilingue. Monica, en ayant de nombreuses activités métalinguistiques, adopte une énonciation singulière pour se dire en tant que locutrice plurilingue. Ses identités plurilingues se traduisent par une forte réflexivité et une hétérogénéité dans son discours ainsi qu'un rapport étroit et nécessaire à l'altérité qui miroite, conforte et construit.

### 4.3 Alina: « oublier (...) ta façon d'écrire dans ta langue maternelle »

Alina, une Majorquine près de la trentaine, est plurilingue (catalan, castillan, français). Elle étudie le français pour faire un Master en France. Alina fait une observation en fin d'entretien en ce qui concerne ses pratiques de littératie en français :

moi ce que je faisais au départ d'abord c'était de faire comme ehm une traduction comme essayer de faire une traduction de mes pensées en espagnol de mettre en papier et : moi je m'ai rendu compte que c'était pas ça le processus qui convient c'est de oublier complètement ta façon de d'écrire dans ta langue maternelle et de plonger dans dans les structures de de la langue que que tu es en train d'apprendre et ça moi je commence maintenant à le faire mais moi je trouve que c'est le plus lourd de et et le plus oui et et . il faut du temps je crois je trouve (entretien 2010)

<sup>5</sup> Nous avons étudié plusieurs extraits de *Nord Perdu* de Nancy Huston qui ont fait l'objet d'échanges vifs. Nous avons notamment discuté l'usage (ou non) pour le locuteur plurilingue de mots vulgaires dans sa langue maternelle et les équivalents dans des langues d'apprentissage. Huston raconte une anecdote d'une amie écossaise devenue corse qui ne peut exprimer de grossièretés dans sa langue maternelle alors qu'en français, elle le fait sans états d'âme : « quand je suis vraiment hors de moi, en français les pires obscénités franchissent mes lèvres sans problème. Dire *putain, salope, enculé de ta mère*, ça ne me fait ni chaud ni froid » (*Nord perdu*, p. 63). Huston déclare elle-même qu'au début de son séjour en France, « la langue française était par rapport à [sa] langue maternelle, moins chargée d'affect » (*Nord Perdu*, p. 64), et lui donnait une grande liberté dans l'écriture.

« *M e n j a z o v u t Nati. I love ma famille, t s h o v r e b a, et j'espère que c'est mutuel* » : variations langagières multiples au sein d'un atelier d'écriture plurielle

Dans cet extrait, Alina partage ses stratégies autour de l'écriture en français. La dimension temporelle est importante, car elle explique ce qu'elle faisait *d'abord* (une traduction de ses pensées en espagnol), puis le moment d'une réalisation (*je m'ai rendu compte que c'était pas ça*) et l'expression fluide d'un constat qu'elle énonce comme une vérité : *le processus qui convient c'est de oublier complètement ta façon d'écrire dans ta langue maternelle*. Elle précise que cela implique qu'il faille adopter les structures de la langue cible, et c'est ce qu'elle pratique *maintenant*. Elle énonce donc un changement, une transformation dans son identité linguistique avec l'adoption de nouvelles stratégies linguistiques. Ricoeur (1990) en proposant la notion d'« identité narrative » vise justement à répondre aux problématiques existentielles de transformation face à la discontinuité à travers le temps tout en assurant une certaine cohérence ou continuité. Et cette continuité se présente par une « mise en intrigue », qui ici est structurée selon un schéma classique: situation initiale, problème, résolution, situation finale. Ce récit est important sur deux points : premièrement, elle énonce une discontinuité entre l'écriture en français et l'utilisation de l'espagnol qu'elle considère comme un obstacle, discontinuité qu'elle ressent comme étant difficile (*c'est le plus lourd*). Deuxièmement, elle construit, dans et à travers son récit, une structure narrative, qui non seulement lui permet une continuité entre deux moments de son identité linguistique, mais qui propose également une cohérence avec les propos même de son récit (*plonger dans les structures de la langue*).

En résumé, Alina instaure une discontinuité entre ses langues dans ses pratiques de littératies. Cependant, elle utilise une énonciation singulière qui se manifeste de manière dialogique à la fois dans son récit sur ses pratiques et dans l'extrait de texte proposé en introduction. En effet, dans ce dernier, on constate enchâssements, multiplicité de points de vue, alternance de types discursifs, discours direct en l'absence de rection et brouillage énonciatif qui requièrent la connivence du lecteur. Si Alina ne s'autorise pas à utiliser plusieurs langues de son répertoire plurilingue dans un même texte, par contraste avec Nati qui utilise l'alternance codique et l'enchevêtrement de langues, on peut dire que ses identités plurilingues se manifestent par l'hybridité textuelle et le dialogisme.

## **5 Pour conclure: regard critique sur l'atelier d'écriture plurielle**

Hornberger et Skilton-Sylvester (2010) insistent sur les enjeux de pouvoir inhérents au modèle des continua de bilittératie. Elles précisent qu'au regard des pratiques de littératie(s) « there tends to be an implicit privileging of one end of the continua over the other such that one end of each continuum is associated with more power than the other » (: 98). Elles soulignent que les pratiques de littératies proposées en salle de classe privilégient l'écriture de textes standardisés, décontextualisés, monolingues et dont les contenus sont issus de groupes majoritaires. Or, précisément, leur objectif est de montrer qu'une telle situation peut être transformée grâce à des réflexions critiques sur ses pratiques par les enseignants. En effet, la salle de classe peut devenir un lieu où la diversité linguistique est conçue comme un atout d'apprentissage (Moore, 2006) et un contexte singulier qui permet la voix, une voix singulière, à ceux et celles qui ne l'ont pas (ou pas encore), et / ou qui ne valorisent pas leurs langues et leurs compétences (même partielles) comme étant des tremplins à la construction de leurs identités. A cet effet, je constate que l'atelier d'écriture plurielle apporte les éléments de réflexion suivants: premièrement, les propositions d'écriture privilégient l'écriture de textes personnels, contextualisés dans l'expérience de chacun à travers la migration et l'apprentissage des langues. Deuxièmement, les textes écrits par les apprenants sont produits suite à des discussions orales ayant pour propos la compréhension des apprenants à des textes d'auteurs étudiés en classe. Troisièmement, les textes d'auteurs sont non seulement écrits par

des écrivains dont le français est la langue d'écriture, mais pas forcément celle d'origine, et aussi dont les thématiques sont en lien même avec les langues, le sentiment d'étrangéité et la rencontre de l'autre. Ainsi, selon le modèle écologique de bilittératie, les pratiques de littératies en atelier d'écriture plurielle se situeraient sur les continua plutôt vers les extrémités traditionnellement dotées de moins de pouvoir, à savoir des contextes d'écriture plutôt oraux et plurilingues, où l'expérience personnelle est primordiale, et où les contenus des textes étudiés reflètent le parcours des apprenants. Selon Hornberger (1989), l'apprenant exposé à des contextes d'apprentissage variés sera en mesure de développer ses compétences plurilittératiées. Je proposerais ainsi que plus les contextes d'apprentissage permettent aux apprenants d'exercer des pratiques plurilittératiées sur tous les points des continua de bilittératie, et notamment sur ceux qui ont eu traditionnellement moins de pouvoir, plus les possibilités d'expression hybride et hétérogène d'identités plurilingues seront présentes pour les apprenants.

## Références

- AUTHIER-REVUZ J. (1995). *Ces mots qui ne vont pas de soi. Boucles réflexives et non-coïncidences du dire*. Deux volumes, Paris, Larousse.
- BARTON D., HAMILTON M. (2010). La littératie: une pratique sociale. *Langage et société*, 133, 45-62.
- BANGE P. (1992). A propos de la communication et de l'apprentissage de L2 (notamment dans ses formes institutionnelles). *AILE, 1*, en ligne dans <http://aile.revues.org/4875>.
- BEMPORAD C. (2012). *Lectures littéraires d'étudiant.e.s plurilingues et appropriations langagières*. Thèse de doctorat, Université de Lausanne.
- BEMPORAD C., MOORE D. (sous presse). Identités plurilingues, pratiques (pluri)littératiées et apprentissages. *Bulletin Vals/Asla*.
- COPE B., KALANTZIS M. (2000). *Multiliteracies : Literacy Learning and the Design of Social Futures*. New York, Routledge.
- COSTE D., MOORE D. et ZARATE G. (1997/2009). *La compétence plurilingue et pluriculturelle*. Strasbourg, Editions du Conseil de l'Europe.
- DAGENAIS D. (2012). Littératies multimodales et perspectives critiques. *Recherches en didactique des langues et des cultures: Les Cahiers de l'Acedle*, 9(2), 15-46.
- DAGENAIS D., MOORE D. (2008). Représentations des littératies plurilingues, de l'immersion en français et des dynamiques identitaires chez des parents chinois. *Canadian Modern Language Review*, 65(1), 11-31.
- HORNBERGER N. (1989). Continua of biliteracy. *Review of Educational Research*, 59, 271-296.
- HORNBERGER N., SKILTON-SYLVESTER E. (2000). Revisiting the Continua of Biliteracy : International and Critical Perspectives. *Language and Education*, 14 :2, 96-122.
- HUSTON N. (1999). *Nord Perdu*. Arles : Actes Sud/Leméac.

« *M e n j a z o v u t Nati. I love ma famille, t s h o v r e b a, et j'espère que c'est mutuel* » : variations langagières multiples au sein d'un atelier d'écriture plurielle

JAFFRE J.-P. (2004). La littéracie : histoire d'un mot, effets d'un concept. In Barré-De Miniac C., Brissot C., Rispaïl M. (dir.). *La littéracie, conceptions théoriques et pratiques d'enseignement de la lecture-écriture*. Paris : L'Harmattan, 21-41.

LE PAGE R.B., TABOURET-KELLER A. (1985). *Acts of identity : Creole-based approaches to language and ethnicity*. Cambridge, London : Cambridge University Press.

LÜDI G., PY B. (2003). *Etre bilingue*. Bern, Peter Lang.

MARTIN-JONES M., JONES K. (2000). *Multilingual Literacies*. Philadelphia, PA: John Benjamins.

MOORE D. (2006). *Plurilinguismes et école*. Paris, Didier, Collection LAL.

MOORE D. (2012). Pratiques plurigraphiées d'enfants chinois en immersion française à Vancouver. *Le français dans le monde, Recherches et applications*, 51, 9-14.

MOORE D., BROHY C. (2013). Identités Plurilingues. In J. Simonin et S. Wharton (éds), *Sociolinguistique du contact : dictionnaire des termes et concepts*. Lyon : ENS-Éditions.

PORQUIER R., PY B. (2004). *Apprentissage d'une langue étrangère : contextes et discours*. Paris, Didier.

RICOEUR P. (1990). *Soi même comme un autre*. Paris, Seuil, Points Essais.

STREET B. (2000). Literacy events and literacy practices : theory and practice in the New Literacy Studies. In M. Martin-Jones & K. Jones (Eds.). *Multilingual Literacies*. Philadelphia, PA: John Benjamins.

## **Une nécessaire contextualisation : le cas d'une étude de la stratégie enseignante d'étayage, en classe de FLE, en milieu universitaire chinois**

Charlotte Vallat<sup>1</sup>

(1) Octogone-Lordat & Université de Toulouse  
charlotte.vallat@hotmail.fr

**Résumé** L'étude présentée porte sur des analyses de pratiques de classe, mises en place dans le cadre d'une expérience contrôlée en milieu universitaire chinois. Nous nous penchons sur la manière dont les enseignants de FLE guident leurs apprenants lors de séances de cours d'oral et sur la nature de leur étayage. L'objectif est de mettre en valeur les diverses stratégies d'étayage utilisées par les enseignants et voir dans quelle mesure elles favorisent la compréhension et la production orales chez les apprenants. Dans une perspective de regards croisés, nous mettons en parallèle deux fois quatre cours (mêmes supports, même nombre d'apprenants, même niveau linguistique) dispensés par deux enseignantes (de culture, nationalité et formation différentes : chinoise et québécoise). La mise en parallèle de pratiques enseignantes permet de nous interroger sur l'efficacité de ces techniques d'enseignement, notamment au regard du contexte d'apprentissage bien spécifique qu'est le contexte chinois. Les résultats obtenus confirment que la prise en compte du contexte est cruciale, c'est la raison pour laquelle la recherche s'inscrit et prône une approche contextualisée de l'enseignement/apprentissage des langues.

**Abstract** This study deals with the analyses of teaching practices during a controlled experiment in a Chinese university. We focus on the way French as a Foreign Language (FFL) teachers guide their students during oral sessions as well as on the nature of their scaffolding. Our aim is to enhance the main strategies used by the teachers and to see how these strategies improve the students' oral comprehension and oral production. We focused on two groups of students with a different teacher for each group (one conducted by a Chinese teacher and the other conducted by a French Canadian teacher). We observed four classes with each of the two groups. With the aid of the analysis grid, we assess the way in which FFL teachers guide their students during oral sessions in the context of a Chinese university. Our results confirm that it is crucial to take the context into account. This is the reason why our research lies within, and advocates, a contextualized approach to modern languages teaching/learning.

**Mots-clés :** Enseignement de FLE en université chinoise, Contextualisation, Culture éducative, Etayage, Pratiques didactiques, Perspective interactionniste

**Keywords:** Teaching of French as a foreign language (FFL) in Chinese universities, Contextualisation, Educative culture, Scaffolding, Didactics practices, Interactionnist perspective

## 1 Introduction

Travailler sur un terrain particulier, c'est prendre en compte le contexte de ce terrain. Analyser un terrain particulier, avec un public d'apprenants spécifiques, c'est prendre en compte le contexte dans lequel a lieu l'apprentissage. Un terrain d'analyse est défini par des caractéristiques qui lui sont propres, par des individus vivant dans une Culture et marqués par une culture éducative et des traditions culturelles et sociales (Castellotti & Moore, 2008 ; Blanchet & Chardenet, 2011). Selon Castellotti & Moore (2008), les notions de situation puis de contexte, occupent une place centrale dans les recherches actuelles en didactique des langues: si elles ne sont pas toujours visibles, ces notions réapparaissent périodiquement.

Notre étude se déroule en Chine, en contexte universitaire chinois. L'expérience contrôlée mise en place au sein de l'Université Normale de Chine du Sud (Canton, Guangdong) et les résultats obtenus sont à étudier et à regarder en fonction du contexte général chinois, du contexte d'apprentissage et de la culture éducative des apprenants.

## 2 Expérimentation en milieu universitaire chinois

### 2.1 Prise en compte de la culture éducative chinoise

Comprendre la problématique actuelle de la présente recherche, nécessite de la replacer dans son contexte spécifique qui est l'enseignement/apprentissage du FLE en milieu chinois. En effet, les codes et héritages culturels d'une société ont vraisemblablement toujours un impact sur le système éducatif en place (Bouvier, 2003). Un système de vie communautaire, des principes du confucianisme toujours bien présents, une certaine conformité, collaboration et recherche du consensus, la notion de face, les habitudes d'apprentissage, sont autant de paramètres à prendre en compte lors d'une étude en contexte chinois (Robert, 2002, 2009 ; Normand-Marconnet, 2011). Le public chinois est marqué par une culture éducative traditionnelle où les habitudes d'apprentissage d'une langue étrangère sont calquées sur l'apprentissage de la langue maternelle. Les poids historique, culturel et institutionnel sont prégnants, tant dans l'enseignement de la langue étrangère (programmes officiels souvent imposés, domination de l'écrit sur l'oral, etc.) que dans la formation des enseignants (espace discursif largement monopolisé par l'enseignant, une grammaire explicite, des exercices de réemploi, etc.).

En Chine, le modèle d'enseignement qui est généralisé depuis de nombreuses années et jusqu'aux années 2000 environ, est qualifié de "traditionnel" dans la mesure où celui-ci est centré sur la place prépondérante et quasi-sacrée accordée au professeur. Il transmet son savoir au travers d'une relation hiérarchique et frontale avec l'apprenant. Ce dernier étant considéré comme passif et simplement censé reproduire un savoir défini et statique. Ce modèle entraîne une absence d'autonomie des apprenants mais, selon Normand-Marconnet (2011) ce modèle est bien accepté car il correspond à la conception traditionnelle chinoise de la relation professeur-élève. La prise de parole spontanée est un véritable défi pour la plupart des étudiants car le questionnement direct ne fait pas partie de leur bagage pédagogique. Cela pourrait entraîner une "perte de face" qui serait susceptible d'entraîner une coupure pratiquement irréversible du lien social qui unit ces deux personnes. En effet, la notion de face est un élément à prendre en compte, dès lors que l'on se trouve en présence d'un public asiatique. D'après Cicurel (2003), cet héritage culturel, fondé sur les principes du confucianisme et du taoïsme amène les étudiants chinois à concevoir leurs professeurs comme les vecteurs privilégiés et incontestables de la transmission des acquisitions du passé. Le proverbe chinois "*Professeur pour un jour, père pour la vie*" (Robert, 2009, p.114), prend ici

tout son sens. Les Chinois pratiquent la reconnaissance et le respect des degrés de hiérarchisation. De ce fait, il paraît nécessaire de replacer notre étude dans son contexte pour ainsi mettre en valeur, la place fondamentale accordée au modèle parfait à reproduire, dans la civilisation chinoise, mais aussi cette culture éducative permettant de comprendre, en termes d'enseignement/apprentissage, nombres de comportements des apprenants mais également des enseignants, en classe. Certains codes culturels, propres au public d'apprenants chinois, permettent de cerner les implications dans le domaine de l'acquisition/apprentissage en classe de langue.

Mais, il n'est pas question ici d'enfermer l'étudiant et l'enseignant chinois dans un moule particulier, strict et immuable, ni d'apporter un modèle stéréotypé de ceux-ci, mais bien d'apporter quelques informations (tant d'ordre culturel que pédagogique) qui peuvent permettre d'éclairer certains comportements assez récurrents en Chine. La Chine change à grande vitesse et certaines attitudes et valeurs sont en voie de disparition, au profit d'une occidentalisation ou mondialisation accélérée. Ainsi, les éléments proposés ne reflètent que des tendances ; tendances sujettes à une transformation rapide. En effet, les traditions éducatives et culturelles, doivent être prises pour ce qu'elles sont : des constructions. Il ne s'agit pas forcément d'une vérité applicable de manière « aveugle » à tous les contextes et apprenants. Ces traditions ne doivent pas être vues comme figées, ni comme des traits culturels absolus, mais davantage comme des tendances qui évoluent et qui ne sont, bien entendu, pas forcément généralisables à l'ensemble des apprenants.

## **2.2 Analyse des pratiques enseignantes à travers l'étayage de l'enseignant**

Dans ce contexte bien particulier qu'est le milieu universitaire chinois, sont observées et analysées les pratiques de classes de deux enseignantes, c'est-à-dire, les techniques utilisées par celles-ci dans leur cours. Ces observations ont lieu, en 2010, à l'Université Normale de Chine du Sud de Canton, dans la province du Guangdong. Pour ce faire, nous nous penchons, plus particulièrement, sur la manière dont ces enseignantes de FLE guident leurs apprenants lors de séances de cours d'oral et sur la nature de leur étayage.

L'étayage (Bruner, 1983 ; Grandaty & Chemla, 2004 ; Bucheton, 2009) correspond à l'ensemble des interventions que l'enseignant accorde à ses apprenants ainsi que tous ses processus d'ajustement. La notion d'étayage (*scaffolding*) introduite par Wood, Bruner & Ross (1976) puis, plus largement développée par Bruner (1983), correspond concrètement à l'aide apportée par un adulte expert, à une autre personne moins experte, afin de lui permettre d'accomplir une tâche qu'elle n'aurait pu effectuer toute seule. Dans le cadre de la classe, le déroulement du cours et les objectifs pointés sont « pensés » par l'enseignant (Courty, 2003). Une fois qu'un contrat didactique (Brousseau, 1989) est établi entre les divers participants à l'interaction, les stratégies d'étayage de l'enseignant peuvent se déployer. L'apprentissage dépend en partie du type d'assistance fourni par l'enseignant ou interagissant le plus compétent. La fonction du tuteur est « *d'assurer l'apprentissage en faisant agir* » (Bange, Carol & Griggs, 2005, p.53). L'activité développée par le tuteur dans l'étayage doit être propre à développer l'autonomie de l'apprenant. L'étayage de l'enseignant est un geste professionnel (Bucheton, 2009). Les gestes professionnels sont des gestes de métier, partagés soit de manière explicite, soit implicite par les membres d'une communauté professionnelle. Ces gestes de métier ont besoin de « formats » (Bruner, 1983) d'étayage qui sont identifiés et reconnus par l'ensemble des élèves. En milieu scolaire, selon Bucheton (2009, p.59), l'étayage « *c'est ce que l'enseignant fait avec son élève pour l'accompagner dans ses*

*apprentissages et dans la mise en place de conduites et attitudes qui leur sont propices. C'est l'intervention du maître dans un espace d'apprentissage que l'élève ne peut mener seul ».*

A l'inverse, il peut arriver que l'enseignant aille à l'opposé de l'étayage dans la mesure où, par exemple, il ne laisse pas de place à la pensée des élèves, il donne la réponse trop rapidement, sans laisser assez de temps aux apprenants pour y réfléchir, coupe la parole ou encore travaille, pendant que les élèves passifs attendent. On qualifie ces situations de « sur-étayage » (Bucheton, 2009) ou encore de « contre-étayage » (Nonnon, 2001 in Grandaty & Chemla, 2004). Pour notre étude, nous retenons la notion de « contre-étayage ».

L'objectif de l'expérience présentée est de mettre en valeur les diverses stratégies d'étayage utilisées par les deux enseignantes observées et voir dans quelle mesure elles favorisent la compréhension et la production orales chez les apprenants.

### **2.3 Méthodologie utilisée**

L'étude se propose d'observer l'enseignement du français oral en milieu universitaire chinois, plus particulièrement au sein de l'Université Normale de Chine du Sud. Le travail consiste à mettre en évidence et à analyser les diverses formes d'étayage à travers les interactions enseignants/apprenants échangées en classe. L'objectif de la recherche est d'analyser les stratégies d'étayage développées par les enseignantes observées et voir quels types de réalisations sont suscités chez leurs apprenants. Dans ce but, diverses données sont récoltées, directement sur le terrain, dans cette université, en Chine. Dans une perspective de regards croisés, nous mettons en parallèle deux fois quatre cours (mêmes supports utilisés, même nombre d'apprenants, même niveau linguistique) effectués par ces deux enseignantes de FLE, de profils différents (de culture, nationalité et formation différentes : chinoise et québécoise). Nous les nommons Professeur A (PA) et Professeur B (PB)<sup>1</sup>.

L'observation est contrôlée dans la mesure où nous avons pris en compte de nombreuses variables<sup>2</sup>. Les observations sont au nombre de huit: quatre cours dispensés par chacune des enseignantes. Les quatre cours sont effectués, chacun, avec le même document vidéo, proposé par nous-même, mais également avec le même nombre d'apprenants ayant le même niveau A2/B1 du CECRL<sup>3</sup>, niveau attesté par des examens. Nous avons également mené des entretiens généraux avec les membres de l'équipe enseignante et des entretiens pré et post cours avec les deux enseignantes observées et nous avons analysé leurs préparations de cours (à l'aide de leurs documents de planification enseignante).

La méthodologie d'observation se déroule en trois temps. Elle permet de considérer la classe dans son intégralité et dans toute sa complexité en partant de la planification du cours (constitution du scénario pédagogique) que l'enseignant compte mettre en place tout au long du cours), en passant par son déroulement (l'agir de l'enseignant dans sa classe) pour en définitive effectuer un retour, une réflexion métacognitive, quant au contenu de celui-ci

---

<sup>1</sup> Professeur A : d'origine chinoise, moins de 30 ans, formation en Chine et en France, obtention d'un master II Sciences du langage en France, expérience dans l'enseignement du FLE de plus de deux ans.

Professeur B : d'origine francophone, moins de 30 ans, formation au Québec, obtention d'une maîtrise de Littérature, expérience dans l'enseignement du FLE de plus d'un an.

<sup>2</sup> Contrôle de diverses variables: même nombre d'apprenants par classe, même niveau des apprenants, attesté par des examens, utilisation du même document audiovisuel dans la classe.

<sup>3</sup> CECRL: Cadre Européen Commun de Référence pour les Langues, Conseil de l'Europe, 2001.

(Blanchard-Laville & Nadot, 2004). Il ne s'agit pas ici de comparer les stratégies d'étayage utilisées par les deux enseignants mais bien d'avoir un panel plus large des stratégies qui peuvent être mises en œuvre dans un cours d'oral.

Une fois les films de classes retranscrits, à l'aide de la convention de transcription ICOR<sup>4</sup>, nous travaillons sur un corpus de huit classes ; corpus analysé à l'aide d'une grille d'observation et d'analyse de l'étayage de l'enseignant. Cette grille est conçue à partir de diverses disciplines contributives, pour les besoins de la recherche (Vallat, 2012). De part la spécificité du corpus (cours de compréhension et de production orales, de niveau A2/B1, objectifs spécifiques à atteindre pour chaque cours), la grille est constituée à partir de différentes données, issues de diverses disciplines telles que les sciences du langage, la psychologie, la pédagogie et la didactique (Bruner, 1983; Crahay, 1999; Grandaty & Turco, 2001; Colletta, 2004; Grandaty & Chemla, 2004; Bange, Carol & Griggs, 2005; Bucheton, 2009). Cette grille présente un aspect novateur car elle envisage l'interaction dans sa globalité et son intégralité (prise en compte des paramètres verbaux, para-verbaux et non verbaux, notamment les gestes co-verbaux<sup>5</sup> de Colletta, 2004). De nos jours, peu de travaux intègrent la totalité de ces données.

D'un autre côté, les caractéristiques des réalisations des apprenants, par rapport aux diverses fonctions d'étayage activées par l'enseignant sont décrites, grâce à une analyse de discours. Chaque tour de parole des apprenants est analysé et caractérisé : diverses sortes de prises de parole sont définies. On peut trouver des réponses sous forme de mots et d'expressions, sous forme d'énoncés contenant un verbe, des réponses simples, des non réponses, des répétitions et des prises de parole spontanées.

La méthodologie d'observation mise en place dans les classes de langues, permet d'observer les diverses stratégies d'étayage assumées par les deux enseignantes observées (grâce aux enregistrements vidéo), de les identifier (grâce à la grille conçue à cet effet), de les interpréter (grâce à l'entretien pré et post cours effectué avec les enseignants relevant une certaine métacognition de leurs pratiques) et de les évaluer (par l'intermédiaire des caractéristiques des réalisations des apprenants).

### **3 Une cruciale prise en compte du contexte d'enseignement/apprentissage**

#### **3.1 Résultats de l'étude**

Tout au long des quatre cours, il est à remarquer que les deux enseignantes qui ont participé à l'expérience contrôlée, ont une attitude et une posture énonciative différente. Il semble que PA, tout en adoptant une posture de sur-énonciateur (Rabatel, 2007) soit un locuteur très actif : elle parle beaucoup et même parfois trop, dans la classe. PB assume aussi une posture

---

<sup>4</sup> Convention ICOR du laboratoire ICAR, Université Lyon 2.

<sup>5</sup> Les gestes co-verbaux sont des gestes qui accompagnent les productions langagières. Ils sont de trois types : les co-verbaux référentiels, qui permettent de représenter des référents concrets ou abstraits c'est-à-dire des gestes de type iconique ou métaphorique (parmi lesquels les gestes déictiques, illustratifs...), les co-verbaux de synchronisation (gestes rythmiques par exemple qui permettent de scander le flux de parole) et les co-verbaux expressifs (gestes qui redoublent, renforcent la parole par exemple).

de sur-énonciateur mais adopte davantage, à certains moments, une posture d'effacement énonciatif (Rabatel, 2004; Grandaty & Chemla, 2004; Bucheton, 2009). Dans cette posture, le professeur se met en retrait de manière à laisser les apprenants interagir. Ceci constitue une stratégie d'enseignement mais aussi une stratégie d'apprentissage, permettant au locuteur de donner l'impression qu'il se retire de l'énonciation. Ici, l'enseignante décentre sa position pour favoriser les interactions horizontales entre élèves et permet ainsi aux apprenants de s'exprimer davantage. Les enseignantes assument toutes deux diverses fonctions d'étayage, à des pourcentages relativement proches. Comme tout enseignant de langue, habituellement grand "gesticulateur", les deux professeurs observées usent énormément de gestes co-verbaux (Colletta, 2004). Les modalités d'étayage sont aussi gestuelles: les enseignantes étayent par leurs paroles mais leurs interventions peuvent également être plus discrètes pour ne pas briser l'échange.

L'analyse du corpus met en évidence la présence de contre-étayages (Grandaty & Chemla, 2004; Bucheton, 2009) effectués par les deux enseignantes et plus particulièrement par le PA. Les contre-étayages prennent la forme d'une absence de consigne, de répétitions directes des extraits de la vidéo (directement après le visionnage), le fait de donner soi-même la solution directement sans attendre la réponse des élèves et le fait de couper la parole aux apprenants.

Il semblerait qu'une partie de ces contre-étayages soient assez caractéristiques de certains aspects de la méthodologie traditionnelle, méthode dite « grammaire-traduction », dans la mesure où, par exemple, la répétition est vue comme une technique d'apprentissage et où l'espace discursif est largement monopolisé par l'enseignant.

Il est intéressant de remarquer qu'en prenant seulement en compte les productions des apprenants sous forme d'énoncés contenant un verbe, soit les productions les plus élaborées (réponses aux propos de leur enseignant ou prise de parole spontanée), de manière globale sur l'ensemble des quatre cours, les apprenants du PA réalisent autant, voire davantage de productions sous forme d'énoncés contenant un verbe que les apprenants du PB. Ceci peut paraître surprenant car le PA effectue de nombreux contre-étayages. Pourtant, d'après les résultats obtenus, cela ne nuit pas à la production des apprenants, bien au contraire.

Ces résultats soulèvent, par conséquent, des questions de méthodologies mises en place dans les classes, et par la même de formation des enseignants de FLE. Rappelons que le PA est partie se former en France en didactique du FLE et notamment aux approches qui ont de plus en plus cours aujourd'hui: les Approches communicatives. De retour en Chine, elle met en pratique cette nouvelle méthodologie. En tout cas, elle tente de le faire. En faisant ces quelques contre-étayages, nous faisons l'hypothèse qu'elle revient peut être, sous certains aspects, à une méthode plus traditionnelle, en rapport avec sa propre expérience de l'apprentissage du français et en prenant en compte les habitudes d'apprentissage des apprenants chinois, qu'elle connaît bien étant elle-même chinoise. Tout enseignant a une conception plus ou moins implicite de l'apprentissage. Cette conception s'est forgée petit à petit par son expérience, la formation qu'il a reçue, ses lectures et ses rencontres pédagogiques (Courtyllon, 2003). La pratique enseignante, qui découle de cette conception, peut donc varier d'un enseignant à l'autre.

La méthode d'enseignement utilisée par le PA semble s'être enrichie de diverses sources: à la fois une approche assez traditionnelle de l'enseignement du français mais également une méthodologie davantage communicative (méthodologie de plus en plus prônée dans certains établissements de langues en Chine). Ainsi, la présente analyse de l'étayage des enseignants, en classe d'oral, en FLE et les résultats obtenus, nous amènent à réfléchir sur la formation des

*Une nécessaire contextualisation : le cas d'une étude de la stratégie enseignante d'étayage, en classe de FLE, en milieu universitaire chinois*

enseignants de FLE en milieu universitaire chinois et met en évidence la nécessité, finalement, de songer à concilier plusieurs méthodologies, plutôt que de vouloir en imposer une seule. C'est la raison pour laquelle nous pensons que les quelques contre-étayages que le PA effectuée ne nuisent pas aux réalisations des apprenants chinois; contre-étayages qui devraient habituellement engendrer des productions moins élaborées chez un autre public d'apprenants étrangers (Nonnon, 2001 in Grandaty & Chemla, 2004; Bucheton, 2009).

Cette analyse amène à réfléchir plus largement sur la formation des enseignants de FLE en milieu universitaire chinois. La présente étude est une étude de cas (un moment particulier, un public donné, dans une seule université en Chine) et les résultats obtenus et les interprétations ne sont pas généralisables à tous les établissements d'enseignement de langues en Chine, ni à l'ensemble des enseignants. Néanmoins, même si cette recherche est une étude d'un cas bien particulier, elle peut être un socle à partir duquel réfléchir et les pistes et propositions qui suivent peuvent venir enrichir les discussions et réflexions déjà en cours depuis quelques années, en Chine. La présente étude s'inscrit dans une recherche praxéologique dont les visées sont, entre autres, d'apporter des propositions didactiques dans le domaine de l'enseignement du FLE, notamment en milieu universitaire chinois.

### **3.2 Quelle didactique du FLE aujourd'hui en Chine ?**

Dans de nombreux établissements de langues en Chine, particulièrement en FLE, la méthodologie d'enseignement utilisée est une méthode assez traditionnelle, qui tente peu à peu de se renouveler. Selon Hu (2011, p.117), *"la méthodologie traditionnelle continue à dominer dans la majorité des classes de français. L'approche communicative reste un décor facultatif de par la relative faible importance accordée à l'interaction, au sens et à l'autonomie"*. Pourtant, depuis quelques années de nouvelles méthodologies d'enseignement apparaissent et essayent de se développer en Chine. En effet, avec l'ouverture sur le monde extérieur et la venue d'enseignants étrangers en Chine (avec de nouvelles méthodologies et manuels de langue), les traditions évoluent. Les Approches communicatives, la perspective actionnelle ou encore l'éclectisme, tentent de s'imposer. Cette évolution peut poser des problèmes dans un pays comme la Chine, où les traditions sont encore bien ancrées. Même si une évolution dans ce domaine est réellement souhaitée par certaines autorités locales, le poids de l'institution en Chine, est tel que le changement est lent.

La didactique des langues en Chine ne paraît pas encore être une discipline reconnue. Même si différentes méthodologies constituées exogènes ont été introduites en Chine (Méthode Audio-visuelle, méthode Structuro Globale Audio-visuelle, etc.), aucune systématisation ne s'est véritablement produite dans le contexte chinois. Les méthodologies sont souvent introduites sans une réelle prise en compte du contexte dans lequel elles sont élaborées, parfois en se superposant, voire même en se contredisant, ce qui explique, probablement, leur manque d'efficacité pour un public chinois. Une application trop brutale d'une méthodologie, dans n'importe quel contexte que ce soit, ne peut rendre service ni aux enseignants, ni aux apprenants. En revanche, il semble que la construction d'une méthodologie, peut partir de bases exogènes pour se constituer. Selon Yang (2007, p.71) *"si l'éclectisme a permis à l'enseignement du français en Chine de sortir de l'impasse du «monopole» d'un quelconque courant méthodologique, dorénavant ce seraient les besoins concrets issus d'une réalité de plus en plus diversifiée qui comptent dans le choix méthodologique pour émettre toute hypothèse didactique."* Une méthodologie doit respecter le contexte d'enseignement/apprentissage dans lequel elle voit le jour, ainsi que la culture d'apprentissage des apprenants. Le riche passé éducatif chinois doit être pris en compte, incluant un certain

savoir-apprendre les langues, que la didactique du français L2 en Chine ne doit pas ignorer. Selon Martin (2007), à travers le dialogue des idées (qui caractérise actuellement nos deux pays), toute construction d'une méthodologie d'enseignement/apprentissage du FLE en Chine ne peut qu'être éclectique dans un premier temps. Néanmoins, il semble que cet éclectisme devrait être raisonné et pensé ensemble.

Ainsi, une réflexion est aujourd'hui amorcée sur la didactique du FLE en Chine. Cette réflexion, en pleine émulation, est le fruit de recherches, de recherches-action, de réflexions multiples qui aboutissent à des pistes, des recommandations tournant autour des notions d'"éclectisme raisonné", de « compromis didactique » (Robert, 2009), de « didactique du compromis » (Cuet, 2008), de « sociodidactique de terrain » (Rispaïl & Blanchet, 2011) ou encore de "pluralisme didactique" (Huver, 2012). Ces concepts mettant bien en évidence une nécessaire contextualisation des recherches. Le présent travail s'inscrit, par conséquent, dans une approche contextualisée (Castellotti & Moore, 2002, 2008; Debono, 2011; Blanchet & Chardenet, 2011 ; Rispaïl & Blanchet, 2011 ; Huver, 2012) de l'enseignement/apprentissage des langues. En effet, la prise en compte des traditions éducatives et culturelles, des représentations et des habitudes d'apprentissage des apprenants présente un intérêt certain pour l'élaboration d'une didactique contextualisée. Par exemple, il est vrai qu'*"au-delà d'approches génériques des situations d'appropriation des langues, les représentations de leur apprentissage semblent de manière unanime, quels que soit les publics concernés et les contextes considérés, très fortement imprégnées du rapport à la langue première et de son rôle, perçu le plus souvent comme prédominant, dans l'accès à d'autres langues"* (Castellotti & Moore, 2002, p.16). Cependant, rappelons que ces traditions éducatives et culturelles doivent être considérées comme ce qu'elles sont réellement: des objets en construction, des mises en regard, permettant de réfléchir à ses pratiques en contexte. Cette approche doit, tout de même, être utilisée avec précaution car selon Debono (2011, p.134) *"pour intéressants et nécessaires qu'ils soient, le maniement de ces travaux de recherche n'en est pas moins délicat: la réflexion sur l'étrangeté des traditions culturelles, éducatives ou autres, est un exercice de funambule, le risque étant de verser dans une réification culturaliste, débouchant sur un déterminisme aliénant"*. Ces traditions ne doivent pas être vues comme figées, ni comme des traits culturels déterminants ou absolus, mais plutôt telles des tendances évolutives, non généralisables à l'ensemble des apprenants.

De ce fait, la question de la formation des enseignants découle de cette approche contextualisée. Huver (2012) parle de diversité formative. Elle souligne le fait que *"la diversité -linguistique, culturelle, formative- semble désormais reconnue comme un phénomène constitutif de l'enseignement/apprentissage des langues"* (Huver, 2012, p.1). D'après Galisson (1999, p. 65 in Huver 2012, p.2) la formation des enseignants reste (sauf quelques rares exceptions) *"confondue avec la formation aux méthodologies et aux méthodes dominantes"*. Ainsi, les enseignants sont formés aux diverses méthodologies mais les formations se fondent encore trop souvent sur un principe de diffusion universel qui ne prendrait pas forcément en compte la complexité et la diversité des expériences et des traditions d'enseignement/apprentissage ou encore des appropriations, interprétations et transformations de ces principes par les acteurs.

*Une nécessaire contextualisation : le cas d'une étude de la stratégie enseignante d'étayage, en classe de FLE, en milieu universitaire chinois*

Concernant la diversité formative, Huver (2012) insiste sur le fait que les représentations sont tenaces quant à l'idée qu'il existerait une seule "bonne voie didactique" alors que ce présupposé même contredit l'idée d'explorer la voie d'un pluralisme didactique<sup>6</sup>.

La réflexion actuellement en cours met très clairement en avant les deux penchants de la didactique du FLE aujourd'hui en Chine: une fondamentale prise en compte de la culture éducative chinoise et par la même, de la « méthode chinoise » (Besse, 2011) ainsi que l'application raisonnée de nouvelles approches davantage communicatives. En effet, pour proposer des pistes pour une didactique du FLE en Chine, il est absolument fondamental de connaître, de comprendre, de respecter les cultures éducatives et spécificités du public d'apprenants et de trouver des passerelles entre didactiques différentes.

## 4 Conclusion

Ainsi, les résultats obtenus concernant les caractéristiques des réalisations des apprenants du PA, ont un impact sur le contexte d'étude chinois. Les résultats ne sont certes pas généralisables à l'ensemble des universités chinoises, mais peuvent constituer une base à partir duquel réfléchir; tant pour la formation des enseignants de FLE en Chine qu'au niveau des méthodologies à mettre en place dans les classes. Ce travail confirme ainsi l'intérêt des recherches en contexte, directement sur le terrain et valorise, par la même, une approche contextualisée (Castellotti & Moore, 2002; Debono, 2011; Blanchet & Chardenet, 2011 ; Huver, 2012).

## Références

BANGE P., CAROL R., GRIGGS P. (2005). *L'apprentissage d'une langue étrangère: Cognition et interaction*. Paris : l'Harmattan.

BESSE H. (2011). Un point de vue sur l'enseignement du français en Chine. *Synergie Chine*, 6, 249-260.

BLANCHARD-LAVILLE C., NADOT S. (2004). Analyse de pratiques et professionnalisation entre affect et représentation. *Connexions*, 82, 119-142.

BLANCHET P., CHARDENET P. (dir.) (2011). *Guide pour la recherche en didactique des langues et des cultures. Approches contextualisées*. Paris: Éditions des archives contemporaines.

BOUVIER B. (2003). Chinois et Français : quand les habitudes culturelles d'apprentissage s'opposent. *Etudes de linguistique appliquée*, 132, 399-414.

BROUSSEAU G. (1989). Utilité et intérêt de la didactique pour un professeur de collègue. *Petit x*, 21, IREM de Grenoble.

BRUNER J. (1983). *Le développement de l'enfant : savoir faire, savoir dire*. Paris : PUF.

---

<sup>6</sup> Huver (2012, p.6) note que ce point de vue se distingue de la notion d'éclectisme telle que développée par Puren (1993), en ce que le pluralisme est ici à entendre comme l'expression politique de la diversité, qu'il suppose une conscience et une réflexivité des acteurs sur leurs pratiques (et leurs représentations) et qu'il s'appuie (notamment) sur une historicisation de celles-ci.

BUCHETON D. (2009). *L'agir enseignant : des gestes professionnels ajustés*. Toulouse : Octarès éditions.

Cadre Européen Commun de Référence pour les Langues (CECRL) : [http://www.coe.int/T/DG4/Linguistic/Source/Framework\\_FR.pdf](http://www.coe.int/T/DG4/Linguistic/Source/Framework_FR.pdf) Mis à jour le 1/09/2009

CASTELLOTTI V., MOORE D. (2002). *Représentations sociales des langues et enseignements. Guide pour l'élaboration des politiques linguistiques éducatives en Europe- De la diversité linguistique à l'éducation plurilingue*. Strasbourg : Division des politiques linguistiques, Direction de l'éducation scolaire, extra-scolaire et de l'enseignement supérieur, DGIV, Conseil de l'Europe.

CASTELLOTTI V., MOORE D. (2008). Contextualisation et universalisme. Quelle didactique des langues pour le XXIème siècle ? In P. Blanchet, D. Moore, & S. Asselah Rahal (Eds), *Perspectives pour une didactique des langues contextualisée* (pp. 197-218). Paris : Editions des archives contemporaines.

CICUREL F. (2003). Figures de maîtres. *Le français dans le monde*, 326, 32-34.

COLLETA J.M. (2004). *Le développement de la parole chez l'enfant âgé de 6 à 11 ans : corps, langage et cognition*. Hayen (Belgique) : Pierre Mardaga éditeur.

Convention ICOR : [http://icar.univlyon2.fr/ecole\\_thematique/tranal\\_i/documents/Mosaic/ICAR\\_Conventions\\_ICOR.pdf](http://icar.univlyon2.fr/ecole_thematique/tranal_i/documents/Mosaic/ICAR_Conventions_ICOR.pdf) Mis à jour en novembre 2007

COURTILLON J. (2003). *Elaborer un cours de FLE*. Paris : Hachette FLE.

CRAHAY M. (1999). *Psychologie de l'éducation*. Paris : PUF (1<sup>ère</sup> édition).

CUET C. (2008). Une didactique du compromis face aux éditeurs et professeurs chinois : concilier deux approches de l'enseignement et innover dans l'élaboration de manuels. *Les cahiers de l'Acedle*, vol. 5, 1, 151-166.

DEBONO M. (2011). Pour une pédagogie du conflit en Chine? *Synergie Chine*, 6, 127-140.

GRANDATY M., TURCO G. (2001). *L'oral dans la classe : discours, métadiscours, interactions verbales et construction de savoirs à l'école primaire*. Paris : INRP.

GRANDATY M., CHEMLA M.T. (2004). Médiation de l'enseignant dans l'apprentissage : les étayages. In C. Garcia-Debanc, & S. Plane (Eds), *Comment enseigner l'oral à l'école primaire ?* (pp. 173-214). Paris : INRP, Hatier pédagogique.

HU Y. (2011). Place du communicatif et didactique dans les pratiques et dans les représentations d'enseignants chinois de français. *Synergie Chine*, 6, 117-126.

HUVER E. (2012). Penser la diversité formative en didactique des langues. La comparaison entre essentialisation et hétérogénéisation. *Actes du Colloque du RFS Dynamiques plurilingues : des observations de terrains aux transpositions politiques, éducatives et didactiques*, Alger, 31 mai-2 juin 2011.

*Une nécessaire contextualisation : le cas d'une étude de la stratégie enseignante d'étayage, en classe de FLE, en milieu universitaire chinois*

MARTIN E. (2007). *Culture(s) éducative(s) et formation continue dans le contexte de la coopération bilatérale, le cas des enseignants de FLE des universités*. Mémoire de master II professionnel, Université du Maine.

NORMAND-MARCONNET N. (2011). Apprenants chinois et pratiques pédagogiques innovantes: une étude en milieu homoglotte. *Synergie Chine*, 6, 151-165.

RABATEL A. (2004). *Interactions orales en contexte didactique*. Lyon : PUL, coll. IUFM.

RABATEL A. (2007). Les enjeux des postures énonciatives et de leur utilisation en didactique. *Education et didactique*, 2, 87-114.

RISPAIL M., BLANCHET P. (2011). Principes transversaux pour une sociodidactique dite " de terrain". In P. Blanchet, & P. Chardenet (Eds), *Guide pour la recherche en didactique des langues et des cultures. Approches contextualisées* (pp. 65-70). Paris: Éditions des archives contemporaines.

ROBERT J.M. (2009). *Manières d'apprendre, pour des stratégies d'apprentissage différenciées*. Paris : Hachette FLE.

VALLAT C. (2012). *Etude de la stratégie enseignante d'étayage dans des interactions en classe de Français Langue Etrangère (FLE), en milieu universitaire chinois*. Toulouse : Thèse de doctorat, Université Toulouse 2-Le Mirail.

WOOD D., BRUNER J.S., ROSS G. (1976). The role of tutoring in problem solving. *Journal of child psychology and psychiatry*, 17, 89-100.

YANG Y. (2007). L'éclectisme, une nouvelle nécessité. Réflexion sur l'enseignement du français en contexte chinois. *Synergie Chine*, 2, 61-72.

## **Cinéma et variabilité de l'expérience émotionnelle des apprenants de FLE**

Claire Del Olmo<sup>1</sup>

(1) Octogone-Lordat & Université de Toulouse  
cdelolmo@univ-tlse2.fr

**Résumé** Aujourd'hui, de nombreuses réflexions portent sur la dimension émotionnelle du médium cinématographique en tant que moyen favorisant l'enseignement et l'apprentissage des Langues Etrangères (LE). Nous intéressant à l'exploitation pédagogique d'extraits de films d'auteurs de la Nouvelle Vague suscitant des émotions négatives chez les apprenants de Français Langue Etrangère (FLE) de niveau B2, nous nous interrogeons sur l'influence du facteur culturel dans l'expérience émotionnelle accompagnant le visionnement d'extraits. Ainsi, un corpus d'extraits de films a été constitué à partir d'un test de jugement réalisé par des natifs. Ensuite, un dispositif expérimental permettant de comparer les émotions induites par ces extraits chez des sujets de différentes cultures a été mis en place. La réception des extraits a été testée auprès de sujets de nationalité chinoise, italienne et vénézuélienne, ayant tous un niveau B2 en français. Cet article présente les démarches méthodologiques adoptées ainsi que les résultats significatifs du test de réception.

**Abstract** A lot of research is being done nowadays on the emotional dimension of the cinematographic medium, as a means to improve the learning and teaching processes of foreign language teaching. With an interest in the pedagogical use of clips taken from Nouvelle Vague directors' movies, which would induce negative emotions for B2 level students in French as a foreign language, we question the influence of cultural factors on the emotional experience associated with the viewing of these clips. A corpus of clips was thus set up on the basis of an evaluation test conducted with native French speakers. Then, an experimental plan allowing to compare the induced emotions on subjects from different cultures was set up. The reactions to the excerpts were tested on Chinese, Italian and Venezuelan subjects, all with a B2 level in French. This article presents the methodological tools used as well as the significant results gathered from this test.

**Mots clé :** Français Langue Etrangère, cinéma, culture, émotion, cognition, apprentissage.

**Keywords:** French foreign language, cinema, culture, emotion, cognition, learning.

## 1 Introduction

Le médium cinématographique représente aujourd'hui un support pédagogique prépondérant dans l'enseignement du FLE. Sa dimension émotionnelle est au cœur des réflexions. Il permet tant de favoriser l'apprentissage de la langue cible que la mise en place d'activités pédagogiques. Le film, combinant les modalités de perception visuelle et auditive, est un inducteur d'émotions plus puissant que ne le sont les images fixes (Gil, 2009). D'après Maury (2012), un des moyens de susciter l'état émotionnel constitutif de la motivation des apprenants est d'utiliser ce support. Selon Arroyo et Avellino (1997), les émotions éprouvées lors du visionnement d'un court métrage par les apprenants favorisent la mémorisation des expressions langagières et des formes de communication non verbales le constituant. Pour Médioni (2012), les émotions des apprenants visionnant un film sont l'occasion de mettre en place des activités développant le lexique des émotions. Cependant, bien que les avantages de ce support en tant qu'inducteur d'émotions soient soulignés, la sélection d'un extrait par les enseignants reste problématique. Lorsque ces derniers veulent choisir un extrait pour induire certaines émotions chez les apprenants et réaliser des activités spécifiques, ils se confrontent à la problématique de la réception. Un extrait suscite-t-il des émotions identiques chez des apprenants de même niveau mais de cultures distinctes ? Peut-il être l'objet d'exploitations pédagogiques semblables auprès d'apprenants de niveau similaire mais de différentes cultures ? Nous intéressant à l'exploitation d'extraits d'œuvres d'auteurs de la Nouvelle Vague engendrant des émotions négatives chez les apprenants de niveau B2<sup>1</sup>, nous avons constitué un corpus d'extraits et réalisé un test de réception.

## 2 Contexte de la recherche

### 2.1 Cadre théorique

Comme le signale Nugier (2009), selon les théories de l'évaluation cognitive développées à partir des travaux d'Arnold et de Lazarus, l'émotion est un phénomène cognitif et l'évaluation cognitive d'un stimulus explique la différence qualitative des émotions ressenties par divers individus face à un même événement. Or selon la théorie des tendances d'action de Beck décrite par Rusinek (2004), l'émotion ressentie par un individu est la conséquence directe des schémas cognitifs mis en place au fil de ses expériences, en interaction avec le milieu, pour traiter les informations en adoptant un comportement efficace.

Ainsi, des émotions ont tendance à être davantage éprouvées par les membres de certaines communautés culturelles. Il existe des émotions universelles (de valence positive ou négative), des émotions de base (l'amour, la joie, la colère, la peur, la tristesse) partagées par la plupart des êtres humains et des émotions propres à certaines cultures comme la tendresse ou l'engouement pour l'amour (Fischer, Shaver et Carnochen, 1990). D'après la théorie socioconstructiviste des émotions, les émotions sont le produit d'une culture et s'acquièrent via le processus de socialisation (Gordon, 1991).

De fait, l'émotion suscitée par une séquence cinématographique ne dépend pas seulement de la manière dont elle est construite, le montage pouvant bien sûr influencer la lecture du

---

<sup>1</sup> Un individu de niveau B2 : « Peut comprendre la plupart des journaux et des magazines télévisés. Peut comprendre un documentaire, une interview [...] et la plupart des films en langue standard. » (CERCL, 2001 : 59).

spectateur et la nature de l'émotion ressentie. Elle dépend de la grille de lecture de ce dernier (Serceau, 2009). Laulan (1976 : 83) postule que : « le cadre de vie de chaque spectateur (le contexte) a façonné en lui un ensemble de jugements de valeur, un faisceau de réactions émotionnelles, un système de pensée qui entrent en œuvre devant chaque « texte » : livre, film, émission... ». Avec le test de réception d'un film intitulé *Les Tziganes*, Laulan (1976) montre que des enfants de milieux modestes jugent la scène d'une veillée funèbre comme « triste » tandis que ceux de milieux aisés comme « solennelle ». L'influence de la culture dans l'expérience émotionnelle est soulignée.

De manière générale, pour mesurer les émotions induites par un stimulus chez des sujets, les chercheurs en psychologie recourent à des questionnaires accompagnés d'échelles donnant lieu à une approche discrète des émotions<sup>2</sup> comme celle du Brief Mood Inventory Scale proposant aux sujets de choisir les adjectifs correspondant le plus à leur ressenti parmi une liste. Ils utilisent aussi des échelles non verbales permettant une approche dimensionnelle des émotions telle le Self-Assessment Manikin Scale, échelle iconique, permettant de mesurer les dimensions émotionnelles de plaisir, d'activation ou de dominance soit le caractère plaisant/déplaisant de l'émotion, l'état de calme/excitation du sujet, le contrôle/non contrôle du sujet sur la situation rencontrée.

En outre, selon Kerbrat-Orrechioni (2002), lorsqu'un individu doit sélectionner une unité de son stock lexical pour faire référence à un objet du monde, il peut s'inscrire dans un discours objectif, n'engageant pas sa personne ou subjectif, l'engageant. Ainsi, les adjectifs peuvent être objectifs ou subjectifs. Les adjectifs objectifs, tels les adjectifs de couleur, ne relèvent pas d'une réaction affective ou d'une évaluation de jugement du locuteur contrairement aux adjectifs subjectifs comme « émouvant » (révélant un engagement émotionnel) ou « beau » (révélant un jugement). Les adjectifs tels qu'« émouvant » sont des adjectifs subjectifs affectifs. Ils révèlent une propriété de l'objet défini ainsi qu'un engagement émotionnel du locuteur.

Ainsi, l'un des moyens d'observer l'émotion éprouvée par un locuteur est d'analyser la subjectivité du lexique utilisé dans son discours. Cette subjectivité peut s'observer grâce à des outils comme Emotaix-Tropes (Piolat et Bannour, 2009)<sup>3</sup>, permettant d'identifier le contenu du lexique émotionnel et affectif ainsi que de diagnostiquer les états émotionnels des locuteurs. La valence positive ou négative et la spécificité de la nature de l'émotion d'un sujet portant sur un objet peuvent être définies en fonction de la valence des unités lexicales<sup>4</sup> qu'il utilise pour le qualifier ainsi que des catégories sémantiques dans lesquelles ces dernières s'inscrivent.

## **2.2 Hypothèses de recherche**

Au vu du cadre théorique exposé, nous formulons l'hypothèse générale suivante : des extraits de films peuvent susciter différentes émotions chez des sujets de cultures distinctes, ayant

<sup>2</sup> Les émotions discrètes correspondent aux émotions de base ; elles sont distinctes les unes des autres.

<sup>3</sup> L'outil définit le contenu de l'expression des émotions « au sens large » ; il prend en compte les émotions de base, sociales mais aussi les sentiments, humeurs, affects car lorsqu'un sujet exprime une émotion, il ne recourt généralement pas à des termes relevant stricto-sensu des émotions de base ou complexes (combinant ces dernières).

<sup>4</sup> La valence d'une unité lexicale est positive si cette lexie révèle un vécu subjectif agréable (une émotion procurant du plaisir et étant de valence positive) ou négative si elle révèle un vécu subjectif désagréable (une émotion procurant du déplaisir et étant de valence négative).

tous un niveau B2 en français. Nous émettons ces sous hypothèses : la culture du sujet peut déterminer la présence ou l'absence d'émotion ressentie face à un extrait, sa valence (positive ou négative) ainsi que sa nature (type d'émotion positive ou négative). Pour vérifier nos hypothèses nous créons un corpus d'extraits à partir d'une tâche de verbalisation de l'expérience émotionnelle de natifs accompagnant le visionnement de ces derniers et mettons en place un test de réception réalisable par des sujets étrangers de niveau B2 en français et comprenant ce même type de tâche, soumettre ces derniers aux échelles citées précédemment semblant complexe<sup>5</sup>.

### 3 Etude n°1 : constitution d'un corpus d'extraits

#### 3.1 Matériel

Pour tester nos hypothèses, nous construisons un corpus d'extraits. Nous pré-sélectionnons 10 extraits de films de Truffaut, Godard et Rohmer dont 5 nous semblent pouvoir induire des émotions négatives. 5 extraits présentant, selon nous, des personnages *violents* et 5 des personnages *calmes*, *doux*<sup>6</sup> sont pré-sélectionnés. Les extraits et les films dont ils sont issus sont :

- Extrait A « Menace de mort après une rupture »/*Jules et Jim* (01:36:06-01:36:26).
- Extrait B « Femme qui n'aime plus sa vie »/*Pierrot le fou* (00:50:06-00:51:17).
- Extrait C « Dispute conjugale »/*Les nuits de la pleine lune* (00:29:46-01:31:08).
- Extrait D « Incendie »/*Les 400 coups* (00:27:03-00:28:04).
- Extrait E « Menace physique »/*Pierrot le fou* (01:10:29-01:11:43).
- Extrait F « Discussion en vacances »/*Pauline à la plage* (00:01:40-00:03:06).
- Extrait G « Rencontre dans un café »/*Les nuits la pleine lune* (01:23:32-01:24:41).
- Extrait H « Déjeuner en terrasse »/*L'ami de mon amie* (01:22:10-01:23:12).
- Extrait I « Discussion père-fils »/*Les 400 coups* (00:49:22-00:49:23).
- Extrait J « Début d'une journée de travail »/*Le beau mariage* (00:09:04-00:10:09).

#### 3.2 Participants

Afin de déterminer la pertinence de notre pré-sélection, les extraits sont jugés par des natifs dans le but d'observer quelles émotions ils ressentent en visionnant ces derniers. 32 français âgés de 19 à 30 ans sont interrogés.

#### 3.3 Procédure expérimentale

Les sujets visionnent les extraits en groupe (8, +/-2 sujets) et, sous forme « papier-crayon », les qualifient en donnant un premier adjectif puis un deuxième voire un troisième adjectifs de plus. Les extraits sont projetés dans un ordre aléatoire.

#### 3.4 Méthodologie d'analyse

---

<sup>5</sup> Soumettre les sujets étrangers aux échelles verbales semble difficile car ces derniers peuvent éprouver des difficultés de compréhension au vu de leur niveau en français et l'explication des trois dimensions du sentiment mesurées par l'échelle non verbale SAM semble complexe.

<sup>6</sup> Pour pré-sélectionner les extraits, nous nous basons sur les définitions de ces adjectifs du Trésor de la Langue Française Informatisé, disponible à l'adresse suivante : [atilf.atilf.fr](http://atilf.atilf.fr).

Pour observer les émotions induites par les extraits chez les natifs, les adjectifs recueillis sont analysés avec le logiciel Emotaix-Tropes (Piolat et Bannour, 2009). L'item donné en premier par chaque sujet est d'abord analysé. S'il a une valence positive ou négative, il est retenu pour diagnostiquer la valence de l'émotion induite par l'extrait chez le sujet. Cependant, s'il n'a pas de valence positive ou négative, c'est le deuxième item donné par le sujet qui est pris en compte ou le cas échéant le troisième.

De plus, si un premier vocable donné par un sujet est polysémique et peut avoir plusieurs valences, c'est la deuxième ou troisième unité lexicale donnée qui, ayant une valence univoque, est considérée pour diagnostiquer la valence de l'émotion éprouvée lors du visionnement d'un extrait. En outre, si le premier item donné par un sujet est lié à une émotion sans valence, c'est le deuxième, voire le troisième item donné qui est pris en compte<sup>7</sup>. Cependant, si aucun de ces items n'a de valence, le premier item est tout de même retenu pour diagnostiquer l'émotion éprouvée par le sujet. Enfin, lorsqu'un sujet donne un premier item à valence non spécifiée et un second lié à une émotion sans valence (relevant de la surprise ou de l'impassibilité), c'est ce dernier qui est pris en compte car la valence non spécifiée ne peut l'être au vu du contexte des réponses, les sujets ayant seulement donné des adjectifs pour qualifier les extraits et non des phrases.

### **3.5 Résultats**

Les statistiques descriptives effectuées sur les analyses menées montrent que la moitié ou plus des sujets ont une émotion négative en visionnant les extraits A, B, C, D, E. Les ayant pré-sélectionné en pensant qu'ils pouvaient induire des émotions de cette valence, nous les retenons pour le corpus. Cependant, plus de la moitié des sujets ont une émotion positive en visionnant l'extrait F. Bien que nous pensions qu'il n'induirait pas d'émotion, il est retenu au vu de l'homogénéité de sa réception. De plus, environ 45% des sujets ressentent une émotion négative en visionnant l'extrait G « Rencontre dans un café » et parmi les sujets restants, environ la moitié ressent une émotion positive et l'autre ne ressent rien. Ayant pré-sélectionné ce dernier en pensant qu'il n'induirait pas d'émotion chez la majorité des sujets, nous ne le retenons pas. Enfin, la majorité des sujets ne ressent pas d'émotion en visionnant les extraits H, I, J. Or afin de ne conserver dans le corpus que des extraits perçus de façon relativement homogène par les natifs, nous notons si les items saisis pour leur qualification partagent un trait définitoire. Nous constatons que les sujets ont tendance à recourir à trois, quatre notions pour qualifier les extraits I et J et à plus de 6 notions pour l'extrait H « Déjeuner en terrasse » ; extrait qui n'est donc pas retenu. Notre corpus comprend, au final, les extraits A, B, C, D, E, F, I, J.

## **4 Etude n°2 : test de réception auprès de sujets étrangers**

### **4.1 Matériel**

Pour observer les émotions induites par les extraits du corpus retenus chez des sujets de niveau B2 en français de différentes cultures, nous créons un questionnaire réalisable par des sujets de ce niveau et adapté à leur diversité linguistique. La première partie comprend des questions ouvertes et vise l'expression du ressenti des sujets suite au visionnement des extraits dans leur Langue Maternelle (LM). La seconde contient des questions fermées et vise la sélection de l'item d'une liste pour qualifier les extraits. La liste inclut les adjectifs des réceptions majoritaires des natifs (positive, négative ou non émotionnelle) de chaque extrait

---

<sup>7</sup> Le but est d'observer si le sujet ressent une émotion positive ou négative.

les plus donnés par ces derniers comme premier adjectif pour les qualifier lors du test de jugement. Si certains sont autant donnés par les natifs, celui apparaissant le plus parmi les seconds adjectifs donnés par les natifs est retenu.

La compréhension de la liste est testée auprès de 13 sujets malaisiens de niveau B2<sup>8</sup>. Au vu de leurs difficultés à comprendre les items *intrigant* et *tendu*, correspondant aux premiers adjectifs les plus donnés par les natifs pour qualifier les extraits C et E, nous remplaçons ces derniers par les autres premiers adjectifs des réceptions majoritaires les plus donnés par les natifs pour qualifier ces extraits. Si certains apparaissent avec une fréquence égale, nous retenons celui apparaissant le plus parmi les seconds adjectifs donnés. S'ils apparaissent autant parmi les seconds adjectifs donnés par les natifs, celui qui apparaît le plus parmi les troisièmes adjectifs donnés est retenu. Nous retenons ainsi l'adjectif *agressif* pour l'extrait C et constatons que pour l'extrait E, les adjectifs *violent* et *menaçant* sont donnés par les sujets autant de fois. Or comme Emotaix-Tropes classe ces derniers dans la même catégorie sémantique de la malveillance, nous retenons l'item *violent*, compris par les sujets malaisiens. La liste finale contient les items *tragique* (retenu pour l'extrait A), *agressif* (retenu pour les extraits B, C), *violent* (retenu pour les extraits D, E), *familial* (retenu pour l'extrait I), *banal* (retenu pour l'extrait J), *calme* (retenu pour l'extrait F). Ceux sont les adjectifs les plus donnés par les natifs pour qualifier les extraits lors du test de jugement et compris par les sujets malaisiens.

## 4.2 Participants

Trois groupes de 20 sujets de culture proche, éloignée et intermédiaire par rapport à la culture française répondent au questionnaire<sup>9</sup>. Les premiers sont italiens. Leur LM est, comme le français, romane et leur culture, comme la française, intègre les cultures dites méditerranéennes. De plus, selon Guy (2008), considérant l'apprentissage du français et les voyages en France, 5 italiens sur 10 présentent une proximité importante avec la France. Les seconds sont chinois. Leur LM, contrairement au français, est à tons et s'écrit avec un système logographique. Leur culture est généralement considérée éloignée de celles occidentales. Harbour (2009) la qualifie de collectiviste, notant qu'elle comprend des valeurs de respect des processus organisationnels, d'identification aux membres du groupe avec des attentes de soutien, réciprocité alors que les sociétés occidentales ont des cultures dites individualistes aux valeurs d'équité, d'indépendance. Les derniers sont vénézuéliens. Leur LM est romane. Leur culture, par son héritage culturel espagnol, est relativement proche de la culture française, ces dernières étant dites méditerranéennes, mais aussi éloignée par sa situation géographique et historique.

## 4.3 Procédure expérimentale

Les sujets participent à deux conditions expérimentales (C1, C2) en groupes (5, +/-3 sujets) sous forme « papier-crayon ». Lors de la C1, ils visionnent les extraits et donnent un mot de leur LM pour exprimer leur ressenti (première partie du questionnaire créé). Lors de la C2, ayant lieu le lendemain de la C1, ils les visionnent de nouveau et sélectionnent un adjectif de la liste construite pour les qualifier (seconde partie du questionnaire créé). L'ordre de projection des extraits est aléatoire.

<sup>8</sup> Les sujets étudient à l'Université Toulouse 2 le Mirail (UTM).

<sup>9</sup> Ceux de culture proche et éloignée sont interrogés à Toulouse, et ceux de culture intermédiaire à Caracas.

#### **4.4 Méthodologie d'analyse**

Les mots donnés par les sujets pour exprimer leur ressenti en visionnant les extraits dans leur LM recueillis lors de la C1 et les adjectifs français choisis pour qualifier les extraits recueillis lors de la C2 sont analysés séparément. De plus, pour observer les émotions des sujets lors de la C1 et les comparer, nous les analysons avec une même grille. Pour cela, les données saisies en langues chinoise, italienne et espagnole sont traduites en français par des experts<sup>10</sup>. Toutes les données sont analysées avec Emotaix-Tropes (Piolat et Bannour, 2009), permettant d'identifier le lexique émotionnel et sa valence.

En outre, lorsque les caractères chinois sont traduits par plusieurs items, nous notons l'inscription de chacun dans le lexique émotionnel et leur valence pour identifier l'émotion d'un sujet visionnant un extrait. Si un nom est suivi d'un adjectif qualificatif et qu'ils ont des valences distinctes, que l'un réfère à un vécu subjectif agréable et l'autre désagréable, nous considérons la valence de l'adjectif pour diagnostiquer l'émotion du sujet car il exprime la qualité de l'être, de la chose à laquelle il se rapporte : le nom. De plus, lorsqu'une traduction est composée de deux noms relevant du lexique émotionnel mais de valences distinctes, nous ne pouvons déterminer la valence positive ou négative de l'émotion ressentie. L'émotion du sujet n'est pas identifiée. Enfin, lorsqu'une traduction contient deux noms de même valence, nous considérons la catégorie sémantique dans laquelle s'inscrit le premier pour diagnostiquer l'émotion du sujet<sup>11</sup>.

Afin d'observer l'influence de la culture des sujets sur leurs émotions, nous notons s'il y a une différence significative entre le nombre de sujets de deux groupes ressentant une émotion ou non, une émotion négative ou un autre type d'émotion, une émotion négative ou positive, des émotions négatives ou positives spécifiques lors du visionnement d'un extrait pendant la C1 et la C2. Le test statistique du Chi 2 d'indépendance ou la correction de Yates soit le test exact de Fisher sont utilisés<sup>12</sup>.

#### **4.5 Résultats**

La significativité des différences entre le nombre de sujets de deux groupes de participants de culture distinctes ressentant certaines émotions en visionnant les extraits durant les C1 et C2 est présentée dans le tableau ci-dessous.

---

<sup>10</sup> Les experts sont professeurs à l'UTM d'espagnol, de chinois et d'italien. Conscients des limites de la traduction en français des données saisies pour l'identification des émotions suscitées par les extraits chez les sujets puisque la perte ou modification d'informations est inhérente à la traduction, elle nous permet toutefois d'adopter un cadre d'analyse commun ainsi qu'une approche comparative.

<sup>11</sup> Nous considérons que l'unité lexicale activée en premier par le sujet illustre sa réception première.

<sup>12</sup> Le Chi 2 permet de vérifier si les variables de la culture et de l'émotion ont une relation de dépendance et si la première influence la seconde. La correction de Yates recalcule le Chi 2 pouvant être surévalué et le test exact de Fisher s'utilise pour les petits échantillons ; ils s'interprètent de façon identique.

Emotion C1	Chinois	Italiens	Vénézuéliens	Significativité
<b>Extrait A « Menace de mort après une rupture »</b>				
Non négative <sup>13</sup>	8		1	$\chi^2$ de Yates (1, N = 31) = 5,109, $p < 0,05$
Négative	8		14	
Positive	4		0	$p < 0,05$ (Test exact de Fisher)
Négative	8		14	
<b>Extrait D « Incendie »</b>				
Positive		5	0	$p < 0,05$ (Test exact de Fisher)
Négative		8	15	
<b>Extrait E « Menace physique »</b>				
Anxiété		4	1	$p < 0,05$ (Test exact de Fisher)
Mal-être		1	7	
<b>Extrait I « Discussion père-fils »</b>				
Emotion	8		2	$\chi^2$ (1, N = 40) = 4,800, $p < 0,05$
Pas d'émotion	12		18	
<b>Extrait J « Début d'une journée de travail »</b>				
Non négative	4	1		$p < 0,05$ (Test exact de Fisher)
Négative	0	5		
<b>Emotion C2</b>				
<b>Extrait C « Dispute conjugale »</b>				
Malveillance	18	13		$P < 0,05$ (Test exact de Fisher)
Mal-être	0	4		

Tableau 1 : Différences significatives entre le nombre de sujets ayant certaines émotions

Les analyses des données saisies via la C1 montrent que plus de sujets chinois que vénézuéliens ressentent une émotion en visionnant l'extrait I, montrant un père et son fils en train de discuter, et que plus de sujets vénézuéliens que chinois ne ressentent pas d'émotion. Or selon les tests statistiques menés, les taux de sujets de ces cultures ressentant ou non une émotion sont significativement différents ( $\chi^2$  (1, N = 40) = 4,800,  $p < 0,05$ ). Leur culture influence le fait qu'ils ressentent une émotion. Les variables de la culture et de l'émotion ont une relation de dépendance d'intensité modérée ( $V = 0,346$ ).

En outre, plus de sujets vénézuéliens que chinois ressentent une émotion négative en visionnant l'extrait A, montrant une menace de mort suite à la rupture d'un couple, et autant de sujets chinois ressentent une émotion négative qu'une émotion non négative. De plus, plus de sujets chinois que vénézuéliens ont une émotion qui n'est pas négative. La différence entre les

<sup>13</sup> Les émotions non négatives comprennent les émotions positives et sans valence dont la surprise et les émotions non spécifiées (non identifiables comme agréables ou désagréables).

taux de sujets de ces cultures ressentant une émotion négative ou non négative est significative ( $\chi^2$  corrigé de Yates (1, N = 31) = 5,109,  $p < 0,05$ ).

Il est aussi noté que 5 sujets italiens ont une émotion négative contre aucun sujet chinois en visionnant l'extrait J, montrant une employée arrivant sur son lieu de travail, et que plus de sujets chinois qu'italiens ont une émotion non négative. Les taux de sujets de ces cultures ressentant une émotion négative ou non négative sont significativement différents ( $p < 0,05$ , test exact de Fisher). Leur culture influence la nature de leur émotion.

Ensuite, plus de sujets vénézuéliens que chinois ressentent une émotion négative en visionnant l'extrait A et seulement 4 sujets chinois ressentent une émotion positive. Les taux de sujets de ces cultures ressentant une émotion négative ou positive sont significativement distincts ( $p < 0,05$ , test exact de Fisher). La valence de leur émotion est influencée par leur culture.

Puis, l'extrait D, montrant un repas interrompu par un incendie, suscite une émotion négative chez plus de sujets vénézuéliens qu'italiens et seulement une émotion positive chez 5 sujets italiens. La différence entre les taux de sujets de ces cultures ressentant une émotion négative ou positive est significative ( $p < 0,05$ , test exact de Fisher). L'émotion des sujets est influencée par leur culture.

Enfin, l'extrait E, montrant une scène de menace, suscite de l'anxiété chez plus de sujets italiens que vénézuéliens et du mal-être chez plus de sujets vénézuéliens. Les taux de sujets de ces cultures ressentant de l'anxiété ou du mal-être sont significativement différents ( $p < 0,05$ , test exact de Fisher). La spécificité de leur émotion négative est influencée par leur culture.

Par ailleurs, selon les analyses des données saisies via la C2, l'extrait C, montrant une dispute conjugale, suscite du mal-être chez quelques sujets italiens mais non les sujets chinois et plus de sujets chinois qu'italiens ressentent de la malveillance. Or la différence entre les taux de sujets de ces cultures ayant ces ressentis est significative ( $p < 0,05$ , test exact de Fisher). Leur culture influence ce qu'ils éprouvent.

## **5 Conclusion et perspectives**

Les résultats de l'étude n°2 mettent en évidence qu'un extrait cinématographique peut susciter des émotions significativement différentes chez des groupes de sujets de niveau B2 en français de distinctes cultures. Ainsi, lorsqu'un enseignant veut mettre en place des activités constituées de débats, théâtralisations, jeux de rôles à partir de l'induction d'émotions spécifiques chez les apprenants via la projection d'une séquence, il doit être conscient des limites présentes dans la sélection de son support et de la possible variabilité des émotions qu'il induira. La variable culturelle doit être envisagée comme facteur pouvant modifier les paramètres de la situation d'enseignement recourant à l'émotion comme moyen d'apprentissage. Enfin, les résultats exposés représentent un indice sur la façon dont les extraits peuvent être perçus par des apprenants de la culture des participants interrogés et ouvrent des pistes de réflexions sur leur exploitation.

## **Références**

ARROYO, F. AVELLINO, C. (1997). Un film contre un virus. *Le français dans le monde* 288, 65-68.

CONSEIL DE L'EUROPE. (2001). Un cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer. Strasbourg : Conseil de L'Europe.

FISCHER, K., W. SHAVER, P. R. CARNOCHEN, P. (1990). How emotions develop and how they organize development. *Cognition and emotion* 4, 81-127.

GORDON, S. (1991). « The socialization of children's emotions : emotional culture, competence and exposure » In Sarni, C. Harris, P.L. *Children's Understanding of Emotion*. New-York : CUP, 319-349.

GIL, S. (2009). Comment étudier les émotions en laboratoire. *Revue électronique de Psychologie Sociale* 4, 15-24. Disponible à : <http://RePS.psychologie-sociale.org/>.

GUY, J.M. (2008). *Cultures croisées*. Disponible à : <http://www.culture.gouv.fr/deps>.

HARBOUR, M. (2009). L'altruisme et le modèle coopératif. *La revue des sciences de gestion* 239-240, 87-95.

KERBRAT-ORECCHIONI, C. (2002). *L'énonciation*. Paris : Armand Colin.

LAULAN, A.M. (1976). Un film et ses filtres culturels. *Communication et langages* 30, 83-96.

MAURY, C. (2012). L'intégration du cinéma Hollywoodien en classe de langue : réflexions sur quelques points de convergence disciplinaires. *Cahiers de l'APLIUT* 31 (2), 26-43. Disponible à : <http://apliut.revues.org/2661>.

MEDIONI, M.A. (2012). De la motivation à la mobilisation intellectuelle. *Cahiers de l'APLIUT* 31 (2), 140-150. Disponible à : <http://apliut.revues.org/2736>

NUGIER, A. (2009). Histoire et grands courants de recherche sur les émotions. *Revue électronique de Psychologie Sociale* 4, 8-14. Disponible à : <http://RePS.psychologie-sociale.org/>.

PIOLAT, A. BANNOUR, R. (2009). EMOTAIX : Un Scénario de Tropes pour l'identification automatisée du lexique émotionnel et affectif. *L'Année Psychologique* 109, 657-700.

RUSINEK, S. (2004). *Les émotions*. Paris : Dunod.

SERCEAU, M. (2009). *Le mythe, le miroir, le divan : pour lire le cinéma*. Villeneuve d'Ascq : Presses Univ Septentrion.

## **Ditransitive Complementation in Medical Research Articles**

Sabiha Choura<sup>1</sup>

(1) University of Arts and Humanities of Sfax  
sabahachoura@yahoo.fr

**Abstract** The present research analyzes ditransitive verbs from a systemic experiential perspective according to six process types which are material processes, mental processes, relational processes, verbal processes, behavioral processes, and existential processes in a particular genre which is the research article genre and in a particular discipline which is the medical discipline. This research attempts to display the distribution of ditransitive process types and to account for the differences as well as the similarities found in these distributions. Both quantitative and qualitative approaches are adopted to the analysis of the corpus. The present research demonstrates that the corpus displays variations in the distribution of the ditransitive process types. Indeed, while material and relational processes are the most frequent in the medical corpus, verbal and mental processes are the least presented. This is accounted for in terms of the specificities of the medical discipline and the norms of research article writing.

**Keywords:** Ditransitive Complementation, Experiential Metafunction, Process Types, Research Articles Genre, Medical Discipline

## 1 Introduction

Since verb complementation is “an all- pervading structural feature of language and thus likely to be more significant in giving a variety in its character than, for example, lexis.” (Olavarria de Ersson & Shaw, 2003, p. 138), studying its distribution in a particular genre is important. Actually, genre determines the choice of linguistic patterns. (Swales, 1990, p. 41)

Research on ditransitive verbs in the research article genre and the scientific writing genre is explained by the gap in the literature as far as this structure is concerned. As Mukherjee (2005, p. 267) puts it: “at the level of language description, it will be useful to complement the language-as-a-whole description offered in [my] study with a more fine-grained analysis of genre-specific trends in using ditransitive verbs.” The analysis of these verbs will be based on a systemic functional analysis as the distribution of process types according to genre is scarcely studied in the literature; while Harvey (2001) investigates relational identifying clauses in English technical manuals, Arùs and Lavid (2001) focus on the grammar of relational processes in English and Spanish along two genres which are machine-aided translation and multilingual generation. From a systemic functional perspective, « clauses which have a Recipient or Beneficiary as an Indirect Object are referred to as ditransitive » (Lock, 1997, p.76). In particular, in the experiential metafunction, ditransitive processes are classified into main types of process including material, relational and mental processes and borderline ones referring to existential, behavioural and verbal processes (Halliday & Matthiessen, 2004, p.171). These processes are semantically different.

- Material processes represent processes of doing and happening. E.g. *I gave my love a ring that has no end.* (Halliday & Matthiessen, 2004, p.191).
- Mental processes are “processes of sensing” (Halliday and Matthiessen, 2004, p.197). E.g. *What really irritates me is that a lot of people go to socialise in pubs.* (Halliday & Matthiessen, 2004, p.205)
- Relational clauses are used for characterization and identification. E.g. *We might call it the authorial voice.* (Halliday & Matthiessen, 2004, p.237)
- Verbal processes are used to say a message. E.g. *Did you repeat that to your parents?* (Halliday & Matthiessen, 2004, p.255)
- Existential processes refer to the happening or existence of something. E.g. *In Bihar, there was no comparable political campaign.* (Halliday & Matthiessen, 2004, p.256)
- Behavioural processes are “processes of (typically human) physiological and psychological behavior” (Halliday and Matthiessen, 2004, p. 248). E.g. *He ’ s always grumbling.* (Halliday & Matthiessen, 2004, p.251)

Since the present research focuses on the scientific writing genre, its main characteristics have to be identified. Indeed, Alley (1996) claims that scientific writing and in particular medical science are based on precision, clarity, familiarity, fluidity and forthrightness. Not only does the present research focus on this genre, but also on the research article genre. Actually, articles in the present corpus conform to the IMRAD format; they are composed of four main sections: Introduction, Methods, Results and Discussion. The Introduction section contains a statement of the problem, references to previous works, a description of the methodology, and a statement of the results and the corresponding conclusions. The Methods section involves a

description of the experimental design, the subjects and the materials. The Results section presents the results and the limitations of the experiment. Finally, the Discussion section includes a discussion of the results, a statement of the limitations, a statement of the conclusion, a summary and comparison with previous works. (Day and Gastel, 2006, p. 58-69)

Obviously, the present research aims at studying the distribution of ditransitive process types in medical research articles from a systemic functional perspective. This study is based on a quantitative analysis conducted thanks to the UAM corpus tool and a qualitative analysis. The analysis shows that material and relational processes are the most frequent ones. It also reveals that mental and verbal processes are rarely present in the corpus while behavioural and existential processes are completely absent. This can be explained by the nature of scientific writing and in particular the medical discipline, the specificities of the research article and the patterns of ditransitive complementation.

## **2 Methods**

### **2.1 Corpus**

The corpus is composed of ten articles conforming to the IMRAD format. These articles published in 2011 contain 41.677 words. They are randomly selected from a peer-reviewed journal which is the British Medical Journal. The selection of reviewed research articles allows the selection of the best articles which conform to the standardized and formal code of each discipline. (Swales, 2004, p. 208)

### **2.2 Procedure**

The analysis is based on three steps. In the first step, research articles are downloaded from the internet. In the second step, I create a new project thanks to the UAM corpus tool. Then, I design a scheme consisting of six systems which are MATERIAL-TYPE, MENTAL-TYPE, VERBAL-TYPE, RELATIONAL TYPE, BEHAVIOURAL-TYPE and EXISTENTIAL-TYPE. After the incorporation of medical research articles into the new project, I annotate them. First, I read each text, and I identify ditransitive verbs. Afterwards, I classify them into the abovementioned types. For example, *sent* is classified as a ditransitive material type in the following sentence: *we sent a questionnaire including the most important outcome measures to 29 of the 34 patients*. Finally, the different frequencies are presented thanks to the corpus statistics pane. The annotation of the articles has been performed individually without evaluation. The third step is devoted to the interpretation of the results. For instance, the differences in the distribution of process types in the medical discipline are explained.

## **3 Results**

Table 1 displays the distribution of the six ditransitive process types in the medical corpus.

	Medical Discipline	
	Number	Percentage
Material	202	41.65%
Mental	18	3.71%
Relational	238	49.07%
Verbal	27	5.57%
Behavioural	0	0.00%
Existential	0	0.00%
Total	485	

Table 1: Distribution of Ditransitive Process Types in the Medical Corpus

The study shows that the six process types are not equally distributed in the medical corpus. In fact, behavioural processes and existential ones are completely absent; while ditransitive verbs are complemented by two constituents, behavioural and existential processes cannot take three participants. While relational processes are the most frequent as they represent 49.07%, material processes are rated second since the frequency of their occurrence is 41.65%. The frequency of relational and material processes goes hand in hand with the findings of Biber et al. (1999, p. 366) who claim that existence and activity verbs are the most common in academic prose in comparison with mental verbs.

The frequency of occurrence of ditransitive relational processes can be explained by the functions they play. Actually, these processes are important in defining or describing a particular variable in the research. In the following example, they are used to introduce and describe one procedure in the experiment.

The intervention ... **was based** on the following four lifestyle behaviours: physical activity, nutrition, media use, and sleep. (Text 1)

Relational processes also serve as a tool to establish a relationship between two entities or things in the medical discipline. Consider, for example, the following sentence where the writer opts for a relational process in order to relate the problem with its reasons.

Adiposity and low aerobic fitness in children **are associated** with a clustering of cardiovascular risk factors. (Text 1)

As for material processes, they have two major roles. The first one is to denote the type of relationship between entities. When exposing the results of the research, writers opt for material processes in order to explain the kind of relationship between two major constituents in the research as is shown in the following example:

Angiotensin receptor blockers **exert** their action on angiotensin I receptors... (Text 9)

The second role is to describe the activities undertaken by the researchers while preparing the material and conducting the analysis. For example, material processes can reveal how researchers select the subject in the experiment. Take, for example, the following sentence:

They **were recruited** from local hospitals or primary care to their nearest university hospital as usual without any supplemental recruitment attempt. (Text 8)

Mental and verbal processes are not as frequent as relational and material processes since they represent 3.71% and 5.57% respectively. For example, mental processes can express the attempts of the researchers to undertake particular measures as is shown in the following subordinate clause:

As surgical procedures **should be evaluated** against non-surgical methods... (Text 8)

As far as verbal processes are concerned, they reveal the talk between the different participants in the research. Indeed, they report the talk between the researchers and the doctors. They also convey the talk between the researchers and the patients. Consider the following example:

We **asked** all competent patients who agreed to participate in the study to provide a specimen of blood for analysis of haematological and biochemical parameters. (Text 4)

Another infrequent function displayed by verbal processes is to make references to other works; in other words, they express the findings stipulated by previous pieces of research in order to give credibility to the present research. Take, for instance, the following sentence:

Most previous palliative prognostic studies... **have reported** only the results from “evaluable” patients. (Text 4)

## 4 Discussion

Unlike previous researches, the present one examines ditransitive complementation from a systemic functional perspective in the medical discipline. The present study shows that material and relational ditransitive process types are the most frequent while mental and verbal processes are the least presented. As for existential and behavioural processes, they are totally absent from the corpus. This can be explained by the communicative purposes of the research article, the characteristics of scientific writing and the patterns of ditransitive verbs. In fact, the absence of existential and behavioural processes is due to the fact that they do not require three participants as ditransitive verbs do. The limited frequency of mental processes can be explained by the nature of the discipline and genre. Indeed, scientific writing and in particular medical science are not based on thoughts and beliefs. They are characterized by precision. In this context, Alley (1996, p. 73) argues that “in scientific writing, precision is the most important goal of language.” Verbal processes are unexpectedly infrequent; normally, they should be more frequent in order to report the findings of the present research as well as the previous ones. The frequency of relational and material processes can be accounted for by the exigencies of the research article genre. Since every piece of research should be reproducible, the research article should describe the conditions under which the experiment takes place including the variables, the procedures and the activities undertaken by the researchers. The research article should also be original in that it conveys a new piece of knowledge to the scientific community. In the medical corpus, relational processes are mainly used for identification and description while material processes are meant for denoting

activities and stipulating the results. This research could be more interesting if it was based on a larger corpus and if it compared the distribution of process types across different disciplines and different genres.

## **References**

ALLEY M. (1996). *The Craft of Scientific Writing* Third Edition. Springer –Verlag New York, Inc.

ARÙS J. and LAVID J. (2001). *The Grammar of Relational Processes in English and Spanish: Implications for Machine-Aided Translation and Multilingual Generation*. *Estudios Ingleses de la Universidad Complutense*, 9, 61-79

BIBER D., JOHANSSON S., LEECH G., CONARD S., & FINEGAN E. (1999). *Longman Grammar of Spoken and Written English*. Harlow: Longman.

DAY D. and GASTEL B. (2006). *How to Write and Publish a Scientific Paper*. Cambridge University Press

HALLIDAY M.A.K. and MATTHIESSEN C. (2004). *An introduction to Functional Grammar*. Third edition. London : Edward Arnold

HARVEY A. (2001). *Relational Clauses in English Technical Discourse: Patterns of Verb Choice*. *Pragmatics* 11:4.379-400 (2001) International Pragmatics Association.

LOCK G. (1997). *Functional English Grammar. An Introduction for Second Language Teachers*. Cambridge University Press.

MUKHERJEE P. (2005). *English Ditransitive Verbs*. Amsterdam, New York: Rodopi.

OLAVARRIA DE ERSSON E. and SHAW P. (2003). *Verb Complementation in Indian Standard English*. *English World-Wide*. 24 (2). p. 137-161.

SWALES J. (1990). *Genre analysis: English in academic and research settings*. Great Britain. Cambridge University Press.

## **Vers une taxonomie phonétique des sibilantes non voisées dans les variétés de l'espagnol**

Anne-Sophie De Clercq<sup>1</sup>

(1) Institut de Recherche en Sciences et Technologies du Langage (IRSTL) &  
Université de Mons  
anne-sophie.declercq@umons.ac.be

**Résumé** Dans le cadre d'une thèse portant sur le traitement des variations phonétiques par les locuteurs francophones, hispanophones et catalanophones en situation de L2, tant au niveau de la perception que de la production, nous étudions les productions de natifs hispanophones en vue de définir une taxonomie des phénomènes phonétiques observables touchant la réalisation des fricatives non voisées. Après avoir tenté, sans succès, de trouver dans la littérature une méthode de segmentation robuste et fiable nous permettant d'extraire de nos enregistrements les sons fricatifs, nous travaillons actuellement à l'établissement d'une méthode d'isolation des phonèmes qui nous intéressent afin de les analyser et de les caractériser.

**Abstract** As part of a thesis about the processing of phonetical variation at a perception as well as at a production level by French, Spanish and Catalan speakers in L2 context, we study the productions of Spanish native speakers with the intention of defining a taxonomy of the observed phenomena affecting the pronunciation of voiceless fricatives. After unsuccessfully trying to find in the literature a robust and reliable segmentation method allowing us to extract the fricative sounds from our recordings, we are currently working on establishing an isolation method for the phonemes we are interested in, in order to analyse and characterise them.

**Mots-clés :** taxonomie, espagnol, fricative, sibilante, segmentation

**Keywords:** taxonomy, Spanish, fricative, sibilant, segmentation

# 1 Système phonologique de l'espagnol et phonèmes fricatifs

Le sous-système des fricatives espagnoles est l'un des plus étudiés diachroniquement et synchroniquement, d'une part parce qu'il a subi des évolutions importantes au cours des siècles et, d'autre part, parce que les manifestations des phonèmes fricatifs présentent une remarquable variabilité.

## 1.1 Bref historique

À l'époque médiévale, l'espagnol compte trois paires sourde/sonore : deux fricatives apicoalvéolaires (/s/ et /z/), deux fricatives prépalatales (/ʃ/ et /ʒ/) et deux affriquées dentoalvéolaires (/tʃ/ et /dʒ/) qui perdent très tôt l'occlusion pour devenir des fricatives prédorsodentales (/s̺/ et /z̺/). Au XV<sup>e</sup> siècle, la similarité entre les apicoalvéolaires et les prédorsodentales entraîne une restructuration du système et son dédoublement. Dans le nord et au centre de la péninsule, les éléments sonores de chaque couple disparaissent. Le /s̺/, conséquemment à une avancée de son point d'articulation, devient /θ/. Le /ʃ/, quant à lui, recule et devient /x/, tandis que le /s/ se maintient. On aboutit donc à un système à trois fricatives sourdes (/s/, /θ/ et /x/), appelé différentiateur. Dans le Sud, en revanche, avant la disparition des sonores, les sourdes apicoalvéolaire et prédorsodentale fusionnent, de même que leurs pendants sonores ; on ne conserve que les formes prédorsodentales. Les deux résultantes de cette transformation fusionnent ensuite elles-mêmes en la seule sourde prédorsodentale /s̺/. Les prépalatales évoluent de la même manière que dans le Nord. Dans le Sud, il subsiste donc deux fricatives sourdes : /s̺/ et /x/ (Real Academia Española, 2011 ; 165-167).

## 1.2 État des lieux

Les deux grandes variétés ne sont pas stables : au sein de chacun des sous-systèmes phonologiques, la réalisation phonétique des phonèmes varie, comme en témoignent les changements qui touchent les sibilantes /s/ et /s̺/. L'articulation apicale ou apicoalvéolaire [s̺], typique du système différentiateur, est plus postérieure et plus concave au nord qu'au centre de la péninsule ibérique et elle est rare au sud. Elle est également avérée dans certaines zones d'Amérique latine, mais elle n'est pas majoritaire sur ce continent. À l'échelle du monde hispanophone, c'est la production non-apicale [s], également dénommée prédorsale ou prédorsodentale qui domine. Plus variable encore que la précédente, elle est très répandue en Amérique latine, dans les Canaries et en Andalousie. Les variantes se différencient également par leur stridence : la non apicale est plus stridente que l'apicale.

Le /s/ non apical présente une importante tendance à l'aspiration (phénomène du *heheo*), voire à l'effacement, touchant principalement le phonème en finale de mot ou devant une consonne, mais pouvant également se manifester lorsqu'il est en position prévoalique (Momcilovic, 2009 ; 11). Ce phénomène est très marqué socialement et varie largement selon les régions et les locuteurs. Le /s/ apical, à l'inverse de son concurrent, est généralement mieux conservé et les réalisations de certains groupes de locuteurs des zones différenciatrices sont considérées comme canoniques, servant de base à l'enseignement de l'espagnol langue étrangère. Pourtant, l'auteur a observé chez des locuteurs madrilènes – peu nombreux – une tendance, comparable à celle des variétés méridionales et américaines, à l'aspiration du /s/ et même à son effacement touchant prioritairement la fricative en fin de mot, puis en position finale ou non devant une consonne, jusqu'en position prévoalique, démontrant que la réduction du

phonème est bien plus importante que ce que l'on pouvait alors penser, en particulier chez les adultes masculins peu scolarisés (2009).

Ainsi, certaines des réalisations propres à des variétés spécifiques semblent émerger dans des régions dont elles sont traditionnellement absentes. Peu de données expérimentales viennent cependant appuyer ce constat. Il est par conséquent nécessaire de construire une taxonomie de ces phénomènes sur base expérimentale afin de pouvoir en étudier la nature et l'ampleur. Cette étude n'est cependant possible que sur des échantillons de parole correctement segmentés et analysés ; il nous faut donc préalablement établir des critères de segmentation du signal nous permettant d'isoler les fricatives : c'est l'objet de cette communication à caractère méthodologique.

## 2 Protocole expérimental

Nous avons enregistré en chambre sourde, via un enregistreur Sony PCM-D50 Linear Recorder (fréquence d'échantillonnage 44.1 kHz, qualité 16 bits) fixé sur un trépied, six sujets hispanophones volontaires aux origines variées, âgés de 18 à 33 ans. Ils avaient pour tâche de prononcer le plus naturellement possible chacun des pseudo-mots inscrits dans le tableau ci-dessous et retranscrits en orthographe espagnole défilant (un toutes les deux secondes) dans une présentation Powerpoint sur l'écran d'un ordinateur installé du côté expérimentateur de la cabine, derrière une vitre, à hauteur du regard, en les insérant dans une phrase porteuse (« ... tiene dos vocales » pour les séries A et B, « ... tiene tres sonidos » pour la série C). Chaque série était répétée cinq fois, ce nombre de répétitions devant être suffisant pour induire une prononciation naturelle et relâchée des stimuli. Chaque sujet a produit 260 items (5\*36 + 5\*11 + 5\*5).

série A	[ifi]	[ifa]	[ifu]	[afi]	[afa]	[afu]	[ufi]	[ufa]	[ufu]		
	[iθi]	[iθa]	[iθu]	[aθi]	[aθa]	[aθu]	[uθi]	[uθa]	[uθu]		
	[isi]	[isa]	[isu]	[asi]	[asa]	[asu]	[usi]	[usa]	[usu]		
	[ixi]	[ixa]	[ixu]	[axi]	[axa]	[axu]	[uxi]	[uxa]	[uxu]		
série B	[aspa]	[asta]	[aska]	[asma]	[asna]	[asɲa]	[asθa]	[asfa]	[asxa]	[asla]	[asra]
série C	[lis]	[les]	[las]	[lɔs]	[lus]						

## 3 Analyse des enregistrements

Tous les sujets ont été enregistrés en une seule prise d'une vingtaine de minutes. Chaque prise a été segmentée manuellement grâce au logiciel Wavosaure (version 1.0.7.0) à partir de son oscillogramme selon différents niveaux : série, répétition, phrase, pseudo-mot. Nous avons résolu de segmenter chaque item de manière à obtenir des extraits bien spécifiques : la fricative seule, la ou les zone(s) de coarticulation entre la fricative et un son la précédant ou la suivant et ce son seul. L'analyse de tous ces segments et non de la fricative uniquement présente un intérêt puisque la coarticulation, tant régressive que progressive, implique que chacun d'eux peut contenir des informations sur le phonème analysé (Mann & Repp, 1980). Nous nous sommes tournée vers la littérature afin d'étudier les méthodes appliquées à l'isolation des fricatives et de nous en inspirer dans le cadre de notre travail.

Par inspection visuelle (du sonagramme et/ou de l'oscillogramme), « à l'oreille » ou par découpe semi-automatisée, trois indices sont couramment utilisés, seuls ou combinés, pour procéder à l'isolation d'une fricative sibilante sourde hors d'un signal : la périodicité (ou plutôt l'apériodicité, qui révèle le bruit de friction), le timbre (présence d'énergie au-delà de 2,5 kHz) et l'intensité (augmentation vs diminution dans certaines zones de fréquence). Si la

procédure de segmentation est couramment évoquée, elle n'est que rarement détaillée (l'article le plus complet en ce sens est celui de Pirello et al., 1997) et nous n'avons pas trouvé d'article méthodologique critique sur le sujet. Nous avons donc entrepris de segmenter les signaux obtenus, visuellement et par vérification auditive d'abord, en utilisant les outils fournis par Praat ensuite, afin de corroborer nos premières analyses. Nous sommes partie du postulat que la fricative isolée est une zone d'apériodicité complète et que les zones de coarticulation doivent contenir les caractéristiques des deux sons en présence.

Pour les pseudo-mots dans lesquels la fricative se trouve en position intervocalique, nous avons considéré que la fricative isolée et « pure » est contenue dans l'intervalle séparant, dans Praat, le dernier *pulse*<sup>1</sup> de la première voyelle du premier *pulse* de la seconde voyelle. En effet, il fonctionne sur tous les items VFV que nous avons testés. Notons tout de même que, si cette méthode est avantageuse car elle nous permet d'extraire un son totalement apériodique, elle inclut parfois une zone « blanche » ne contenant effectivement aucun son périodique, mais ne contenant pas non plus de bruit fricatif, à éliminer en recourant au nombre de passages à zéro ou au repérage de la zone de fréquence, et elle exclut dans d'autres cas une importante partie du signal et notamment du bruit de friction. Cette zone éliminée peut tout de même être reprise lors des analyses puisqu'elle peut être considérée comme un lieu de coarticulation entre la fricative et la voyelle qui la précède ou la suit.

Peu d'auteurs ont travaillé sur des items où la sibilante sourde /s/ se trouve en position finale, et ceux qui l'ont fait n'ont pas manqué de souligner le problème majeur inhérent à cette position : la continuité du voisement jusqu'à la fin du signal chez certains locuteurs. La présence ou l'absence de *pulses* dans Praat ne peut donc pas être systématiquement utilisée pour délimiter la zone de friction. C'est la friction elle-même qui doit alors, selon nous, servir de marqueur. Le nombre de passages à zéro peut dans ce cas être un indice intéressant à exploiter ainsi que l'intensité dans les zones de fréquence correspondant au profil acoustique du /s/.

Quant aux pseudo-mots V[s]CV, nous proposons de placer la marque d'ouverture de la fricative au dernier *pulse* de la voyelle qui la précède. La marque de fermeture se placera selon des critères dépendant de la nature de la consonne qui suit. Si la consonne suivante est une approximante ou une nasale (donc voisée), les *pulses* peuvent être utilisés pour la première et la dernière marque, comme si on segmentait un VFV, avec les mêmes avantages et les mêmes inconvénients. S'il s'agit d'une occlusive, c'est bien sûr le début du silence précédant l'explosion qui constitue la limite, le « début » restant à fixer (le nombre de passages à zéro pourrait être ici utile de même que la chute d'intensité). Si la consonne suivant la fricative est une autre fricative, il faut jouer sur l'intensité : on voit généralement dans le sonagramme une marque bien claire du passage de l'une à l'autre, grâce à une intensité plus importante dans une zone propre à chaque fricative.

Les zones de coarticulation, enfin, sont repérables grâce au marqueur d'intensité de Praat : une première zone stable marque la présence de la première voyelle ou de la nasale ; vient ensuite une pente descendante, typique de la zone de coarticulation, suivie d'une nouvelle zone stable, la fricative elle-même. La seconde zone de coarticulation se marque en miroir lorsque la fricative est suivie d'une voyelle : une pente montante, menant à une dernière zone de stabilité. Quant aux transitions de la fricative à une plosive ou à une autre fricative, elles sont également marquées par des variations d'intensité, mais aussi de hauteur.

---

<sup>1</sup> L'outil *pulses* disponible via le logiciel Praat permet de détecter la périodicité dans un signal ; cette périodicité est marquée par une série de barres verticales s'intégrant dans l'oscillogramme.

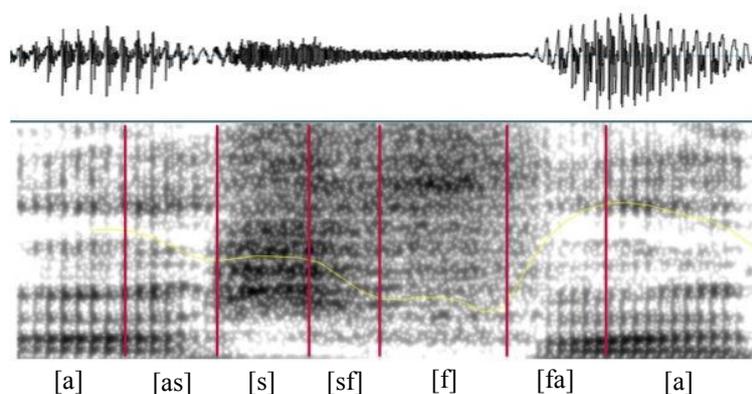


Figure 1: proposition de segmentation de asfa (outil « intensity » de Praat)

## 4 Perspectives en guise de conclusion

Nos premières analyses nous ont permis de déterminer un certain nombre de critères permettant de segmenter le signal en vue d'isoler une fricative sourde ainsi que les zones de coarticulation l'entourant. Il nous reste à présent à fixer définitivement nos critères de segmentation par un calibrage précis de Praat, puis à utiliser ce système sur nos enregistrements avant de pouvoir *in fine* analyser chacune des fricatives extraites et établir une taxonomie des phénomènes qui les touchent.

## Remerciements

Je tiens à remercier mes superviseurs au Laboratoire de Phonétique de l'Université de Mons, Véronique Delvaux, Kathy Huet, Myriam Piccaluga et Bernard Harmegnies, pour leurs relectures attentives et leurs commentaires toujours éclairés et éclairants.

## Références

- BROERSMA M. (2010). Perception of final fricative voicing: Native and nonnative listeners' use of vowel duration, *Journal of the Acoustical Society of America* 127 (3), 1636-1644.
- MANN V. & REPP B. (1980). Influence of vocalic context on perception of the [ʃ]–[s] distinction, *Perception & Psychophysics* 28 (3), 213-228.
- MOMCILOVIC N. B. (2009). *A sociolinguistic analysis of /s/-aspiration in Madrid Spanish*, Munich, Lincom.
- PIRELLO K., BLUMSTEIN S. & KUROWSKI K. (1997). The characteristics of voicing in syllable-initial fricatives in American English, *Journal of the Acoustical Society of America* 101 (6), 3754-3765.
- Real Academia Española (2011). *Nueva gramática española. Fonética y fonología*, Barcelone, Espasa Libros.

## **Comparaison de types d'erreurs orthographiques en FLM et FLE : une revue d'articles**

Katayoon Katoozian<sup>1</sup>

(1) DILTEC & Université Sorbonne Nouvelle – Paris3  
kathy.katoozian@laposte.net

**Résumé** Les difficultés de l'orthographe française poussent les usagers de cette langue à produire des « variantes » par rapport à la norme, c'est-à-dire des « erreurs orthographiques ». Selon les différentes typologies pour la classification des erreurs d'orthographe, il existe plusieurs types d'erreurs dont l'étude a fait l'objet de recherches variées. Le propos de cette contribution, basée sur une revue d'articles, est de comparer les types d'erreurs orthographiques en FLM (Français Langue Maternelle) et FLE (Français Langue Étrangère) décrits par quelques recherches descriptives dont l'instrument d'observation était la dictée traditionnelle. Notre objectif est d'identifier les types d'erreurs les plus fréquents chez les francophones et les non francophones et ceux fréquents seulement chez ces derniers. Cette révision a permis de constater que les erreurs d'orthographe grammaticale sont parmi les plus fréquentes en FLM et FLE tandis que les erreurs de type phonétique ne le sont que pour le second.

**Abstract** Because of the difficulties of French orthography, users of this language tend to produce "variants" in relation to the norm, i.e. "orthographical errors". Depending on different methods of classifying orthographical errors, there are several types of errors that have been the object of various investigations. The purpose of this presentation, based on a review of articles, is to compare types of orthographical errors between FFL (French as a First Language) and FFL (French as a Foreign Language) described by a few investigations using dictation as an instrument of observation. Our aim is to identify the most frequent types of orthographical errors for both native and non-native speakers of French and those most frequent only for the latter. This review found that grammatical errors are among the most frequent types of orthographical errors for the two groups, whereas phonetic errors are frequent only for learners of French as a Foreign Language.

**Mots-clés :** orthographe française, erreur, type, Français Langue Maternelle, Français Langue Étrangère

**Keywords:** French orthography, error, type, French as a First Language, French as a Foreign Language

## 1 Introduction

À partir des années 1970, les linguistes étudiant le système de l'orthographe française se sont intéressés à l'étude de la variante. Non seulement les variations dans l'histoire de l'orthographe elle-même mais aussi celles des formes non conformes à la norme, c'est-à-dire les erreurs d'orthographe, sont devenues un sujet digne d'analyse surtout pour la linguistique descriptive (analyse du corpus). Les erreurs deviennent ainsi « matière à réflexion » (Manesse, 2007). Plusieurs recherches ont été menées se penchant sur les erreurs orthographiques d'un public donné dans une gamme d'objectifs variée : comparaison du niveau orthographique des élèves de telle ou telle époque, comparaison entre les différents niveaux scolaires ou chez les publics différents, simple description des erreurs d'orthographe d'un public donné, etc. Parmi les différents types de recherches, cet article se propose de réviser quatre recherches descriptives sur l'étude des erreurs orthographiques, deux dans le domaine du FLM et deux dans celui du FLE, et ceci dans un ordre chronologique. L'objectif poursuivi est de comparer les résultats afin de connaître les types d'erreurs les plus fréquents chez les francophones et les non francophones et ceux qui sont fréquents seulement chez ces derniers. Les recherches descriptives constituent près de 39% de la totalité des recherches en didactique de l'orthographe (Jaffré, 1992). Dans ce type de recherche, comme son nom l'indique, il s'agit d'observer la situation didactique afin de la comprendre. Le point commun entre les recherches révisées dans cet article est ceci qu'elles ont utilisé, toutes, la « dictée traditionnelle » comme instrument d'observation et ont étudié les « types d'erreurs d'orthographe ». Cette révision sera suivie d'une discussion autour des types d'erreurs observés dans ces deux grands groupes de recherches.

## 2 Recherches en FLM

### 2.1 Roberge : 1961-1982<sup>1</sup>, Canada

Dans le but de comparer les acquisitions orthographiques chez les écoliers canadiens francophones, Roberge (1984) au cours de deux enquêtes dicte un même texte en 1961 et en 1982 à respectivement 426 candidats au certificat d'études primaires et 348 écoliers de la première année du secondaire dans une même région (ville de Sherbrooke et région de l'Estrie). À côté de plusieurs questions secondaires relatives aux différences dans les sexes, dans les types d'erreurs, etc., la question principale reste à savoir si l'orthographe des écoliers québécois est meilleure ou pire qu'avant. Pour l'étude des erreurs d'orthographe de son échantillon, ce chercheur élabore la typologie suivante : 1) erreurs de règle ; 2) erreurs d'usage ; 3) erreurs mixtes (cumul des deux types précédents) et 4) oubli ou ajout de mot. Le résultat final met en évidence une chute verticale de la compétence orthographique entre ces deux époques. Quant aux types d'erreurs, les résultats montrent que les erreurs grammaticales sont plus nombreuses que les erreurs d'usage et celles-ci plus fréquentes que les erreurs mixtes (les erreurs du quatrième type étant négligeables).

### 2.2 Manesse et al. : 2005, France

L'étude comparative de la compétence orthographique des élèves français du XIX<sup>e</sup> et du XX<sup>e</sup> siècle a fait l'objet d'une grande recherche à l'échelle nationale en 1986-1987 (Chervel, Manesse, 1989a, 1989b). Avec le même instrument d'observation, Manesse et al. (2007)

---

<sup>1</sup> La date figurant dans les titres correspond à celle de la réalisation de l'enquête et pourrait être différente de la date de publication des résultats.

lancent une nouvelle recherche en 2005 en s'appuyant cette fois sur les données de l'enquête précédente. Les chercheurs, à côté d'une comparaison entre le niveau orthographique des élèves de 1987 et de 2005, s'intéressent au système des erreurs. L'échantillon de 1987 a été composé de 3048 des élèves du CM2<sup>2</sup> à la 3<sup>e</sup> contre 2767 pour les élèves du même niveau scolaire de 2005. La comparaison des erreurs orthographiques a montré la baisse du niveau en orthographe pour ce second public et ceci malgré l'amélioration de la compétence orthographique de classe en classe dans chacun des échantillons de 1987 et 2005. Quant aux types d'erreurs, les résultats obtenus pour l'échantillon de 2005 sont les suivants<sup>3</sup> :

Types d'erreur	Définitions	%
1	mauvais découpage du mot ; mot sauté ou tronqué ; mot oublié (« *sans force » pour « s'enfoncent »)	4
2	aberration dans la représentation des sons (« *s'ovonce » pour « s'enfoncent »)	5
3	substitution de mot (« *sont font » pour « s'enfoncent »)	3
4	cumul de fautes grammaticale et lexicale (« *s'enffons » pour « s'enfoncent »)	8
5	faute d'orthographe grammaticale ; la catégorie grammaticale n'est pas représentée ou est mal représentée : problèmes d'accord en genre et en nombre, de fautes de singulier et de pluriel, ou d'erreurs dans la conjugaison verbale (« *s'enfonce » pour « s'enfoncent »)	44
6	faute grave d'orthographe lexicale (« *s'enffonssent » pour « s'enfoncent »)	6
7	faute d'orthographe lexicale : forme approchante ou quasi-réussite (« *s'enfoncent » pour « s'enfoncent »)	16
8	faute sur les signes orthographiques (accents, apostrophe, trait d'union, cédille) et les majuscules (« *s'enfoncent » pour « s'enfoncent »)	14
Total		100

Tableau 1 : Répartition des erreurs orthographiques par type, Manesse et al., 2005

### 3 Recherches en FLE

#### 3.1 Ghellai : 1997, Libye

L'objectif principal de la recherche de Ghellai (1997) consiste à analyser la nature des erreurs orthographiques des étudiants libyens en dictée et en production écrite tout en étudiant leurs causes interlinguales et intralinguales d'une part et la représentation que se font ces étudiants et leurs enseignants sur l'orthographe française de l'autre. À cette fin et au moyen de deux dictées, ce chercheur étudie les erreurs orthographiques de 27 étudiants libyens inscrits en première année au département de français de l'Université « el Fateh » de Tripoli. Ces erreurs sont ensuite classées selon la typologie fondée sur la notion de « plurisystème » de l'équipe HESO<sup>4</sup> du CNRS :

<sup>2</sup> Cours moyen deuxième année.

<sup>3</sup> Les graphies erronées seront désormais précédées d'astérisques.

<sup>4</sup> Histoire des Écritures et des Systèmes d'Orthographe.

Types d'erreur	Définitions	%
Phonétiques	(« *cal » pour « car »)	27.3
Phonogrammiques	graphèmes chargés de transcrire les phonèmes (« *gérer » pour « guérir »)	35.5
Morphogrammiques	morphogrammes grammaticaux : marques de féminin/masculin, singulier/pluriel, désinences verbales (« *je fait » pour « je fais »)	17.9
	morphogrammes lexicaux : suffixes, préfixes, radicaux/dérivés (« *anterrement » pour « enterrement »)	7.4
Logogrammiques	homophones hétérographes (« *voix » pour « voie »)	10.2
Lettres non justifiables	lettres historiques et étymologiques (« *téâtre » pour « théâtre »)	1.7
Total		100

Tableau 2 : Répartition des erreurs orthographiques par type, Ghellai, 1997

### 3.2 Katoozian : 2010, Iran

Katoozian (2013) se donne comme objectif principal de recherche la description et l'analyse qualitative et quantitative des types d'erreurs d'orthographe de 313 étudiants iraniens de la première à la quatrième et dernière année de licence de langue française de toutes les universités publiques et privées de Téhéran. L'étude de l'évolution de ces erreurs révèle que les types d'erreurs restent les mêmes de la première à la quatrième année et que leur fréquence diminue mais n'arrive cependant pas à zéro pourcent. Le tableau suivant montre les résultats obtenus quant aux types d'erreur pour l'ensemble de cet échantillon :

Types d'erreur	Définitions	%
1	non-reconnaissance du mot en tant qu'unité lexicale (« *mes ses sair » pour « nécessaires »)	25.6
2	erreur phonétique (« *neceseure » pour « nécessaires »)	22.6
3	substitution de mot (« *sinceres » pour « nécessaires »)	0.7
4a	erreurs de grammaire et de lexique altérant la valeur phonique : cumul des erreurs « type5 » et « type6 » (« *nessaiser » pour « nécessaires »)	1.4
4b	erreurs de grammaire et de lexique n'altérant pas la valeur phonique : cumul des erreurs « type5 » et « type7 » (« *nessessaire » pour « nécessaires »)	4.0
5	orthographe grammaticale (« *nécessaire » pour « nécessaires »)	26.3
6	orthographe lexicale altérant la valeur phonique (« *neccessaires » pour « nécessaires »)	3.3
7	orthographe lexicale n'altérant pas la valeur phonique (« *nessessaires » pour « nécessaires »)	8.9
8a	accents, tréma, cédille (« *neccessaires » pour « nécessaires »)	6.5
8b	apostrophe, trait d'union, majuscule (« *néCessaires » pour « nécessaires »)	0.8
Total		100

Tableau 3 : Répartition des erreurs orthographiques par type, Katoozian, 2010

## **4 Discussion et conclusion**

Dans cette contribution ont été revues quelques recherches descriptives qui se penchaient toutes sur l'étude des erreurs orthographiques et dont l'instrument d'observation était la dictée traditionnelle. Cette révision avait pour objectif principal de connaître les types d'erreurs orthographiques les plus fréquents chez les francophones et les non francophones et ceux fréquents seulement pour ces derniers. En d'autres termes, est-ce qu'on peut parler d'une sorte de « tendance » en ce qui concerne les types d'erreurs orthographiques en FLM et FLE ? Afin de mener cette discussion, les résultats obtenus par les recherches révisées seront rapprochés de la typologie utilisée par Katoozian (2013).

La recherche de Roberge (1984) montre que chez les élèves canadiens francophones les erreurs d'orthographe grammaticale sont les plus fréquentes suivies des erreurs d'orthographe lexicale et des erreurs mixtes de grammaire et de lexique. Les erreurs les plus fréquentes dans la recherche de Manesse et al. (2007)<sup>5</sup> sont dans l'ordre : erreurs d'orthographe grammaticale, erreurs d'orthographe lexicale n'altérant pas la valeur phonique du mot et erreurs d'accentuation. Les erreurs les moins fréquentes sont les erreurs de substitution de mot, les erreurs de segmentation et les erreurs phonétiques. Les résultats de la recherche de Ghellai (1997)<sup>6</sup> révèlent que les erreurs orthographiques les plus fréquentes des étudiants libyens sont les erreurs d'orthographe lexicale altérant ou pas la valeur phonique du mot, les erreurs phonétiques et les erreurs d'orthographe grammaticale. Et finalement, les erreurs d'orthographe grammaticale, les erreurs de non-reconnaissance du mot et les erreurs phonétiques sont parmi les erreurs orthographiques les plus fréquentes dans la recherche de Katoozian (2013)<sup>7</sup>. De l'ensemble de ces constatations, il ressort les remarques suivantes :

- Les erreurs d'orthographe grammaticale semblent être parmi les erreurs d'orthographe les plus fréquentes tant dans le domaine du FLM que dans celui du FLE. Ceci serait en partie imputable à la difficulté de ce secteur de l'orthographe française ;
- Les erreurs orthographiques de type phonétique semblent être plus fréquentes chez les non francophones. Chez les francophones également elles existent mais disparaissent avec l'âge et le niveau scolaire ;
- Il semble que la fréquence de certains autres types d'erreurs d'orthographe varie en fonction du texte de dictée et de la typologie adoptée.

Donc, en ce qui concerne les erreurs orthographiques de type grammatical on peut parler d'une sorte de « tendance » due à la difficulté de ce secteur de l'orthographe française tant pour les francophones que pour les non francophones. Pour ce qui est des erreurs orthographiques de type phonétique, une sorte de « parallélisme » peut être évoqué entre les résultats obtenus par les recherches menées auprès des apprenants du FLE. En effet, il semble qu'il n'existe pas d'erreurs orthographiques essentiellement spécifiques au FLE. C'est dans leur pourcentage que certains types d'erreurs d'orthographe deviennent plus importants pour ce secteur. Plus de recherches sur les types d'erreurs d'orthographe en FLM et FLE (surtout

---

5 Voir Tableau 1.

6 Voir Tableau 2.

7 Voir Tableau 3.

des recherches parallèles avec des méthodologies identiques) sont nécessaires pour mieux élucider la situation et pour mieux répondre à la problématique de cet article.

## Références

- CHERVEL A., MANESSE D. (1989a). *Comparaison de deux ensembles de dictées (1873-1987) : Méthodologie et résultats*. Paris : INRP.
- CHERVEL A., MANESSE D. (1989b). *La Dictée, les Français et l'orthographe 1873- 1987*. Paris : INRP/CALMANN-LÉVY.
- FAYOL M., JAFFRE J.-P. (1<sup>re</sup> éd. 2008). *Orthographier*. Paris : Presses Universitaires de France.
- GHELLAI M. (1997). *Analyse des erreurs et des représentations orthographiques du français par des intellectuels arabophones*. Thèse de doctorat en Linguistique et didactique des langues, Université Stendhal – Grenoble 3.
- JAFFRE J.-P. (1991). Les recherches en didactique de l'orthographe. *Études de linguistique appliquée* 84, 55-63.
- JAFFRE J.-P. (1992). *Didactiques de l'orthographe*. Paris : Hachette Éducation.
- KATOOZIAN K. (2013). *Analyse des erreurs orthographiques des étudiants iraniens en langue française*. Thèse de doctorat en Didactique des langues et des cultures, Université Sorbonne Nouvelle – Paris 3.
- MANESSE D. (2007). La dictée, résistance et avatars d'un exercice scolaire. In H. PEYRONIE, A. VERGNIOUX dir., *Éducation et longue durée*. Caen : PUC, 167- 182.
- MANESSE D. ET AL. (2007). *Orthographe : à qui la faute ?*. Paris : ESF.
- ROBERGE A. (1984). *Étude comparative sur l'orthographe d'élèves québécois*. Québec : Conseil de la langue française, disponible sur : [http://www.cslf.gouv.qc.ca/bibliotheque-virtuelle/publication-html/?tx\\_iggcpplus\\_pi4\[file\]=publications%2Fpubc137%2Fc137ch1.html#table](http://www.cslf.gouv.qc.ca/bibliotheque-virtuelle/publication-html/?tx_iggcpplus_pi4[file]=publications%2Fpubc137%2Fc137ch1.html#table) (consulté le 29 mai 2011).

## **« Non les gens (ils) sont pas dangereux » : variation et redoublement du sujet**

Géraldine Zumwald Küster<sup>1</sup>

(1) Université de Fribourg  
geraldine.zumwaldkuester@unifr.ch

**Résumé** Ce papier s'interroge sur différents aspects de la variation touchant les propositions canoniques du français construites à l'aide d'un sujet SN. Après une brève description de la notion de « variation syntaxique » et de ses diverses interprétations, nous donnons un aperçu des propriétés de l'alternance entre sujet SN simple et sujet SN redoublé, pour souligner finalement les difficultés qui apparaissent lorsqu'il s'agit d'étudier cette variation, notamment à travers des approches quantitatives.

**Abstract** This paper raises some issues about the variation observed within the French canonical clauses composed of DP subjects. First, the notion of “syntactic variation” is looked at, along with its several different interpretations. Then, I expose some properties of the alternation between simple and doubled DP subjects and finally, I show that many methodological problems can arise when it comes to studying this variation phenomenon, especially when using quantitative measures.

**Mots-clés :** double marquage, dislocation, accord sujet-verbe, variation syntaxique, sociolinguistique variationniste

**Keywords:** subject doubling, dislocation, subject-verbe agreement, syntactic variation, variationist sociolinguistics

En français contemporain, les séquences [SNsujet + SV] font l'objet de deux réalisations concurrentes. Soit le SN sujet est directement suivi par le verbe (on parlera de « simple marquage » du sujet) :

- (1) simplement euh + pendant la mise en scène pendant après que enfin tout le t- **tout le monde était** là parce que euh ce qui se passe c'est que + on peut pas euh [oral, CRFP]

Soit le SN sujet est doublé d'un indice personnel (désormais IP) sujet de la série [*je, tu, il, nous,...*] (on parlera alors de « double marquage » ou « redoublement » du sujet) :

- (2) je suis comédien sur un autre spectacle + **tout le monde il est** comme ça + dispatché par rapport à sa profession [oral, CRFP]

La coexistence d'exemples comme (1) et (2), énoncés par un même locuteur lors du même entretien, suggère que la présence vs l'absence d'un IP subséquent au SN sujet peut être envisagée comme un cas de variation : on aurait ainsi, pour la séquence [SNsujet + SV], deux variantes en concurrence :

V1 = SN SV[Ø, V,...] = simple marquage

V2 = SN SV[IP, V,...] = double marquage

La reconnaissance d'un phénomène de variation impliquant V1/V2 et l'étude des modalités de cette alternance posent un certain nombre de questions que ce papier vise à mettre en lumière.

## 1 La variation syntaxique<sup>1</sup> : problématique générale

L'expression « variation syntaxique » renvoie souvent aux travaux variationnistes d'inspiration labovienne, qui ont pour principal objectif de quantifier l'influence de facteurs sociaux ou stylistiques sur l'usage de constructions dites « variantes ». Le cadre théorique de ces travaux a été élaboré dans le domaine de la phonologie, et son extension à des données qui ne concernent pas seulement des phonèmes mais aussi des morphèmes porteurs de sens a causé de nombreux débats (cf. Lavandera 1978, Romaine 1984, Gadet 1992). Au cœur de ces débats, on trouve la notion d'*équivalence* des variantes, posée par Labov (1972) comme condition nécessaire à l'exploitation sociale de la variation, mais critiquée par ceux qui considèrent « qu'à toute différence de forme correspon[d] une différence de sens » (Blanche-Benveniste, 1997, 20).

Suite à ces débats, plusieurs idées ont été proposées pour adapter le cadre théorique labovien à l'étude d'alternances en morphologie ou en syntaxe : Labov a suggéré de faire reposer l'équivalence des variantes sur une identité de sens descriptif uniquement, tandis que d'autres (Lavandera ou Romaine) ont proposé de prendre en compte le sens pragmatique des variantes. Face à ces désaccords, des chercheurs comme Jacobson (1989) ou Winford (1996) ont suggéré que c'était au chercheur de choisir, suivant les propriétés de l'alternance observée, le type de sens (descriptif, pragmatique, connotatif ou autre) sur lequel se base l'équivalence.

<sup>1</sup> Les IP sujets étant assimilés ici à des affixes, l'alternance entre V1 et V2 concerne plutôt la morphologie : nous parlons donc de « variation syntaxique » dans un sens large, pour désigner des alternances qui impliquent des morphèmes grammaticaux et/ou des constructions syntaxiques.

On voit donc que la notion de « variation syntaxique » héritée des travaux de Labov en phonologie connaît des interprétations et des applications diverses. Toutefois, la plupart des études qui traitent de variation syntaxique ont comme intérêt commun l'identification et parfois la quantification des différents facteurs qui peuvent influencer l'usage de l'une ou l'autre des constructions variantes. Suivant les propriétés des variantes et la constante de sens choisie, plusieurs types de facteurs peuvent être pris en compte : si les premières études de Labov se concentraient sur l'influence du contexte linguistique et des caractéristiques sociales des locuteurs, d'autres travaux ont cherché à mettre au jour des facteurs de nature pragmatique liés notamment à la fonction communicative des énoncés ou à l'organisation de l'information. Comme le soulignent Jacobson et Winford, le recours à des outils formels destinés à quantifier l'influence de tel ou tel facteur dans un phénomène de variation donné peut être plus ou moins approprié selon les cas.

## 2 Description des variantes

L'existence même des constructions à sujet redoublé est controversée : pour certains (par ex. Rizzi 1986, De Cat 2007), la présence d'un IP coréférent et cooccurrent avec un SN sujet indique que l'on a affaire à une construction disloquée. Selon cette analyse, lorsqu'un SN sujet est doublé d'un IP, il s'interprète toujours comme un élément syntaxiquement détaché, c'est-à-dire non incorporé à la proposition qui le suit.

Les constructions à SN initial disloqué repris par un IP coréférent sont généralement assimilées à des *thématisations*, c'est-à-dire, à des dispositifs qui signalent explicitement le statut de thème du SN sujet initial (voir par exemple Lambrecht 1981, Creissels 2006). Or, certaines unités comme les SN indéfinis du type *personne, quelqu'un, tout le monde, etc.*, sont incompatibles avec un statut thématique, de sorte qu'elles ne peuvent pas apparaître en position disloquée initiale. Les cas comme (2) *supra* ou (5) *infra*, où un sujet du type *tout le monde* ou *chacun* est doublé de l'indice *il* forcent donc l'analyse des SN indéfinis en termes de sujets intra-propositionnels, attestant de ce fait l'existence des phénomènes de redoublement du sujet (voir Berrendonner 2008 pour une argumentation plus détaillée).

On voit donc que les constructions V1 et V2 peuvent être envisagées comme deux variantes d'une seule et même séquence syntaxique, à savoir l'enchaînement tout à fait ordinaire d'un SN sujet et d'un SV constituant ensemble une proposition (ou l'initiale d'une proposition) canonique. L'identité de contenu descriptif et pragmatique est en principe assurée entre V1 et V2 : en effet, l'IP est redondant par rapport au SN et sa présence ne modifie pas les propriétés référentielles ou informationnelles de la proposition (voir toutefois les exceptions mentionnées plus bas). Nous sommes donc apparemment face à un phénomène de variation au sens le plus « strict » du terme, et il est dès lors pertinent d'envisager son traitement à travers des outils quantitatifs tels que ceux mis en place par les variationnistes laboviens.

L'alternance entre V1 et V2 est susceptible d'être conditionnée par différents critères, les uns ayant trait au système linguistique même et les autres concernant le domaine extralinguistique. Les propriétés oppositives de V1 et V2 sont nombreuses, mais la brièveté de ce papier ne nous permettra d'en exposer qu'un nombre très restreint.

Bien que l'information transmise par l'IP soit en principe redondante par rapport à celle qu'apporte le sujet SN, la variante V2 peut, dans certains cas, être plus informative que V1. Par exemple, lorsque le genre d'un SN sujet n'est pas marqué formellement, la variante V1 (3) peut présenter une ambiguïté que ne partage pas la variante V2 (3'):

- (3) L'artiste utilise tous ses sens.      (3') L'artiste elle utilise tous ses sens.

La variante doublement marquée peut également apporter des informations sur le caractère spécifique ou générique du SN sujet ; ainsi en (4), le SN peut être assimilé à *un camembert* en particulier ou à *du camembert* en général. En (4'), la présence de *il* ou de *ça* permet d'orienter l'interprétation du SN vers un sens plutôt spécifique ou plutôt générique (exemple tiré de Blasco-Dulbecco, 1999) :

- (4) Le camembert sent très fort.      (4') Le camembert il/ça sent très fort.

Parallèlement à ces conditionnements internes au système, il semble que les occurrences de V1 et V2 puissent aussi être influencées par des facteurs de nature extra-linguistique. Ainsi, Nadasdi (2000) a observé, dans un corpus de français parlé ontarien, que la répartition de V1 et V2 était sensible à l'origine sociale des locuteurs : il note 54% d'usage du redoublement chez les locuteurs de la classe ouvrière, contre 30% pour la petite bourgeoisie et 25% pour la moyenne bourgeoisie. Coveney (2003) observe une tendance similaire dans un corpus de français parlé de Picardie.

Le caractère socialement marqué du redoublement du sujet se vérifie aussi dans les textes littéraires. Souvent, les doubles marquages apparaissent dans des contextes bien spécifiques, notamment pour représenter le langage oral spontané des personnages (cf, en (5), la présence d'autres marques d'oral familier telles que *eh ben*, *pis*, *s'rait*, etc.) :

- (5) [Sarah] s'efforçait d'imposer, consensuel pourtant, un point de vue selon lequel eh ben **chacun il** avait qu'à faire comme y voulait et pis comme ça **tout l'monde y** s'rait content, y s'rait content, y s'rait content, y s'rait cont... [Benoziglio, *La voix des mauvais jours et des chagrins rentrés*, 2004]

### 3 Difficultés dans le traitement de la variation

Pour calculer la répartition des variantes V1 et V2, Nadasdi (2000) et Coveney (2003) s'inspirent de la méthodologie labovienne. Cette démarche a l'avantage de mesurer l'influence de divers types de facteurs sur l'alternance V1/V2, mais elle comporte un certain nombre de simplifications et d'approximations qui nuisent à la fiabilité des résultats.

Une première difficulté réside dans la distinction entre les constructions à SN disloqué à gauche et celles à SN redoublé : en effet, si cette distinction s'établit clairement en théorie, elle n'est pas toujours perceptible dans la pratique. Ainsi, de nombreuses séquences qui font apparaître un SN sujet susceptible d'être thématiqué et un IP sujet coréférent peuvent se prêter aux deux analyses. Pour exclure les dislocations et ne garder que les redoublements, Nadasdi se base principalement sur des critères d'ordre prosodique.<sup>2</sup> Or, Avanzi (2012) a montré que la prosodie ne permettait pas d'établir une distinction systématique entre dislocations et doubles marquages.<sup>3</sup> Coveney considère quant à lui que la présence de l'IP sujet est variable

<sup>2</sup> Il classe les suites SN+IP+V comme des dislocations si l'un des trois points suivants apparaît : 1) pause entre le SN et l'IP 2) absence d'enchaînement entre le SN et l'IP 3) accent tonique sur la dernière syllabe du SN.

<sup>3</sup> Son étude montre notamment qu'un enchaînement n'est pas un signe d'intégration syntaxique puisqu'il peut s'observer entre un SN clairement disloqué et un IP subséquent (par ex., entre l'objet *les livres* et le sujet *elle* dans *les livres elle les avait envoyés*). On voit aussi que la présence d'un accent tonique sur la dernière syllabe du SN ne dépend pas seulement de son statut disloqué vs intégré mais aussi du nombre de syllabes dont il est composé.

tant dans le cas des dislocations que dans le cas des doubles marquages, et il prend le parti de ne pas distinguer les deux constructions dans son étude. On voit donc que la difficulté de différenciation entre les structures à SN sujet intégré et celles à SN sujet disloqué pose problème pour la récolte des données servant de base à l'analyse quantitative.

Un autre point difficilement conciliable avec une étude quantitative concerne les occurrences de V2 faisant apparaître un IP *il/elle* et celles faisant apparaître un IP *ce/ça*. Conscients que l'usage de *il/elle* vs *ce/ça* est susceptible de transmettre une opposition du type *spécifique* vs *générique*, et en vertu du principe d'équivalence, Nadasdi et Coveney écartent de l'analyse quantitative tous les cas de SN doublés par *ce* ou *ça*. Cette mesure ne résout cependant pas le problème des alternances entre *il/elle* et  $\emptyset$ , qui peuvent faire l'objet d'une distinction *spécifique* vs *générique*, comme en (4) vs (4'). De plus, les propositions composées d'un SN sujet simple qui, s'il était doublé, le serait probablement par *ce* ou *ça*, posent problème :

- (6) Perso, je ne suis pas un grand pratiquant de planche à voile. Je débute et ne demande qu'à en apprendre davantage. Mais **la planche à voile me plaît bien** et quand on m'a proposé de bosser la dessus, j'ai préféré ça à l'aérodynamique et structure des ailes d'avions. [www.directwind.com]

Si les redoublements par *ce* ou *ça* sont exclus de l'analyse, on devrait s'attendre, parallèlement, à ce que les occurrences du type de (6) le soient aussi. Or, aucune mesure de ce type n'est mentionnée dans les études quantitatives que nous avons consultées. Du reste, même si une telle mesure apparaît comme une contrepartie nécessaire à l'exclusion des redoublements en *ce/ça*, sa mise en œuvre nécessiterait un travail de reconstruction et d'imagination (qu'est-ce qui *aurait pu être dit* mais qui ne l'a pas été ?) qui pourrait affaiblir la valeur scientifique des résultats.

La complexité de la question des redoublements en *il/elle* vs *ce/ça* est accentuée encore par l'existence de cas comme (7), où les deux types d'IP ne semblent pas s'opposer :

- (7) vous savez il nous manquait de tout pendant la guerre et alors je me laissais un peu courtiser, je crois que j'avais treize-quatorze ans, par deux garçons, **un il était** le fils du boucher, donc j'étais sûre toujours de d'avoir de la charcuterie à la maison, et **l'autre c'était** le fils du boulanger, car nous avions des tickets de pain [oral - radio]

Dans cet exemple, *il* et *ce* sont employés dans deux énoncés construits de manière parallèle, et ils accompagnent deux pro-SN de nature identique, qui renvoient à deux référents humains envisagés dans leur individualité. Il semble donc que dans certains contextes tels que (7), *il* et *ce* peuvent alterner sans être associés à des distinctions du type *spécifique* vs *générique*.

Ce bref survol des différents aspects de l'alternance entre V1 et V2 laisse entrevoir la complexité d'un traitement quantitatif de la variation touchant les séquences [SNsujet + SV] et les coûts théoriques qui peuvent en résulter. Cela ne veut pas dire qu'il faille abandonner tout essai de mesure quantitative dans le domaine de la variation entre V1 et V2. Pour mieux comprendre cette variation, il nous semble important de répertorier toutes les propriétés oppositives des variantes à l'intérieur du système linguistique, pour pouvoir ensuite interpréter au mieux les données relatives à leur répartition extra-linguistique. Ainsi, on pourra déterminer dans quelles mesures les variantes veulent *dire la même chose* et dans quelles mesures leur répartition inégale parmi les différents locuteurs et/ou les différents registres est associée à des différences au niveau du sens.

## Références

- AVANZI M. (2012). *L'interface prosodie/syntaxe en français : dislocations, incises et asyndètes*. Bruxelles ; Bern : Peter Lang
- BERRENDONNER A. (2008). « Dislocation et conjugaison en français ». *Cahiers de praxématique* 28, 85-110
- BLANCHE-BENVENISTE C. (1997). « La notion de variation syntaxique dans la langue parlée ». *La variation en syntaxe*. F. Gadet (ed.). Paris. *Langue française* 115, 19-29
- BLASCO-DULBECCO M. (1999). *Les dislocations en français contemporain : étude syntaxique*. Paris : H. Champion ; Genève : Diff. Slatkine
- COVENEY A. (2003). « Le redoublement du sujet en français parlé. Une approche variationniste ». *Structures linguistiques et interactionnelles dans le français parlé*. A.B. Hansen & M.-B. Mosegaard-Hansen (eds.). Copenhague : Museum Tusulanum Press, 110-143
- CREISSELS D. (2006). *Syntaxe générale, une introduction typologique* (vol.2). Paris : Lavoisier
- DE CAT C. (2007). *French Dislocation, Interpretation, Syntax, Acquisition*. Oxford : Oxford University Press
- GADET F. (1992). « Hétérogénéité et variation : Labov, un bilan ». *Langages* 108, 127 p.
- JACOBSON S. (1989). « Some approaches to syntactic variation ». *Language Change and Variation*. R.W. Fasold and D. Schiffrin (eds.). Amsterdam/Philadelphia : John Benjamins Publishing Company, 381-393
- LABOV W. (1972). *Sociolinguistic Patterns*. Philadelphia : University of Pennsylvania Press
- LAMBRECHT K. (1981). *Topic, antitopic and verb agreement in non-standard French*. Amsterdam : J. Benjamins
- LAVANDERA B. (1978). « Where does the sociolinguistic variable stop? ». *Language in society* 7, 171-182
- NADASDI T. (2000). *Variation grammaticale et langue minoritaire : le cas des pronoms clitiques en français ontarien*. Munich : Lincom Europa
- RIZZI L. (1986). « On the Status of Subject Clitics in Romance ». *Studies in Romance Linguistics*. Jaegli O. & C. Silva-Corvalán (eds.). Dordrecht : Foris Publications, 391-419
- ROMAINE S. (1984). « On the Problem of Syntactic Variation and Pragmatic Meaning in Sociolinguistic Theory ». *Folia Linguistica* Vol. XVIII, n° 3-4, 409-438
- WINFORD D. (1996). « The problem of syntactic variation ». *Sociolinguistic variation: data, theory, and analysis*. J. Arnold & al (ed.). Stanford : Center for the Study of Language and Information, 177-192

## Comment traiter la variation dans la communication par SMS ? Le cas de l'interrogation totale

Alexander Guryev<sup>1</sup>

(1) Université de Neuchâtel, projet FNS « SMS Communication in Switzerland » & Université de Paris III Sorbonne Nouvelle  
alexander.guryev@unine.ch

**Résumé** Le français est connu pour la variété de ses formes interrogatives. Dans le seul cas des questions totales, il existe les trois tours classiques *Tu vas au cours? Est-ce que tu vas au cours? Vas-tu au cours?* Une hypothèse courante consiste à dire que chacune de ces structures est porteuse d'une valeur socio-stylistique particulière (familiale, neutre ou soutenue). L'étude présentée ici, fondée sur l'examen des questions totales dans un corpus de SMS, montre cependant que la réalité est plus complexe. Dans le corpus considéré, la production des interrogatives est soumise, selon toute apparence, à des contraintes pragmatiques, découlant du type de communication concerné ; il faut compter en outre avec le figement de certaines structures, et avec l'interférence de facteurs personnels et extralinguistiques.

**Abstract** In French, there are multiple ways to construct the same question: *Tu vas au cours? Est-ce que tu vas au cours? Vas-tu au cours?* A possible explanation would be to consider these variants as expressing different socio-stylistic values (informal, neutral, or elevated). However, the reality seems to be much more complex. This study of yes-no direct questions in text messages shows that there are at least two scenarios. Firstly, the production of structures may be subject to various text messaging constraints. Secondly, in some uses interrogative structures seem to be routinized and their production rather depends on extralinguistic (personal) factors.

**Mots-clés :** SMS, structures interrogatives, variation syntaxique, études variationnelles

**Keywords:** SMS, interrogative structures, syntactic variation, variationist studies

# 1 Introduction

Cet article s'inscrit dans le cadre d'un travail de thèse — sous la direction de Marie-José Béguelin (Université de Neuchâtel) et de Florence Lefeuve (Université de Paris 3 Sorbonne Nouvelle) — ayant pour objet la variation des structures interrogatives dans un corpus suisse de SMS en français.

L'étude des structures interrogatives du français retient depuis longtemps l'attention des linguistes, de par la variété des formes interrogatives à disposition : trois structures pour les questions totales, et jusqu'à 16 formes (*cf.* Quillard, 2000 : 45) pour les questions partielles, et cela sans compter les structures focalisantes ou segmentées (*C'est au cours que tu vas ? T'y vas au cours ?*). Pour la question totale, les structures de base sont les suivantes : (i) maintien de l'ordre 'Sujet Verbe' : *Tu vas au cours ?*, et marquage exclusivement intonatif ; (ii) inversion du clitique sujet, simple ou « complexe » : *Vas-tu au cours ? L'étudiant va-t-il au cours ?* ; (iii) particule 'est-ce que' : *Est-ce que tu vas au cours ?*

Les problèmes posés par cette situation de variation sont principalement les suivants : (i) À quelles causes faut-il attribuer la variété des formes interrogatives en français ? (ii) Les structures en concurrence *Tu vas au cours ?*, *Est-ce que tu vas au cours ?*, *Vas-tu au cours ?* sont-elles à considérer comme des variantes fonctionnellement interchangeables, renfermant le même contenu propositionnel, ou faut-il au contraire y voir des structures particulières, dotées chacune d'une fonction propre ? (iii) Y-a-t-il des corrélations à l'œuvre entre certains facteurs externes et l'occurrence de telle ou telle variante ?

Les structures qui nous occupent ont donné lieu à trois grands types d'approches : les approches générativiste, variationniste et fonctionnaliste. Chacune d'entre elle s'est centrée sur l'aspect privilégié par son programme : les contraintes internes pour la première, les facteurs externes pour la deuxième, et les rapports 'forme/fonction' pour la troisième.

Or, les études de Coveney (2002) et de Quillard (2000) ont montré que le cas des structures interrogatives relève d'une analyse multifactorielle. Leur constat est corroboré par les travaux de Druetta, qui écrit que « l'interrogation est un mécanisme psycholinguistique complexe », et pour qui « une analyse exhaustive de ce mécanisme doit [...] intégrer des composantes de plusieurs niveaux : syntaxiques, prosodiques, sémantiques et pragmatiques » (2009 : 9).

Nos premiers résultats de recherche vont dans le même sens et confirment que plusieurs paramètres sont en jeu. Dans le cadre de ce bref article, nous examinerons plus particulièrement l'influence des contraintes communicatives et des facteurs externes sur l'emploi des trois structures interrogatives totales classiques, sur la base du corpus de SMS suisse en français. Par commodité, les trois types d'interrogatives seront abrégés comme suit : SV (*Tu vas au cours ?*), V-CL (*Vas-tu au cours ?*), et ESV (*Est-ce que tu vas au cours ?*)<sup>1</sup>.

Pour commencer, nous nous intéresserons à la fréquence d'emploi des trois structures, à partir de l'analyse des questions contenues dans les 500 premiers messages du corpus (présenté au § 2.1). Ensuite, nous tenterons de montrer que dans la communication par SMS, les contraintes socio-stylistiques sont secondaires par rapport aux contraintes communicatives. De ce point

---

1 Nous empruntons ces abréviations aux travaux de Coveney (2002), Mosegaard-Hansen (2001) et Quillard (2000, 2001). SV renvoie donc à l'ordre formel 'Sujet Verbe', V-CL à 'Verbe-Sujet Clitique', ESV à 'Est-ce que Sujet Verbe'.

de vue, nous verrons que le choix des variantes est soumis à des contraintes propres au type de communication concerné : le scripteur privilégie la variante qui lui permet de réaliser ses intentions le plus efficacement possible. Cependant, il y a aussi des cas dans lesquels le rôle des facteurs pragmatiques semble être minimal, les structures employées semblant « routinisées » ou en voie de figement.

Étant au début de notre recherche, nous souhaitons, sans prétention à l'exhaustivité, présenter aux lecteurs quelques pistes de réflexion et un aperçu des problèmes posés par les structures interrogatives totales dans les messages SMS.

## **2 Structures interrogatives totales dans le corpus suisse de SMS**

### **2.1 Le corpus suisse de SMS**

Le corpus suisse de SMS comprend 25'947 messages rédigés dans plusieurs langues (essentiellement dans les quatre langues nationales de la Suisse : allemand et dialectes alémaniques, français, italien, romanche), dont 4'624 messages en français. Les données – qui ont été collectées par les Universités de Zurich et de Neuchâtel, sous l'égide du CENTAL de l'Université de Louvain – sont actuellement étudiées dans le cadre d'un projet de recherche intitulé 'SMS communication in Switzerland: Facets of Linguistic Variation in a Multilingual Country'<sup>2</sup>.

### **2.2 Tri des données et résultats provisoires**

Après avoir examiné les 500 premiers messages du corpus, nous avons relevé 123 occurrences de questions totales, avec les fréquences d'emploi suivantes pour les trois types : SV : 78% (96 occurrences) ; V-CL : 17.1% (21 occurrences), et ESV : 4.9% (6 occurrences)<sup>3</sup>.

Notre première tâche a consisté à trier les exemples, afin de ne retenir que les structures véritablement sujettes à variation. Pour ce faire, nous nous sommes basé sur les critères d'interchangeabilité grammaticale et de fréquence. D'une part, les constructions devaient être formellement transformables et grammaticalement acceptables : ainsi, pour SV '*Tu vas au cours ?*', nous avons les commutations possibles V-CL '*Vas-tu au cours ?*' et ESV '*Est-ce que tu vas au cours ?*'. D'autre part, si une construction syntaxique donnée était constamment réalisée avec l'une des structures, nous l'avons considérée comme un contexte catégorique. En conséquence, les constructions suivantes n'ont pas été retenues dans les comptages :

1. ***Sujet pronominal en ça.*** Nos données fournissent 40 occurrences de cette configuration, qui est réalisée 39 fois avec SV *Ça avance ton mémoire ?*, et une seule fois avec ESV (*Est-ce que ça joue ?*). Or, ça ne peut être soumis à l'inversion (*\*Joue-ça*), au contraire des clitiques je, tu, il(s), elle(s), nous, vous, on, ce (Kayne, 1983).

---

<sup>2</sup> Ce projet interuniversitaire Sinergia est soutenu par le FNS et dirigé par Elisabeth Stark, professeur à l'Université de Zurich (<http://www.sms4science.ch>).

<sup>3</sup> Si, pour les besoins de l'analyse qualitative, nous avons emprunté quelques exemples au corpus global, les comptages présentés ici ne concernent que les formes contenues dans les 500 premiers messages.

2. **Sujet pronominal en ce.** Dans nos données, cette construction est toujours produite avec SV : *C'était la même ?* (8 occurrences). Il est certes possible de la transformer en V-CL : *Etait-ce la même ?* ou en ESV : *Est-ce que c'était la même ?* ; cependant, vu l'absence dans le corpus de ces deux dernières structures, nous n'avons pas retenu la construction. Certains auteurs considèrent d'ailleurs que l'usage de *ce* est à peine acceptable dans l'inversion (du moins en langue non littéraire), sauf avec 'est' (Coveney 2002 : 209 cité par Elsig 2009 : 19).
3. **Sujet pronominal en je suivi d'un verbe non-modal au présent.** Nous avons relevé 7 occurrences de ce type de construction, toutes de forme SV *Je viens vous rejoindre ?* À part, comme le remarque Wilmet, avec un verbe courant : *Ai-je ?*, *Suis-je ?*, etc., les usages inversés de cette configuration « frôlent la dérision » : *Aimé-je ?*, et sont « déviants » : *\*Cours-je ? / ?? Couré-je, Pars-je / ?? Parté-je ?* (2007 : 584).

Il faut remarquer que si nous avons choisi de retenir ces usages, la fréquence de l'emploi des questions de type SV (*Tu vas au cours ?*) aurait dépassé 90%.

### 3 Que nous apprend le corpus de SMS ?

Nous avons pour base de données un corpus de SMS, et il est loisible de penser que la performance fait l'objet de contraintes propres à ce type de communication : le scripteur privilégie la structure (ou la variante, parmi d'autres à sa disposition) qui lui permet de réaliser au mieux ses intentions, par le biais d'opérations à la fois économiques et optimales (Frei 1929<sup>4</sup>, Berrendonner 1988, Haspelmath 1999).

Toutefois, il y a des cas où le rôle des contraintes est minimalisé, les structures étant presque ritualisées et figées (Cheshire 2005, Quillard 2001). En ce cas, la répartition des variantes dépend probablement aussi de facteurs d'ordre socio-personnel.

En lien avec ces postulats, nous examinerons d'abord les contraintes de la communication par SMS qui sont susceptibles d'avoir un impact sur la performance du scripteur dans son choix de telle ou telle structure interrogative totale. En second lieu, nous montrerons qu'il peut y avoir des cas où la production des structures interrogatives ne semble pas « générée par la grammaire »<sup>5</sup>, mais tout simplement « routinisée ».

#### 3.1 Contraintes communicatives vs contraintes socio-stylistiques

D'après la répartition socio-stylistique classique (largement exploitée, par exemple, en didactique du FLE), SV *Tu vas au cours ?* serait familier, V-CL *Vas-tu au cours ?* soutenu, et ESV *Est-ce que tu vas ?* plutôt neutre (Coveney 2011 : 124). Or, il n'est pas toujours évident, dans nos données, de mettre en corrélation l'opposition classique « soutenu/non-soutenu » et la variation syntaxique. Nous observons ainsi, dans le corpus suisse de SMS, certains emplois de V-CL — structure réputée être la variante de prestige — qui surviennent dans des contextes informels :

<sup>4</sup> Cité par Gadet (2007 : 76) : « Pour [Frei], la langue en usage obéit à des besoins fondamentaux, ou « constantes du langage » (assimilation, clarté, brièveté, invariabilité, expressivité) ».

<sup>5</sup> Ici, nous utilisons le terme de Cheshire 2005.

1. a. *Hello en fait dem s cous ou pa si oui à quel heur si oui ve tu que je prévienn antoine ciao bonn soiré (12847)*
- b. *Salut Denis Markus,c'est Garold du local de la vallée du gottéron.J'espère que tu vas bien.Je ne sais ps si tu es au courant ms Yannick ns a demandé de bosser un peu pr son parcours ce w-e.As-tu déjà Itache?sinon ca t'ennuie ps si on partage le boulot? (9193)*

De telles données semblent indiquer que les structures interrogatives totales ne sont pas *per se* porteuses d'une valeur socio-stylistique. C'est pourquoi nous nous rallions aux propos de Frei (1929), Berrendonner (1988) et Haspelmath (1999), auteurs selon lesquels la variation servirait au locuteur de ressource d'ordre pragmatique en vue de réaliser ses intentions communicatives, en fonction de ses besoins et des contraintes découlant de la situation de communication. Compte tenu des contraintes en question, les variantes permettraient de parvenir à des opérations moins « coûteuses » et tendanciellement « optimales » :

Berrendonner (1988 : 50) : « Disposer de plusieurs variantes [...], et pouvoir à tout moment choisir entre elles la plus adaptée aux circonstances contextuelles et aux besoins de la parole, c'est bénéficier d'une marge de manœuvre commode, qui permet de rechercher et d'atteindre une certaine optimisation des procédures d'encodage. »

Haspelmath (1999 : 203) : « In language change, variants are created from which speakers may choose. Being subject to various constraints on language use, speakers tend to choose those variants that suit them best. »

De ce fait, nous n'envisagerons pas les variantes de l'interrogation comme des structures sémantiquement (ou pragmatiquement) équivalentes, mais plutôt comme des « formes en concurrence » (Quillard 2000 : 297). Les faits que nous observons dans le corpus suisse de SMS semblent corroborer ce postulat. En effet, la répartition des structures interrogatives totales telle qu'elle se présente dans nos données diffère de la répartition constatée dans d'autres corpus (tableau 1) :

	SV	V-CL	ESV
Corpus de SMS suisse	78 %	17,1% <sup>6</sup>	4,9%
Français parlé dans le corpus de Coveney, 2002	79,4%	-	20,6%
Radio Corpus de Behnstedt <sup>7</sup> , 1973	41%	20%	39%

Tableau 1 : Fréquence des structures interrogatives totales dans 3 corpus différents

On voit notamment que la communication par SMS est loin d'être assimilable à de l'oral conversationnel, bien qu'il s'agisse dans les deux cas d'échanges spontanés et informels. Le texto représente un genre spécifique, avec ses propres contraintes pesant sur la performance.

<sup>6</sup> Nous pourrions nous demander si cette représentation relativement élevée par rapport à d'autres corpus de français informel ne résulte pas des particularités du français pratiqué en Suisse ; cependant, comme le remarque Elsig (2009), l'inversion est devenue un fait rare en français parlé d'Europe.

<sup>7</sup> Cité par Coveney (2002 : 106 ). Le corpus de Behnstedt comprend environ 4000 interviews et conversations enregistrées à la radio.

Selon Anis (2007), les contraintes en question sont économiques, techniques, communicatives, linguistiques et psychosociales.

Ainsi, en lien avec les contraintes techniques, on peut observer que l'échange par SMS n'implique pas de 'feed-back' immédiat, au contraire des échanges oraux ; du point de vue des facteurs communicatifs et psychosociaux, les auteurs de textos sont souvent en relation proche avec leurs destinataires ; enfin, en lien avec les contraintes économiques et linguistiques, ils tendent à être le plus rapide et le plus concis possible. Ceci conduit souvent à la transgression de l'orthographe standard (2) : l'usage de 'esk' pour 'est-ce que', 't la' pour 'tu es là', 'peutu mdir' pour 'peux-tu me dire', 've tu' pour 'veux-tu', etc.

2. a. *Hello!ca va?Esk t'as les bac&love suivants(je viens de finir le 7)?tu pourra mles preter un dc jour stp?jsuis tro pressé dlire la suite!à demain bonne soirée:-) (10320)*
- b. *Mé fiche de Geo son ché toi?t la se soir pr ke je viene les chercheR?*
- c. *Hello, comen va? Jespèr ke ton operation c bil pacé é ke ta pa du alé en pédiatri come moi!lol,ya dé joli pti desin o mur la ba! Peutu mdir kel tail de chemis il te fau pr lé scout? Merci biz a+ (18395d. Hello en fait dem s cous ou pa si oui à quel heur si oui ve tu que je prévienn antoine ciao bonn soiré (12847)*

Nos données laissent à penser qu'il y aurait *grosso modo* deux types de contraintes pesant sur le choix des variantes de l'interrogation totale : (i) des contraintes d'ordre interactionnel et informationnel, et (ii) des contraintes liées aux particularités médiales de la communication par SMS. Toutefois, ces contraintes semblent réciproquement conditionnées : l'écrit est caractérisé par l'absence de 'feed-back' immédiat, ce qui a des effets sur l'état des connaissances partagées et partant sur la structuration informationnelle<sup>8</sup>.

## 3.2 Motivations d'ordre interactionnel et informationnel

D'après Gadet, « l'interrogation constituant un point d'enjeu dans les rapports interpersonnels, elle requiert souplesse et variété de formes » (1997 : 15). L'interrogation est liée de près à l'organisation informationnelle et interactionnelle du discours, aspect auquel Mosegaard-Hansen (2001) et Abeillé *et al.* (2012) ont consacré une bonne part de leurs études respectives.

Suivant en cela les hypothèses de Mosegaard-Hansen (2001), nous considérerons que l'interaction entre le scripteur et le destinataire est sensible à la proximité psychosociale et à l'accès aux connaissances partagées. Ne disposant que de peu d'exemples avec 'est-ce que', nous évoquerons surtout ici les particularités d'emploi de SV et V-CL.

### 3.2.1 Proximité psychosociale

Selon Mosegaard-Hansen, SV serait préféré dans les situations où les interlocuteurs sont « socialement égaux » ou « relativement proches », alors que l'usage de V-CL ou ESV marquerait au contraire une « distance » entre les personnes (2001 : 477). En effet, dans la plupart des échanges par SMS, le scripteur et le destinataire appartiennent au même milieu socio-culturel : ils partagent certaines connaissances, valeurs sociales et références communes, et sont souvent en relation étroite (Anis 2007). Ces facteurs justifient que SV (78 %) domine

<sup>8</sup> Comme le postule Lambrecht, « written language lacks many of the grammatical features which reflect the speaker's beliefs and assumptions about the hearer's state of mind at the time of the utterance » (1994 : 19).

dans les SMS par rapport aux autres variantes, son usage étant de nature à souligner la proximité entre les interlocuteurs<sup>9</sup> :

3. a. *Yop. Tu bouges avec nous demain soir? A+ (9140)*  
 b. *Salut chérie! Tu n'aimes pas les pates aquatiques :)? Mdr (9232)*  
 c. *Yo madame!qu'est-ce qui se passe?tu es malade?tu déprime? (10763)*  
 d. *Tcho,tu bosse demain?ptet je vais a vevey pour fair les mag.tu viens?ou bien un peu de grimpe?bizou (9451)*

En revanche, l'usage de V-CL marquerait plutôt une « distance » inter-individuelle :

4. a. *Bonsoir, des cours sont annulés pour raison d'un conseil de lycée extraordinaire. Je finis donc les cours à 15h15 au lieu de 17h. Est-il possible d'avancer notre rdv? Bonne soirée rick (8310)*  
 b. *Bonjour Clerc, viens-tu au cours de francais aujourd'hui? Nous n'avons PAS de cours le mercredi 11novembre (la semaine prochaine). Cordialement. (8155)*

Nous n'avons pas relevé d'autres exemples de ce type avec V-CL, ce qui découle probablement du fait que la communication par SMS concerne le plus souvent des personnes relativement proches.

### 3.2.2 Accès aux connaissances partagées

Pour mesurer l'influence de l'accès aux connaissances partagées sur la production des structures interrogatives totales, Mosegaard-Hansen (2001 : 474) se réfère au schéma développé par Labov & Fanshel (1977 : 100), où A réfère au locuteur, et B à l'interlocuteur (tableau 2) (notre traduction) :

Type d'évènement	Accès informationnel
A-évènements	connus par A mais pas par B
B-évènements	connus par B mais pas par A
AB-évènements	connus par A et B
O-évènements	connus par toutes les personnes présentes
D-évènements	disputables

Tableau 2 : Accès à l'information des participants dans une interaction

Ainsi, selon Mosegaard-Hansen, SV porterait souvent sur des B-évènements, alors que V-CL et ESV seraient réservés davantage à d'autres types d'évènements. La dominance de SV dans nos données peut découler du fait que les interlocuteurs s'interrogent sur des contenus qui ne sont pas nouveaux pour eux, mais dont la validité est susceptible d'être mise en question au cours de l'interaction, ou nécessite des précisions (demandes de confirmation) :

5. a. *Mdr...alors...donne moi bonne conscience: tu les laisse tomber?? (9492)*  
 b. *Hey mamie jeanrenaud, il fait quel temps en haut??tu prend le bus de 01? Je vais essayer de le choper..mais je promet rien...le train arrive à 58...quand il est à l'heure... (10786)*

<sup>9</sup> Cependant, il ne faut pas perdre de vue les cas où SV s'emploie dans des contextes formels, comme lors de débats politiques, pour des raisons interactionnelles ou informationnelles, et non pour des raisons de proximité : *Non, vous me permettez là-dessus ? Vous les embauchez pour cinq ans ?* [des enseignants] (Débats présidentiels 2012).

c. *Alors tu viens faire ce massage?* (9210)

d. *Hi baby!Sorry I'm in New York City at the moment!Hihi!Je te raconterai!Tu t'arretes à Lsne après tes cours mercredi?Moi j'arrive à 19h15 depuis Neuch!Gros bisou* (10687)

Par ailleurs, dans nos données, tous les emplois du passé composé vont de pair avec la structure SV (17 occurrences), les interlocuteurs s'interrogeant souvent sur des faits qui ont eu lieu, ou devraient avoir eu lieu (6) (cf. la présence de marques temporelles comme 'alors' ou 'déjà' (6b-c)) :

6. a. *Me voilà bien arrivée à luni et motiv pr lire un peu!toi t-a fait bonne impression avec t croissants?;) bonne journée j'taime* (10036)
- b. *Ta déjà fait le gato? Sinon un cake ou d croissant o jambon!* (10327)
- c. *Coucou alors tu as reçu mon âne? Bisoux je tm* (11062)

L'emploi de V-CL signale au contraire que le scripteur ne sait pas si le contenu propositionnel sous la portée de l'interrogation est activé dans la mémoire discursive<sup>10</sup> de l'interlocuteur (7). En ce cas, le recours à V-CL vise plutôt à susciter des informations nouvelles.

7. a. *Hey, comment va? As-tu une adresse e-mail pour que je puisse te répondre? Biz* (9370)
- b. *Hello,comment va l'archi à LS? Connais-tu un étudiant qui serait d'accord de suivre les devoirs de math? Est-ce tes soeurs connaissent qq? Merci d'av. Jojo* (9575)
- c. *Helo il y aura finalement pa dséance ac lé vieux-castel samedi proch. On a réussi a tt planifié samedi.je vous envoie le program 2m1.par contr pr la séance du matin ac lé éclai as tu dé idé dactivité? A+* (10455)
- d. *Salut les ami(e)s. Seriez vous là demain soir pour une bouffe chez nous. J'ai envie de feter avec vous ma récente obtention de mon titre de psychothérapeute. [...]* A+ (9872)
- e. *L un de vous aura t il son ordi avec lui? Sinon j prends le mien, pr vous montrer les fantastiques statistiques de notre site :)* (10807)

En effet, dans (7a-d), l'usage de l'article indéfini permet l'introduction d'un nouvel objet-de-discours, mentionné pour la première fois dans l'interaction. Ceci semble plutôt caractéristique de la structure V-CL (vs SV) : 28,6% des occurrences de V-CL (6 sur 21) contiennent en effet des constructions verbales avec un SN indéfini comme complément, tandis que dans le cas de SV la proportion tombe à 6,25 % (6 occurrences sur 96). Cela laisse à penser que nos scripteurs suisses de SMS recourent préférentiellement à V-CL pour introduire un contenu en tant qu'information nouvelle (7a-d) ou pour (ré)activer un objet dans la mémoire discursive alors que celui-ci n'a pas été récemment évoqué (7e).

Nous sommes conscient que ces explications sont à avancer avec prudence, une partie seulement du corpus ayant été exploré : nos analyses demandent à être confirmées par l'examen du corpus entier et par celui d'autres corpus. Cependant, nous pouvons d'ores et

<sup>10</sup> Voir Berrendonner 2005 pour les explications de la Mémoire discursive et les opérations effectuées par les structures interrogatives totales.

déjà nous appuyer sur Weinrich (1989 cité par Quillard 2000 : 5), aux yeux de qui les structures SV « cherchent à combler une lacune dans l'information préalable », alors que ESV et V-CL « appellent en outre une information supplémentaire ou une prise de position ».

### 3.3 Motivations liées aux particularités médiales

Le choix des structures utilisées dans la communication par SMS est également influencé par des contraintes d'économie, liées aux particularités médiales de ce type de communication. De manière générale, l'écrit semble caractérisé par une volonté du scripteur d'être le plus concis possible (Lambrecht 1994 : 18) ; dans le cas du SMS, il s'y ajoute une contrainte économique liée au nombre limité de caractères par message<sup>11</sup> et, souvent aussi, une contrainte de rapidité dans la rédaction (Anis 2007 : 94).

Dans nos données, et sous bénéfice d'inventaire, il résulte de ce fait que l'inversion simple V-CL semble préférée à ESV<sup>12</sup>, la première permettant un gain de caractères – '*connais-tu un étudiant... ?*' vs '*est-ce que tu connais... ?*' :

8. *Hello, comment va l'archi à LS? Connais-tu un étudiant qui serait d'accord de suivre les devoirs de math? Est-ce tes soeurs connaissent qq? Merci d'av. Jojo (9575)*

Cependant, quand le sujet est un SN, le scripteur de (8) opte pour ESV tout en abrégant l'élément '*est-ce que*'<sup>13</sup> : '*est-ce Ø tes sœurs connaissent qq?*', ce qui représente apparemment (au niveau graphique et pourquoi pas cognitif) une opération moins « coûteuse » que ne le serait l'inversion complexe *tes sœurs connaissent-elles...* . Les deux variantes, que nous retrouvons dans le même message, servent ainsi d'options au locuteur et ne sont pas introduites arbitrairement. Il est remarquable que dans nos données, les constructions avec sujet SN ont été réalisées soit par SV, soit par ESV, évitant aux utilisateurs l'usage de l'inversion complexe :

9. a. *Mé fiche de Geo son ché toi? t la se soir pr ke je viene les chercheR? (6108)*  
b. *Est-ce que A veut venir patiner avec M et p-ê hans-henning? (9571).*

### 3.4 Contraintes externes

Dans certains cas, le choix des variantes ne semble pas particulièrement guidé par des motivations pragmatiques, mais relève selon toute apparence de contraintes externes, d'ordre personnel. Ainsi, dans le corpus étudié par Quillard, l'inversion avec le verbe '*vouloir*' exprime souvent l'offre et intervient dans les contextes informels (Quillard, 2001 : 66) : « Quant aux structures plus "formelles", correspondant aux structures par inversion [...], elles sont présentes dans des contextes jugés "non formels", uniquement pour exprimer des demandes d'acceptation d'offre, qui s'avèrent être des formules figées ».

---

<sup>11</sup> Cela s'applique à la totalité des 500 SMS analysés.

<sup>12</sup> Cependant, nos premières observations montrent que rien n'est catégorique : certains utilisateurs peuvent ne pas faire usage de l'inversion, en optant pour ESV.

<sup>13</sup> Comme on peut s'y attendre, '*est-ce que*' fait l'objet d'une variation graphique importante : '*est ce que*', '*est-ce ke*', '*est ce ke*', '*eske*', '*esk*', '*est-ce*', etc.

Dans nos données, c'est le verbe '*pouvoir*' qui est concerné par le figement, et l'usage de l'inversion ne semble pas refléter forcément la volonté du locuteur d'avoir recours à une structure châtiée. Les structures V-CL semblent en ce cas « quasi-figées » (voir la soudure orthographique dont témoigne (10b)), et la répartition des variantes semble obéir à des préférences personnelles.

10. a. *peu tu traduir en anglai ce ki sui?* (17699)

b. *Peutu mdir kel tail de chemis il te fau pr lé scout? Merci biz a+* (18395)

Autre indice révélateur, près de la moitié des emplois V-CL (47,6 %) concernent le verbe '*pouvoir*', lequel n'entre en revanche que dans 17,7% des occurrences de la structure SV.

Pour montrer que les usages peuvent varier d'une personne à l'autre, nous avons contrôlé 5 profils et examiné les requêtes, demandes d'autorisation ou propositions exprimées dans leurs textos à l'aide de différentes formes du verbe '*pouvoir*' au présent, au conditionnel ou encore au futur simple. Les résultats sont présentés dans le tableau 3, qui intègre quelques données personnelles : âge, sexe et canton d'origine (tableau 3).

	Variantes employées (l'orthographe est conservée) <sup>14</sup>	SV	V-CL	ESV
Age 17, F, Jura	<i>peux-tu... ?</i> <i>peux tu... ?</i> <i>peu tu... ?</i> <i>peutu... ?</i> <i>est ce kon pourrait... ?</i>	0	<b>80% (4)</b>	20% (1)
Age 22, F, Neuchâtel	<i>tu px... ?</i> <i>tu pourrais... ?</i> <i>on peut... ?</i> <i>on peu... ?</i> <i>est-ce que je peux... ?</i>	<b>80% (4)</b>	0	20% (1)
Age 40, F, Neuchâtel	<i>je peux... ?</i> <i>tu pourrais... ?</i> <i>pvez-vs... ?</i> <i>est-ce que je pourrais... ?</i>	60% (3)	20% (1)	20% (1)
Age 14, H, Neuchâtel	<i>tu peux... ?</i> <i>tu pourras... ?</i>	<b>100% (6)</b>	0	0
Age 28, H, Fribourg	<i>puis-je... ?</i> <i>puis je... ?</i> <i>pourrais-je... ?</i> <i>pourrais tu... ?</i>	0	<b>100% (7)</b>	0

Tableau 3 : Réalisation du verbe '*pouvoir*' avec variantes privilégiées par différents auteurs

Ces résultats provisoires montrent que certains auteurs ont tendance à favoriser une des structures possibles avec le verbe '*pouvoir*'. Seule une personne (40, F) alterne les trois variantes, alors que les autres manifestent une préférence pour une seule variante : (14, H) et (22, F) pour SV vs (28, H) et (17, F) pour V-CL. Il arrive ainsi qu'un scripteur choisisse une

<sup>14</sup> Ici, nous n'avons pas mis toutes les occurrences, une même variante (par exemple '*tu peux... ?*') pouvant être employée par une même personne plusieurs fois.

variante et l'applique dans la plupart des situations, qu'il s'agisse de requêtes, d'offres, de demandes d'autorisation, de propositions, etc. À ce propos, nous renvoyons à Cheshire, qui considère ces actes comme des « routines conversationnelles » où les unités ne sont pas générées par la grammaire mais sont plus ou moins « préfabriquées » : « [...] speakers use prefabricated expressions to help them cope with the demands of fast speech production. These expressions include conversational routines with clear social or cultural functions, such as forms conventionally expressing apologies, thanks, compliments or requests » (2005). Nous pouvons en conclure qu'en plus des contraintes pragmatiques, il faut tenir compte de facteurs et de tendances personnelles, qui peuvent être liés aux habitudes des locuteurs.

## **4 Conclusions provisoires**

De cette étude préliminaire, nous tirerons les conclusions provisoires suivantes :

1. S'il y a une nette différence de fréquence pour les trois structures interrogatives totales du français, on peut se demander si elles sont fonctionnellement interchangeables. Les remarques faites par Quillard à ce propos nous semblent très justes : « Plutôt de dire que la variation fait référence à plusieurs formes voulant signifier la même chose, ne suffirait-il pas de proposer que la variation désigne tout simplement plusieurs formes en concurrence ? » (2000 : 297).
2. Le choix entre plusieurs variantes permet au locuteur de mieux s'adapter à ses besoins communicatifs ; il recourt à une variante plus efficace que les autres dans un contexte de communication particulier.
3. Dans certains cas, l'usage des variantes n'est pas fonctionnellement motivé, ce qui, il faut l'admettre, n'est pas toujours facile à vérifier. Le choix d'une variante semble alors essentiellement dicté par des facteurs personnels, et les structures semblent « préconstruites » ou « ancrées dans la mémoire » (voir tableau 3). À ce propos, Cheshire (2005) préconise une collaboration entre les syntacticiens et les variationnistes, en vue de définir des critères permettant de distinguer les structures « préfabriquées » de celles qui sont « générées par la grammaire ».

## **Remerciements**

Je tiens à remercier Marie-José Béguelin, Florence Lefevre, François Delafontaine ainsi que Géraldine Zumwald Küster et les deux experts anonymes pour leurs remarques et suggestions lors de la relecture de cet article. Je demeure responsable des propos qui y sont exprimés.

## **Références**

ABEILLE A., CRABBE B., GODARD D., MARANDIN J.-M. (2012). French questioning declaratives: a corpus study. *Proceedings of SemDial 2012 (SeineDial) : The 16th Workshop on the Semantics and Pragmatics of Dialogue*. Université Paris-Diderot (Paris 7), Paris Sorbonne-Cité, September 2012.

ANIS, J. (2007). Neography – Unconventional Spelling in French SMS Text Messages. In : Brenda Danet and Susan C. Herring (Hrsg.). *The Multilingual Internet – Language, Culture and Communication Online*. New York : Oxford University Press, 87-115.

- BEGUELIN, M.-J. (2012). La variation graphique dans le corpus suisse de SMS en français. In : Sandrine Caddéo, Marie-Noëlle Roubaud, Magali Rouquier, Frédéric Sabio (éds). *Penser les langues avec Claire Blanche-Benveniste*. Presses de l'Université de Provence, 47-63.
- BERRENDONNER, A. (2005). Questions et mémoire discursive. In : Corinne Rossari, Anne Beaulieu-Masson, Corina Cojocariu, Anna Razgoulieva (éds). *Les Etats de la question*. Editions Nota Bene, 147-173.
- BERRENDONNER, A. (1988). Normes et variations. In : *La langue française est-elle gouvernable ?* sous la dir. de G. Schoeni, J.-P. Bronckart, P. Perrenoud. Neuchâtel/Paris : Delachaux & Niestlé, 43-62.
- CHESHIRE, J. (2005). Syntactic variation and spoken language. In : L. Cornips and K. Corrigan (eds.) *Syntax and Variation: Reconciling the Biological and the Social*. Amsterdam : John Benjamins, 81-106.
- COVENEY, A., (2002). *Variability in Spoken French: interrogation and negation*. Bristol : Intellect Books.
- DRUETTA, R. (2011) Les formes interrogatives au début du XXI<sup>e</sup> siècle : évolution ou continuité ? *L'Information grammaticale*, 129, 26-34.
- ELSIG, M. (2009). *Grammatical variation across space and time: the French interrogative system*. Amsterdam : John Benjamins.
- FREI, H. (2003). La grammaire des fautes. Rennes. 1<sup>ère</sup> édition : 1929.
- GADET, F. (1997). La variation plus qu'une écume. *Langue française* 115, 5-17.
- GADET, F. (2007). *La Variation Sociale en Français*. Collection l'Essentiel Français. Paris : Ophrys.
- HASPELMATH, M. (1999). Optimality and Diachronic Adaptation. *Zeitschrift für Sprachwissenschaft* 18, 180-205.
- KAYNE, R. S. (1983). Chains, Categories External to S, and French Complex Inversion. *Natural Language and Linguistic Theory* 1, 107-139.
- LAMBRECHT, K. (1994). *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse*. Cambridge University Press.
- MOSEGAARD-HANSEN, M.-B. (2001). Syntax in interaction. Form and function of yes/no interrogatives in spoken standard French. *Studies in language* 25/3, 463-520.
- QUILLARD, V. (2001). La diversité des formes interrogatives : comment l'interpréter ? *Langage et société* 2001/1, 95, 57-72.
- QUILLARD, V. (2000). *Interroger en français parlé: études syntaxique, pragmatique et sociolinguistique* (thèse de doctorat). Université de Tours, France.
- WILMET, M. (2007). *Grammaire critique du français*. 4e éd. Bruxelles: De Boeck Université.

## ***Je saicherais pas te dire : erreurs et variations flexionnelles dans la structure des paradigmes verbaux français***

Xavier Bach<sup>1</sup> Louise Esher<sup>1,2</sup>

(1) Research Centre for Romance Linguistics, University of Oxford

(2) University of Nottingham

xavier.bach@oxon.org, louise.esher@sjc.oxon.org

**Résumé** Cet article propose une analyse de données sur les erreurs de flexion verbale concernant les bases des tiroirs futur et conditionnel (FUEC). Les données sont issues d'Internet par la recherche de formes erronées surgénérées. Elles révèlent l'absence d'erreurs refaisant le futur et le conditionnel sur la base de l'infinitif lorsque leurs bases sont différentes, ce qui plaide en faveur d'une position périphérique de l'infinitif. Les erreurs sont pour l'essentiel assignées à un modèle flexionnel de classe 1. Un lien est établi entre FUEC et le N-pattern complément, en accord avec les observations proposées par Albright (2002, 2009) pour l'analogie. L'hypothèse est faite que ces réfections demeurent des erreurs par blocage morphologique de l'analogie. Les processus analogiques ne pourraient se produire que si d'autres facteurs de type sémantique ou peut-être phonologique intervenaient.

**Abstract** This article analyses data from errors in verb inflection for the paradigm categories future and conditional (FUEC). The data are drawn from online sources by means of systematic searches for a set of overgenerated erroneous forms. The data show a lack of errors remodelling the future and conditional on the infinitive stem if the FUEC and infinitive stems differ, and thus argue in favour of a peripheral status for the infinitive. In the majority of cases, errors display inflections characteristic of the first conjugation. The study establishes a link between FUEC and the N-pattern complement, consistent with Albright's (2002, 2009) observations on analogy. We advance the hypothesis that these remodellings remain as errors by virtue of morphological blocking of analogy, and that licensing of analogical processes may be partially dependent on semantic or phonological factors.

**Mots-clés :** morphologie, paradigmes, variation, erreurs, flexion

**Keywords:** morphology, paradigms, variation, errors, inflection

## 1 Introduction

Les erreurs de syntaxe ont été dans une large mesure l'objet d'intérêt, particulièrement dans le but de prédire les évolutions futures de la langue (Blanche-Benveniste, 2002). Les erreurs de morphologie ont en comparaison attiré peu d'intérêt. On peut néanmoins relever en français contemporain un certain nombre d'erreurs de flexion du type *je puisserais* ou *tu puisseras*. Ces formes, rares et qui sont à classer dans le domaine de l'erreur, présentent néanmoins d'intéressantes reformatations analogiques sur d'autres bases de leur paradigme, et peuvent nous renseigner sur le fonctionnement des processus analogiques en cours, pour compléter les connaissances acquises à partir d'analogies figées.

Cet article, qui s'inscrit dans le cadre plus large d'une étude en cours sur les erreurs de flexion en français, se limite à analyser les erreurs concernant des formes de futur et de conditionnel, groupés dans un morphème FUEC (Esher, 2012 ; voir aussi Maiden *et al.*, 2011 pour la notion de morphème) correspondant à la fraction de l'espace thématique que partagent de façon systématique ces deux tiroirs en français (Bonami et Boyé, 2002). Le présent travail ne concerne que les alternances de bases, pas les questions de désinences ou d'indexation des arguments du verbe. Il ne concerne que les erreurs sporadiques ; les formes lexicalisées comme les participes *prendu* ou *mourru* sont donc exclues de notre analyse. Le cadre théorique utilisé est celui des théories Mot et Paradigme (MP, Hockett, 1954 ; Aronoff, 1994) et des travaux sur l'organisation morphomique des paradigmes romans (Maiden 2011).

Après avoir explicité la méthodologie suivie pour le recueil des données (2), nous présentons une typologie des erreurs relevées, qui nous amène à examiner les rapports entre FUEC et les autres parties du paradigme (3). Une hypothèse est avancée sur le statut de ces erreurs et leur non-généralisation, qui prend en compte les modèles récents proposés pour l'analogie (Albright, 2002, 2009) comme les travaux sur le morphème FUEC dans les langues romanes (Esher, 2012).

## 2 Méthodologie et recueil des données

Les données analysées ici ont été recueillies sur Internet, en utilisant le moteur de recherche Google paramétré pour ne chercher que des pages en France et en français, sans filtre de sécurité. Les formes utilisées pour la recherche concernent 50 lexèmes verbaux et ont été obtenues au moyen d'une surgénération de formes du futur et du conditionnel sur l'ensemble des bases disponibles dans le paradigme, avec assignation aux quatre principaux groupes, en présupposant la présence d'un segment *-r-* pour ces tiroirs en français (Camussi-Ni, 2006 ; Surcouf, 2012 entre autres). Des formes refaites sur l'infinitif ont également été générées pour tous les lexèmes dont l'infinitif est clairement séparé de la base FUEC (ex. *savoir*, *saurai* > *\*savoirai*), et pour tous les lexèmes des formes refaites sur la base de FUEC avec adjonction d'un segment *-erai*, *-erais*. Près de 9000 formes ont ainsi été générées. Ont été exclus de l'échantillon de lexèmes les verbes ne présentant pas d'alternance de base, ainsi que les verbes présentant une alternance de type vocalique difficilement distinguable à l'écrit (ex. *céder*, *cèderais*)<sup>1</sup>. Les formes surgénérées présentant des formes phonologiquement impossibles (ex.

<sup>1</sup> Les lexèmes inclus dans la base sont : *accroître*, *acquérir*, *agir*, *aller*, *apercevoir*, *apprendre*, *atteindre*, *boire*, *bouillir*, *clore*, *conduire*, *connaître*, *courir*, *craindre*, *croire*, *devoir*, *dire*, *dormir*, *écrire*, *faire*, *falloir*, *finir*, *haïr*, *joindre*, *mettre*, *mourir*, *naître*, *offrir*, *paraître*, *partir*, *perdre*, *pouvoir*, *prendre*, *recevoir*, *réduire*, *rejoindre*, *résoudre*, *retenir*, *réussir*, *satisfaire*, *savoir*, *suffire*, *suivre*, *tenir*, *valoir*, *vendre*, *venir*, *vivre*, *voir*, *vouloir*. Les lexèmes ont été choisis pour représenter chaque classe flexionnelle (exception faite des classes régulières en *-er* qui ne présentent pas d'alternance de base), et en fonction de leur fréquence en français.

\**vdrai*), ou qui sont identiques à une forme d'un verbe existant (ex. *moudrai* généré à partir de *mourir*) sont généralement écartées. En revanche, les formes non distinguables de la forme correcte à l'oral (\**suiverez / suivrez*) ont été conservées comme pouvant indiquer des phénomènes d'assignation de classe. Le *Tableau 1* présente un exemple des formes de futur première personne du singulier pour le verbe *aller* (les formes grisées ont été écartées) :

	<b>er</b>	<b>ir</b>	<b>r</b>	<b>dr</b>
<b>all</b>	alleraï	allirai	allrai/alrai	alldrai/alledrai/aldrai
<b>aill</b>	ailleraï	aillirai	aillrai/alrai	ailldrai/ailledrai
<b>vais</b>	vaiseraï	vaisirai	vaisrai / vairai	vaisdrai/vaidrai/vaiderai
<b>vas</b>	vaseraï	vasirai	vasrai	vasdrai/vasedrai
<b>va</b>	va(t)eraï	va(t)irai	va(t)rai	vadraï/vaderait
<b>v-</b>	veraï / varai	virai	vrai	vdrai

Tableau 1 : Formes générées pour ALLER

Au vu du nombre de formes générées, nous avons choisi de procéder à l'étude d'un échantillon, afin de déterminer les types d'erreurs qui se produisent effectivement. Ne seront donc abordées dans un premier temps que des données quantitatives sur les erreurs concernant les 50 verbes de la base, mais limitées aux cases 1sg.fut et 2pl.cond, recherchées avec leur pronom personnel en séquence stricte sur Google, sur des pages en français et en France, et en excluant les pages contenant le mot « conjugueur », afin de limiter le bruit dans les données. Le choix de ces cases s'est fait sur le critère d'avoir une forme de chaque tiroir, la forme 1s.fut pouvant aussi retourner par erreur graphique des formes de 1s.cond, d'avoir une forme au singulier et une au pluriel, et d'avoir une forme de basse fréquence (2pl.cond). Le nombre d'occurrences dans la recherche Google a été intégré à la base après correction (nettoyage du bruit, erreurs lexicales, apprenants L2 et conjugueurs ou exercices de langue pour apprenants de L2). Google est limité dans l'affichage de 1000 occurrences. Néanmoins, une seule forme ayant atteint ce nombre d'occurrences avant nettoyage des données (*falloir, \*il faura*), nous l'avons codée pour 1000 occurrences et nous avons considéré que cela ne remettait pas en cause la validité des données après avoir comparé les résultats incluant ce lexème et ceux l'écartant de la base. La base comprend ainsi 38.248 occurrences de formes erronées, soit une moyenne de presque 24 occurrences par forme. Plus de la moitié des formes générées n'ont retourné aucune occurrence.

### 3 Une typologie des erreurs sur FUEC

Les erreurs sur FUEC sont pour l'essentiel groupées autour d'un même type, qui implique une assignation quasi générale à la première classe, quelle que soit la base sur laquelle la forme est construite (3.2). Quasiment aucune erreur n'assigne au verbe des modèles qui présentent une dentale avant le *-r-* présent dans ces tiroirs. La plupart des erreurs semblent être liées au N-pattern complément (3.3), et éventuellement au subjonctif (3.4). Les erreurs relevées indiquent par ailleurs une absence de réfection de ces formes sur la base de l'infinitif (3.1).

### 3.1 FUEC et l'infinitif

Historiquement, les formes de FUEC sont issues de périphrases latines incluant une forme infinitive (Lausberg, 1965) :

CANTARE HABET	>	<i>il chantera</i>	(futur)
CANTARE HABEBAT	>	<i>il chanterait</i>	(conditionnel)

En synchronie, un certain nombre d'analyses lient encore FUEC à l'infinitif, de façon plus ou moins étroite (voir Camussi-Ni 2006, 21-26). D'autres travaux tendent à séparer FUEC de l'infinitif, en se fondant sur des données de segmentation rattachant le segment *-r-* à la désinence (Surcouf, 2012 ; cf. Camussi-Ni, 2006, 34-43 pour un sommaire de chaque position). Bonami et Boyé (2002) séparent les cases de l'espace thématique correspondant au futur et au conditionnel de l'infinitif, sur le critère qui distingue les espaces thématiques, à savoir que certains verbes ont deux bases différentes pour ces ensembles de formes.

Maiden (2011) indique que l'infinitif tend à être considéré comme non-segmenté par les locuteurs, du fait de la faible récurrence de finales identiques d'une part, et surtout de l'absence d'association possible entre une désinence donnée (ou un choix de base) et un ensemble de traits morpho-syntaxiques. Si FUEC et l'infinitif étaient liés en synchronie, on devrait voir apparaître un certain nombre d'erreurs produisant des formes de FUEC sur la base de l'infinitif, du type *j'avois* ou *il fallois*, en prenant l'infinitif comme base, ou *je pourrais* s'il s'agissait de la base de l'infinitif<sup>2</sup>. Il apparaît qu'aucune erreur n'est produite dans ce sens, et qu'on ne trouve pas non plus de réfections de l'infinitif sur la base de FUEC<sup>3</sup>.

	base Inf	base FUEC
<i>1s.fut</i>	132	1.763
<i>2p.cond</i>	24	301
<b><i>cumul</i></b>	<b>156</b>	<b>2.064</b>

Tableau 2 : Nombre d'erreurs FUEC sur la base de l'infinitif et sur celle de FUEC (lexèmes ayant deux bases différentes)

Exception faite de *mourir* et *courir* pour lesquels les erreurs sont mieux analysées comme une incertitude de classe flexionnelle (3.2), les erreurs de formes de FUEC refaites sur un infinitif n'ayant pas de base commune avec FUEC (19 lexèmes) ne représentent qu'à peine 0,5% des erreurs relevées. Ces verbes présentent une moyenne de 6 erreurs par forme, ce qui est tout à fait négligeable au vu du corpus. Le *Tableau 2* donne quelques indications sur la proportion d'erreurs relevées entre les erreurs refaisant une forme de FUEC sur une base FUEC existante en ajoutant une désinence *-erai*, *-erais*, et les formes refaites directement sur infinitif.

<sup>2</sup> Il est difficile de considérer ce cas, puisqu'aucun verbe n'a de base spécifique à l'infinitif en français. L'analogie pourrait se faire sur la base d'un autre tiroir.

<sup>3</sup> Les seules formes du type *avois* sont le fait d'enseignants qui craignent de les voir apparaître dans les copies de leurs élèves, d'apprenants non-natifs, et de création que l'on peut sans ambiguïté assigner au registre humoristique, comme dans le cas du gnome Halumgalom, un avatar de jeu vidéo en ligne dont la langue se caractérise entre autres par une base verbale unique, l'infinitif.

Il semble donc que les erreurs sur FUEC soient dans une plus grande mesure sur la base même de FUEC que sur l'infinif. Par ailleurs, une large majorité des erreurs sur la base de l'infinif concernent le seul lexème *voir* (1sg.fut 98, 2pl.cond 15 ; soit 113 erreurs sur base infinitif<sup>4</sup> pour 101 erreurs sur base FUEC pour ce lexème, qui représente donc 72.4% des erreurs sur l'infinif).

### 3.2 Variation de classe : absence d'erreurs en *-drai, -traï, -irai*.

Une grande majorité des erreurs relevées consiste à assigner le verbe en question à la première classe pour FUEC, quel que soit le choix de base qui est fait. Le *Tableau 3* présente la répartition en classe des erreurs, en pourcentage du total d'erreurs. La base de données ne comprenant que 8 verbes du deuxième groupe et aucun du premier groupe, les résultats en sont d'autant plus frappants :

-er-	-ir-	-r-	-dr-	inf
78,61 %	4,47 %	13,82 %	2,70 %	0,40 %

Tableau 3 : Répartition des erreurs par classe

Il semble donc qu'on assiste à une sorte de régularisation générale sur la première classe (Frei, 1929), sauf partiellement dans le cas de verbes du deuxième groupe. Les formes erronées se voient donc prioritairement assignées aux classes régulières (Bonami *et al.*, 2008). En effet, les lexèmes présentant une séquence *-dr-* dans les erreurs sont des verbes qui présentent cette séquence dans leur forme régulière : la majeure partie de ces erreurs concernent le lexème *vouloir*, et dans une moindre mesure *prendre* et *valoir*. Le nombre d'erreurs semble cohérent entre les deux cases du paradigme systématiquement recherchées pour les formes assignées à la première classe. Une explication partiellement phonologique, portant par exemple une insertion possible et régulière d'un schwa épenthétique comme alternative à la diérèse dans une séquence problématique comme /vrje/ au 2pl.cond (ex. *suivre*, /sqi'vrje/) est donc à exclure comme explication unique, puisque le contexte du futur est lui non-problématique de ce point de vue. Le cas des verbes sans voyelle thématique et du deuxième groupe demande un traitement plus poussé.

La plupart des occurrences d'erreurs ne présentant pas de voyelle thématique pourraient être considérées comme un assignement par défaut à la première classe, à cause de la position particulière du schwa en français contemporain. Les verbes de la première classe sont en effet, sauf partiellement dans le sud de la France, quasi exclusivement prononcés sans schwa au futur et au conditionnel (ex. *manger*, *je mangerai* /ʒə-mãʒə're/ [ʒmãʒ're]), à l'exception des 1pl+2pl.cond/fut sur la plupart des bases à finale en groupe consonantique (Cote, 2000).

Le pourcentage de lexèmes assignés à la deuxième classe dans les erreurs est relativement bas, et plus encore quand on examine de près les lexèmes touchés. Plus de 95% des erreurs concernent *mourir* et *courir*, plus la case 2pl.cond de *devoir*. Il est probable que dans ce

<sup>4</sup> Selon la segmentation adoptée, ces erreurs pourraient être attribuées à une réfection sur la base *voi-* du présent.

dernier cas une explication partiellement phonologique soit envisageable<sup>5</sup>. Il est aussi à noter qu'un plus grand nombre d'erreurs assignent cette case au premier groupe pour *devoir* (875 erreurs au premier groupe contre 257 au deuxième). Nous pensons que le cas de *mourir* et *courir* ressort plus d'une variation dans la classe flexionnelle suivie pour ces verbes sur le morphème FUEC que d'erreurs, et qu'il faut donc considérer qu'un nombre infime d'erreurs assignent une deuxième classe à des verbes qui ne remplissent pas les critères d'assignation généralement définis pour cette classe (Bonami *et al.*, 2008). Le *Tableau 4* donne le nombre corrigé d'erreurs relevées dans notre base pour ces deux verbes sur FUEC et sur chacun des deux tiroirs Futur et Conditionnel, pour des désinences en *-erai*, *-irai*, *-drai*, et les formes standard, et pour toutes les personnes avec leur pronom personnel (*il* et *elle* pour les troisièmes personnes)<sup>6</sup> :

		-er-	-ir-	-dr-	std
FUEC	<i>mourir</i>	3.594	4.510	0	13.636
	<i>courir</i>	4.122	2.155	0	14.400
Futur	<i>mourir</i>	2.375	3.065	0	6.613
	<i>courir</i>	2.615	1.415	0	6.785
Cond	<i>mourir</i>	1.219	1.445	0	7.023
	<i>courir</i>	1.507	700	0	7.615

Tableau 4 : Nombre d'erreurs de classe pour mourir et courir (pour les 12 cases de FUEC)

Le nombre équivalent et parfois supérieur de réfections en *-erai* indique que la réfection en *-irai* est pareillement un phénomène de classe flexionnelle, et qu'il ne s'agit pas d'une réfection sur l'infinitif (3.1). Le nombre d'erreurs est aussi nettement plus important pour le tiroir futur que pour le tiroir conditionnel, ce qui tendrait à indiquer que ces lexèmes doivent être traités à part. Le fait que le verbe *mourir* présente 8.104 formes erronées de FUEC contre seulement 13.636 formes standard plaide pour traiter ces formes comme des variantes courantes. Par comparaison, *résoudre*, fréquemment décrit par les locuteurs comme un lieu majeur d'incertitude de forme, ne présente que 37 erreurs du type *je résolverai* pour 887 formes justes en *résoudrai*.

### 3.3 FUEC et N-pattern ou N-pattern complement ?

Le morphème dit N-pattern (Maiden 2005, 2009, 2011) fait partie des morphèmes les plus robustes dans les langues romanes. Il existe en général sous au moins deux formes en français : une qui oppose l'ensemble 1sg+2sg+3sg+3pl.prs.ind/subj au reste du paradigme (ex.

<sup>5</sup> Une explication phonologique par des processus appliqués à la séquence problématique /rj/ n'est pas envisageable pour *courir* et *mourir*, puisque les erreurs sont distribuées sur toutes les cases, y compris celles qui ne contiennent pas de yod.

<sup>6</sup> Le nombre d'occurrences pour la forme standard est à prendre avec quelques précautions. Pour *mourir*, une seule forme a dépassé la limite de 1000 résultats imposée par Google. Pour *courir* en revanche, 7 formes ont dépassé ces 1000 résultats, dont 5 pour le conditionnel. Les données présentées pour *courir* au conditionnel sont donc à prendre avec réserve ; les autres données sont statistiquement valides. Le tableau ne présente que les réfections sur une base *cour-* et *mour-*.

*Je sacherais pas te dire : erreurs et variations flexionnelles dans la structure des paradigmes verbaux français*

*mourir*, base en *meur-*), et une où ces formes présentent une base différente à l'indicatif et au subjonctif, distinctes du reste du paradigme (ex. *aller*, formes supplétives en *v-*/base en *aill-*, reste du paradigme en *all-*).

Le N-pattern est cependant moins répandu en français que dans les autres langues romanes. De nombreux verbes présentant en ancien français une alternance vocalique de type N-pattern se sont vu régulariser, et ne présentent aujourd'hui qu'une seule base (ex. *trouver*). Plusieurs verbes ont au présent une distinction singulier/pluriel plutôt que N-pattern (ex. *savoir*, *finir*). Le N-pattern lui-même existe souvent sous une version dite hybride (Smith 2011), où, au présent toujours, les formes du singulier présentent une base distinctive, les 1pl et 2pl en présentent une autre, et celle du 3pl regroupe des éléments de ces deux bases (ex. *boire* : *boi-*, *buv-*, *boiv-*).

La soustraction des cases constitutives du N-pattern au paradigme dessine le morphème appelé N-pattern complément (Maiden, 2005 et 2011), dont la forme fluctue selon celle du N-pattern. Peuvent également être soustraits au N-pattern complément le morphème PYTA (Maiden 2005), représenté en français par le passé simple et l'imparfait du subjonctif, et FUEC lorsque ces morphèmes présentent une base distinctive. Le cœur stable du N-pattern complément est ainsi composé de l'imparfait de l'indicatif, associé aux 1pl et 2pl du présent de l'indicatif ; il est parfois rejoint par les 1pl et 2pl du présent du subjonctif, voire du présent du subjonctif tout entier.

Pour certains lexèmes, le radical présent dans FUEC correspond à celui du N-pattern : c'est notamment le cas des verbes à alternance vocalique *é/è*, tels que *lever*, *jeter*, etc. (voir ex. Bonami *et al.*, 2008). Dans l'histoire du français, l'identité entre ces deux radicaux a motivé un certain nombre de réfections analogiques, auxquelles sont dues entre autres les formes modernes *tiendra*, *viendra*, *boira* (Martin Maiden, p.c. ; Price, 1998). Nous pourrions ainsi nous attendre à trouver des erreurs où la base de FUEC est remplacée par une base construite sur le radical du N-pattern.

On constate au contraire une basse fréquence de ce type d'erreurs pour les verbes dont les formes standard de FUEC ne partagent pas déjà la base du N-pattern, comme pour ceux qui la partagent déjà. La plupart des erreurs se produisent sur le N-pattern complément. Le *Tableau 5* indique le nombre d'erreurs de chaque type, à l'exclusion des réfections sur une base FUEC pré-existante :

	base standard de FUEC		nouvelle base pour FUEC	
	N-pattern	N-p. complément	N-pattern	N-p. complément
occurrences	734	10.512	1.660	4.305
lexèmes	10	19	11	30

Tableau 5 : Erreurs de FUEC sur le N-pattern et le N-pattern complément

Onze lexèmes sont touchés par des erreurs refaisant FUEC sur une base de N-pattern que les cases constitutives de FUEC ne partagent pas dans les formes standard. Néanmoins, la quasi-totalité de ces erreurs touche *faire* avec 1453 erreurs sur 1660, soit environ 87,5%. Ce cas est problématique car les graphies <*fai*>- et <*fe*>- sont ambiguës pour ce lexème, ne permettant pas vraiment de distinguer des erreurs de graphie des erreurs de base (ex. *nous faisons* [nu-

fə'zɔ̃] ou [nu-'fzɔ̃], *vous faites* [vu-'fɛt]). Exception faite de ce cas problématique, le nombre d'erreurs refaisant FUEC sur une base de N-pattern est de 107 sur 10 lexèmes, donc extrêmement rare. Par ailleurs, seuls quelques lexèmes présentent des erreurs de FUEC basées sur le N-pattern quand ce n'est pas le cas dans la forme standard, alors que la quasi-totalité des lexèmes présentent des erreurs basées sur le N-pattern complement, que FUEC partage la même base dans les formes standard ou pas. Il semble donc qu'en synchronie le lien FUEC/N-pattern soit plus faible que par le passé, et que les erreurs tendent à aller vers la base du N-pattern complement. Ce type d'erreur était déjà mentionné comme prévalent par Frei (1929, 168-9).

### 3.4 FUEC et le subjonctif

Un certain nombre de lexèmes qui présentent la rare particularité d'avoir une base distinctive pour le présent du subjonctif, soit en totalité (ex. *pouvoir*, *savoir*), soit sur les cases correspondant au N-pattern (ex. *vouloir*, *aller*, *boire*, *devoir*) présentent des erreurs où FUEC est refait sur des bases présentes au subjonctif. Néanmoins, lorsque ces verbes présentent une distribution de type N-pattern, c'est toujours sur le N-pattern complement que ces erreurs sont formées. Dans tous les cas le nombre d'erreurs de ce type reste restreint (ex. 57 sur *savoir*, 70 sur *pouvoir*).

Dans ces conditions, il ne semble pas envisageable de postuler une possible réfection due à des valeurs sémantiques et des emplois syntaxiques partagés entre le subjonctif, le futur et le conditionnel. Le subjonctif présent est parfois inclus dans le N-pattern complement, ne serait-ce que pour les 1pl et 2pl. Il s'agit donc d'erreurs spécifiques à un nombre restreint de lexèmes qui ont un nombre important de bases (7 bases pour *pouvoir*, 5 pour *savoir*). Il s'agit d'un type d'erreur nettement moins probable, qui est proportionnellement moins fréquent que les autres types d'erreurs pour les lexèmes touchés (ex. *pouvoir*, 196 N-pattern, 425 N-pattern complement, 70 Subjonctif, 0 Infinitif et 104 FUEC).

## 4 Hypothèse : un modèle multi-facteur pour l'analogie

Il a depuis longtemps été noté que la fréquence de type d'un schéma morphomique (base ou classe) compte plus pour un modèle de l'analogie que la fréquence d'occurrences (Bybee, 1995 ; Albright, 2002, 2009). Les types d'erreur relevés en français sur FUEC concordent avec cette analyse. La première classe a une fréquence de type bien plus importante que les autres classes, ce qui explique donc la généralisation des erreurs sur ce modèle. De la même façon, le nombre de cases présentant une base de N-pattern complement est toujours supérieur au nombre de cases touchées par le N-pattern, et explique la nette préférence pour la reformation de FUEC sur le N-pattern complement dans les erreurs.

Ce que ces données quantitatives n'expliquent pas, c'est pourquoi ces lexèmes ne présentent pas une tendance à développer une base unique, ou du moins à réduire l'idiosyncrasie de la base de FUEC lorsqu'elle est différente. Il est probable que la fréquence d'occurrence joue un rôle dans la résistance d'une base distincte pour des verbes comme *aller*. Mais le même constat ne peut être fait pour des verbes comme *résoudre*, qui manifestent à la fois une certaine insécurité des locuteurs quant à la base à choisir pour FUEC, et une nette préférence pour les formes standard en *résoudr-* sur FUEC. Notre hypothèse, qui sera testée dans les prochains mois, est que plusieurs facteurs entrent en ligne pour autoriser un changement analogique, au nombre desquels sont les structures morphomiques et un probable renforcement de celles-ci par des questions phonologiques ou sémantiques. En effet, hormis les cas d'asymétrie entre les bases du futur et du conditionnel (Esher, 2012), le morphème

FUEC semble bien former un bloc homogène pour les changements analogiques. Les réformations analogiques qui ont touché *tenir*, *venir*, *boire* et *croire* vers le XIV<sup>e</sup> siècle ont ainsi touché FUEC dans son ensemble (Price, 1998, 202-3). Il est donc probable que le poids de FUEC en terme de fréquence de type soit à mettre en relation avec celui des autres bases sur lesquelles une analogie pourrait s'établir. De la même façon, des conditionnements sémantiques ou fonctionnels pourraient intervenir si des formes de FUEC et des formes d'autres tiroirs ayant des valeurs sémantiques partagées peuvent être utilisées dans des contextes identiques, par exemple le conditionnel et l'imparfait de l'indicatif sur les hypothétiques introduites par *si* (Poplack, 2001). Enfin, l'absence de préférence nette pour le choix de la base de FUEC dans le français standard (parfois base commune au N-pattern, parfois au N-pattern complément, parfois spécifique, parfois partagée avec tout le reste du paradigme) limite probablement la possibilité d'une analogie fondée sur un type unique de distribution.

## 5 Conclusion

L'analyse de la production d'erreurs de bases sur FUEC en français permet de révéler ou de confirmer un certain nombre de liens intraparadigmatiques. Le statut de l'infinitif est ainsi clairement séparé de FUEC en synchronie. Il semble aussi que les locuteurs tendent à produire des formes erronées qui ont toutes la même classe pour FUEC (schwa en voyelle thématique). Les verbes *courir* et *mourir* sont eux l'objet, sur ce morphème, d'une variation majeure entre les trois classes qui indique probablement un changement en cours. Contrairement à ce qui a pu se produire dans l'histoire du français, le lien entre N-pattern et FUEC ne semble plus tenir en synchronie. De façon cohérente avec les modèles récents d'analogie (Albright, 2002 et 2009), les erreurs produites sur des formes FUEC refont très majoritairement ces formes sur une base de N-pattern complément, qui a une plus forte fréquence de type. Ces éléments n'expliquent cependant pas pourquoi ces erreurs, dont certaines sont présentes en français depuis longtemps (Frei, 1929) ne se développent pas en innovation, et semblent indiquer que FUEC n'est pas pour l'instant susceptible de réformations de base majeures. C'est là ce que la suite de ce travail s'attachera à établir, par le développement d'une analyse multi-facteur pour l'analogie.

## Remerciements

Nous remercions Martin Maiden et deux relecteurs anonymes. Toute erreur reste de notre fait.

## Références

- ALBRIGHT A. C. (2002). *The Identification of Bases in Morphological Paradigms*. Unpublished Ph.D. thesis. University of California, Los Angeles.
- ALBRIGHT A. C. (2009). Modeling analogy as a probabilistic grammar. In BLEVINS J., BLEVINS J. (2009). *Analogy in Grammar*. Oxford : O.U.P., 185-213.
- ARONOFF M. (1994). *Morphology by Itself. Stems and Inflectional Classes*. Cambridge (MA) : M.I.T. Press.
- BLANCHE-BENVENISTE C. (2002). Quel est le rôle du français parlé dans les évolutions syntaxiques ? *L'Information grammaticale* 94, 11-17.

- BONAMI O., BOYÉ G. (2002). Suppletion and dependency in inflectional morphology. *Proceedings of the 8<sup>th</sup> International HSPG Conference*, 51-70.
- BONAMI O., BOYÉ G., GIRAUDO H., VOGA M. (2008). Quels verbes sont réguliers en français ? In DURAND J., HABERT B., LAKS B. (EDS.) *Actes du Congrès Mondial de Linguistique Française – CMLF'08*, 1511-1523.
- BYBEE J. (1995). Regular Morphology and the Lexicon. *Language and Cognitive Processes* 10, 425-455.
- CAMUSSI-NI M.-A. (2006). *Analyse formelle et conceptuelle des formes verbales du français contemporain : A la croisée du passé simple et de l'imparfait, du futur et du conditionnel, les concepts « ±potentiel » et « ±défini »*. Thèse de Doctorat, Université de Rennes.
- COTE M.-H. (2000). *Consonant cluster phonotactics: a perceptual approach*. Unpublished Ph.D. thesis, M.I.T. Retrieved online <http://roa.rutgers.edu/article/view/558>
- ESHER L. (2012). *Future, Conditional, and Autonomous Morphology in Occitan*. Unpublished D.Phil. thesis, University of Oxford.
- FREI H. (1929). *Grammaire des fautes*. Genève : Slatkine.
- HOCKETT C. (1954). Two models of grammatical description. *Word* 10, 210-234.
- LAUSBERG H. (1965). *Lingüística románica*. Madrid : Gredos.
- MAIDEN M. (2005). Morphological autonomy in diachrony. *Yearbook of Morphology 2004*, 137-175.
- MAIDEN M. (2009). From pure phonology to pure morphology the reshaping of the Romance verb. *Recherches linguistiques de Vincennes* 38, 45-82. Retrieved online <http://rlv.revues.org/1765#text>
- MAIDEN M. (2011). Morphological persistence. Morphological innovation. In MAIDEN M., SMITH J.-C., LEDGEWAY A. (2011). *The Cambridge History of the Romance Languages*. Cambridge : C.U.P., 155-215 et 216-267.
- MAIDEN M., SMITH J. C., GOLDBACH M., HINZELIN M. O. (ÉDS.) (2011) *Morphological Autonomy. Perspectives from Romance Inflectional Morphology*, Oxford : O.U.P.
- PRICE G. (1998) *The French language: present and past*. London : Grant & Cutler.
- POPLACK S. (2001). Variability, frequency and productivity in the irrealis domain of French. In BYBEE J., HOPPER P., ÉDS. *Frequency Effects and Emergent Grammar*. Amsterdam : Benjamins, 405-428.
- SMITH J. C. (2011). Variable Analyses of Verbal Inflection in (mainly) Canadian French. In MAIDEN M., SMITH J. C., GOLDBACH M., HINZELIN M. O. (ÉDS.) *Morphological Autonomy. Perspectives from Romance Inflectional Morphology*, Oxford : O.U.P., 311-326.
- SURCOUF C. (2012). « Agisseriez-vous de même avec des garçons ? » : de quelques (ir?)régularités dans la formation du Futur et du Conditionnel. Poster aux *Décembrettes 8*, Bordeaux, 6 décembre 2012.

## **Neoclassical compounds and language registers**

Marine Lasserre<sup>1</sup>

(1) CLLE-ERSS & Université de Toulouse  
marine.lasserre@univ-tlse2.fr

**Résumé** Les composés néoclassiques occupent une place à part dans la morphologie du français. Quand ils mettent en jeu des éléments finaux néoclassiques récurrents, ils sont à la frontière de la composition et de la dérivation. S'ils étaient au départ utilisés pour désigner des concepts scientifiques ou techniques, ils sont aujourd'hui pleinement utilisés dans la langue courante. Il s'agit ici de voir comment ces différents registres de langue interfèrent dans la construction non seulement formelle mais aussi sémantique des mots complexes comportant des éléments néoclassiques.

**Abstract** Neoclassical compounds are special in French morphology. With regard to their final elements, they are at the border of composition and derivation. If originally they were used to designate scientific or technical concepts, today they are commonly used in everyday language. In this paper, we will see how these different language registers come into play in the construction of these words containing neoclassical elements, both from a formal and from a semantic point of view.

**Mots-clés :** Morphologie, Composition néoclassique, Registres de langue, Corpus, Analogie, Sémantique, Grammaticalisation.

**Keywords:** Morphology, Neoclassical compounding, Language registers, Corpus, Analogy, Semantics, Grammaticalization.

# 1 Introduction

The purpose of this presentation is to study some French neoclassical compounds which apparently display the same behavior. These compounds are built with final neoclassical elements and enter in the construction of large series of lexemes. Consequently, we consider that the final neoclassical element is undergoing grammaticalization, and will eventually become a suffix.

In this article, we adopt the framework of lexeme-based morphology ((Boyé *et al.*, 2011) among others). In this framework, the basic unit of morphology is the lexeme, which is seen as a unit having three properties: a phonological form, a semantic content and a grammatical category.

After defining what is a neoclassical compound (1.1) and presenting our corpus (section 2), we will present two different case studies: first, a morphophonological study (section 3) which shows how neoclassical compounds from different registers are formally treated in the same way, and second, a morphosemantic analysis (section 4) showing how non lexicalized words produced by native speakers allow us to refine the semantic models built on the basis of lexicalized compounds.

## 1.1 What is a neoclassical compound?

Traditionally, native compounds are defined on the basis of a main property: they result from the combination of two (or more) native lexemes. Thus, we have something like [Lexeme + Lexeme]<sub>Lexeme</sub>. In French, we find four productive patterns of native compounds (Villoing, 2012) : (1) [NN]<sub>N</sub>, (2) [VN]<sub>A/N</sub>, (3) [AN]<sub>A</sub> and (4) [AA]<sub>A</sub>.

- |     |              |              |                         |
|-----|--------------|--------------|-------------------------|
| (1) | poisson-chat | ‘fish-cat’   | <i>catfish</i>          |
| (2) | ouvre-boîte  | ‘open-can’   | <i>can-opener</i>       |
| (3) | rouge-brique | ‘red-brick’  | <i>red like a brick</i> |
| (4) | aigre-doux   | ‘sour-sweet’ | <i>sweet-and-sour</i>   |

The particularity of Romance languages is that the head of the compound is on the left.

Neoclassical compounds are different (cf. (Bauer, 1998; Iacobini, 2004) among others). First, they imply at least one Greek or Latin constituent which is no longer autonomous in modern languages, here in French. Moreover, the order of the constituents is reversed as compared to Romance native compounds. (Villoing, 2012) found seven regular patterns of neoclassical compounding in French: (5) [NN]<sub>A</sub>, (6) [NN]<sub>N</sub>, (7) [AN]<sub>A</sub>, (8) [AN]<sub>N</sub>, (9) [NA]<sub>A</sub>, (10) [AA]<sub>A</sub> and (11) [NV]<sub>A/N</sub>.

- |     |                |                      |
|-----|----------------|----------------------|
| (5) | androgynie     | <i>androgynous</i>   |
| (6) | vélodrome      | <i>velodrome</i>     |
| (7) | macrocéphale   | <i>macrocephalic</i> |
| (8) | microorganisme | <i>microorganism</i> |

- (9)    photosensible            *light-sensitive*  
(10)    médico-social            *medico-social*  
(11)    xénophobe                  *xenophobic*

In addition, neoclassical compounds differ from the native ones in the presence of a linking vowel between the two constituents, generally /o/ with Greek constituents and /i/ with Latin ones.

## **1.2 Problems in defining the status of neoclassical constituent**

The neoclassical constituent can be defined through four properties (Amiot & Dal, 2007). First, these elements were lexemes in the source languages (Greek or Latin) but second, they have no syntactic realization in modern languages, since they can only appear as bound constituents. Third, they serve to construct the learned vocabulary of the language. Finally, they display a linking vowel between the two constituents in a particular phonological context. So, neoclassical constituents inherit from Greek or Latin lexemes a semantic content just like autonomous French lexemes. Yet, their non-autonomous nature makes them look like affixes. We argue here that neoclassical constituents do not constitute a homogenous set since the status of each of these elements may vary. A constituent like *xéno-* seems to be closer to modern lexemes since it has a stable lexical meaning ('foreigner') whereas a constituent like *-logue* tends to be closer to affixes, and can even be considered as an affix: it lost its original semantic content as it became very productive, as we will see below.

## **1.3 Different language registers**

The third property stated by (Amiot & Dal, 2007) cannot be applied to all neoclassical constituents. If originally these constituents were used in specialized lexicons (medical, learned, technical vocabularies), this is no longer the case. In fact, the same constituent can be employed in learned language, e.g. *-logue* is primarily used in the medical vocabulary (*endocrinologue*, 'endocrinologist'; *cardiologue*, 'cardiologist'), but the current language adopted it either to imitate the medical vocabulary (*chevillologue*, 'specialist of ankles') or to form the names of other kinds of specialists (*chaussettologue*, 'specialist of socks'). Most of the words constructed on the model of "pure" neoclassical words are used with an ironical sense.

## **2 The Corpus**

In order to account for the differences between these language registers, we chose six French neoclassical elements which seem to have the same behavior (Table 1).

All these elements are inherited from Greek: *-phage* comes from a verb and *-crate*, *-cratie*, *-phone*, *-logue* and *-logie* from nouns.

Our corpus is built on two different resources. First, we looked at the neoclassical compounds contained in two dictionaries : *Le Grand Robert* (online version) and *Le Trésor de la Langue*

French neoclassical element	<i>-crate</i>	<i>-cratie</i>	<i>-phone</i>	<i>-logie</i>	<i>-logue</i>	<i>-phage</i>
English gloss	‘-crat’	‘-cracy’	‘speaker of a language’ <sup>1</sup>	‘-logy’	‘-logist’	‘-phagous’
Original lexeme	κράτος		φωνή		λόγος	φαγεῖν

Table 1: Studied elements and their origins

*Française informatisé* (TLFi). Then, we looked for non lexicalized compounds in a corpus built on the Google Ngrams resource. Finally, the corpus was completed by occasional findings. Table 2 gives the total number of lexemes for each element.

	<i>-crate</i>	<i>-cratie</i>	<i>-phone</i>	<i>-logie</i>	<i>-logue</i>	<i>-phage</i>
Dictionaries	18	30	15	535	226	78
Google	146	230	212	455	562	92

Table 2: Number of entries in the corpus for each element

### 3 Morphophonological analysis

At first, we studied the compounds of our corpus from a formal point of view. In particular, three main constraints were observed that determine the output form of these complex lexemes (3.1): Fidelity to the base, Fidelity to the output form, which has to be, preferentially, /ɔkʁat/, /ɔkʁasi/, /ɔfɔn/, /ɔfaʒ/ and so on, and finally a prosodic constraint on the size of the base. The variation observed in the attested forms is the result of the interaction of these three constraints (3). What we can conclude about morphophonology is that the differences in register or the origin of the bases do not play any role in the formal construction of the neoclassical compounds of our corpus (3.3).

#### 3.1 The constraints involved

The first constraint involved is a constraint of Fidelity. It stipulates that the first element of the compound has to be maximally intact. Here, we refer to the first element as the base, as we consider that the final elements involved are quasi suffixes. If we consider that the Fidelity-to-the-base constraint is respected when there is no truncation, no epenthesis, no haplology and no learned allomorphy, this constraint is not violated in 65% to 85% of the cases, depending on the different elements.

The second constraint relates to the output form, which has end in /ɔkʁat/, /ɔkʁasi/, /ɔfɔn/, /ɔlɔg/, /ɔlɔʒi/ or /ɔfaʒ/, according to the element involved. Obviously, this constraint has a diachronic justification. The presence of the vowel -o itself is part of the definitional properties of neoclassical compounds. In ancient Greek (and still in Modern Greek), the vowel -o marked the compositional character of a word. Hence, the output form of a neoclassical compound, whether built with modern or ancient material, necessary display this vowel. We see in fact that /o/ is present in more than 80% of the words in our corpus at the boundary between the first and

the second element. In the remaining cases, neoclassical compounds mainly display a different vowel (mostly /i/ or /a/) and in a large minority of cases a consonant (mostly liquids /l/ or /ʁ/, which are close to vowels in the universal scale of sonority).

The third constraint that comes into play concerns the size of the complex lexeme. In the light of what we have studied in our corpus, we can infer that words constructed with the neoclassical elements *-crate*, *-cratie* and *-phage* are preferentially preceded by two syllables and *-phone*, *-logie* and *-logue* by three syllables.

### 3.2 The interaction of constraints

In the OT framework (cf. (Prince & Smolensky, 2008); for an application to morphology cf., among others, (Plenat, 1996)), constraints are hierarchically ordered and can be violated. Here we do not order them in a very precise hierarchy, but we can nevertheless observe their interaction. Thus, *pédagogocratie* (‘pedagogue’ + *-cratie*) does not respect either the Fidelity to the base constraint or the size constraint. However, the truncation of the base permits respecting Fidelity to the output constraint, which states that the output has to end in /ɔkʁasi/.

On the other hand, a lexeme like *médiocratie* (‘media’ + *-cratie*) respects the Fidelity to the base constraint and violates the Output constraint, maybe to avoid homophony with *médiocratie* (*medio-* for ‘middle class’ + *-cratie*).

Table 3 shows how the same base may give different outputs words depending on the constraints which are respected or violated<sup>2</sup>.

<i>corporation</i> ‘ship’ + <i>-cratie</i>	(i) Fidelity to the base	(ii) Fidelity to the output /ɔkʁasi/	(iii) Size of the base
corporatocratie (10355)	+	+	--
corporacratie (1830)	-	-	-
corpocratie (1308)	-	+	+

Table 3: Different outputs words for a same base

Let us now look at the behavior of bases containing /o/, once again based on examples in *-cratie*. If the vowel /o/ occurs in the first syllable a base is never truncated after it (e.g. *politocratie* ← ‘politician’ + *-cratie*). If the /o/ occurs in the second or in the third syllable, in a large majority of cases the base is truncated after it (e.g. *consocratie* ← *consommateur* ‘consumer’ + *-cratie*; *tortiocratie* ← *tortionnaire* ‘torturer’ + *-cratie*; *pédagogocratie* ← *pédagogue* ‘pedagogue’ + *-cratie*).

If the base ends in a consonant, the vowel /o/ allows a resyllabification and hence, respect for the Size constraint by one-syllable bases (e.g. *fricocratie* ← *fric* ‘dough’ + *-cratie*).

<sup>2</sup>In Table 3, + indicates that a constraint is respected and – that it is violated. Concerning constraint (iii), – indicates each syllable which violates the two-syllables preference. In brackets we indicate the number of occurrences found on Google (research of 1<sup>st</sup> March 2012)

### 3.3 Different registers, same constraints

The Fidelity to the output constraint is automatically respected with bases of Greek origin, which systematically end in /ɔ/. However, when a suppletive Greek stem is not available, French native stems are subject to the same constraint. For instance, in our corpus very few classical bases are available for words in *-phone*, since their bases are language names, often containing sequences which are uncommon in French. We can compare them with the words in *-logie*, for which large number of bases of classic origin<sup>3</sup> are available. Table 4 shows the difference between the two subcorpora:

	<i>-phone</i>	<i>-logie</i>
Classical bases	3 (1.32%)	396 (40.04%)
French native bases	224 (98.68%)	593 (59.96%)

Table 4: Words in *-phone* and *-logie* with classical and French native bases

In spite of these differences, words constructed with the two elements display /ɔ/ at the boundary between the first and the second element in a comparable proportion<sup>4</sup>:

	<i>-phone</i>	<i>-logie</i>
Output containing /ɔ/ between the two elements	81.94%	94.40%
Of which, with /ɔ/ not present in the base	90.86%	62.30%

Table 5: Percentage of words in *-phone* and *-logie* containing /ɔ/ between the two elements

So, even if the origin of the bases varies, the outputs of neoclassical compounds respect the same formal constraints. We observe the same phenomenon when we look at the size of the bases, whether they are of classic origin or not.

In what follows, we will see that the homogeneity observed in morphophonology is not found when semantics is concerned.

## 4 Morphosemantic analysis

### 4.1 Lexical constraints

The neoclassical elements considered here contribute to lexical enrichment. Each new word is built in connection with the other lexemes in the lexicon. This connection may take different ways. For instance, some lexemes may acquire the status of "leader words" (Roché, 2011). In general, a "leader word" corresponds to the first (or one of the first) derivate which is attested, which becomes the source of a word formation pattern by virtue of its frequency or saliency. Once the pattern is established in the language, new words which are coined can be positioned

<sup>3</sup>We choose to consider bases as French ones when they actually exist in French, even if they originally come from Greek or Latin (e.g. "théâtre" in *théâtrologie*).

<sup>4</sup>We consider here that the segment -o- is present in the base when the base actually contains this segment (e.g. "bobo" in "bobologie") or when the base is a Greek one (e.g. "psycho-" in "psychologie").

with respect to the original model. For instance, we may consider that, as far as meaning is concerned, leader words function as semantic catalysts.

Concerning *-crate* and *-cratie*, two leader words seem obvious : *démocratie* (‘democracy’), first documented in 1370, may have generated the meaning "power of X", while *aristocratie* (‘aristocracy’), also first documented in 1370, the meaning "self-valorizing closed social group"; this meaning has been documented since the end of the eighteenth century. On the other hand, for words in *-phage*, it is more difficult to find a clear leader word (or a set of leader words) on a diachronic or on a logical basis. However, we can also observe the emergence of different semantic patterns. The main semantic patterns of *-phage* words are given below:

- |      |  |   |
|------|--|---|
| (12) | ‘who eats X’                           | e.g. <i>anthropophage</i> (‘anthropophagous’) |
| (13) | ‘who/which consumes X’                 | e.g. <i>chronophage</i> (‘time wasting’)      |
| (14) | ‘who loves X in a pathological manner’ | e.g. <i>téléphage</i> (‘TV addict’)           |
| (15) | ‘who hates X / who annihilates X’      | e.g. <i>turcophage</i> (‘who hates Turks’)    |

If we look at the words constructed with *-logue* or *-logie*, no clear leader word seems to emerge. In this case, lexical organization is semantically "distributed": the lexemes in our corpus display a range of different but connected meanings without any clear attracting pole emerging. In all these cases, new lexemes are constructed by analogy with the already existing lexicon.

What neoclassical compounds tell us about lexical construction is that although lexicalized compounds give us a global vision of the semantic models involved, words productively coined by speakers allow us to refine this models.

## **4.2 How non lexicalized compounds contribute to semantic modelization**

If we look at compounds containing the element *-cratie*, we see that *démocratie*, as we said before, functions as a leader word for the semantic reading ‘power of’. The description of this model can be refined looking at non lexicalized words. In fact, speakers coined such words as *corruptocratie* (‘power exercised by means of corruption’), *fricocratie* (‘power exercised in order to get money’) or *ruecratie* (‘power exercised in the streets’). Along with this original model, another semantic pole emerged, with *aristocratie* as its leader word. The original notion of power is absent from the semantic reading of words in this class, which refer to ‘self-valorizing closed social groups’. Several new words were coined in every-day language according to this model, such as *cathocratie* (‘Catholics as a self-valorizing social group’). On their turn, these two main poles generated some sub-models. Thus, we found such words as *Sarkocratie* (‘a state in which Sarkozy has the power’), modeled on the *démocratie* type, or *footocratie* (‘omnipresence of football’) modeled on *aristocratie* type.

As far as words in *-phage* in our corpus are concerned, the original meaning is ‘who eats X’, as illustrated by words like *anthropophage*. Two antithetical models were generated by metaphor from this original sense, mainly in common language. On the one hand, we have the reading ‘who hates X / who annihilates X’ exemplified by *turcophage* (see (15)). On the other hand, we have the reading ‘who/which consumes (a lot of) X’ illustrated by *chronophage* (see (13)). This model, in turn, generated a fourth reading ‘who loves X in a pathological manner’. Such

lexemes as *cinéphage* ('movie addict') or *téléphage* ('TV addict') exemplify this pattern. From a semantic point of view, words in *-phone* in our corpus are not so interesting, as we limited our observation to words indicating the speaker of a language, ignoring, for the moment, such models as 'music instrument using something' (*xylophone* 'xylophone') or 'measuring device' (*tensiophone* 'blood pressure meter'). However, common language also tends to refine the first model, to a lesser extent. Thus, a lexeme like *belgophone* can refer to someone who speaks Belgian (the variety of French spoken in Belgium) or, in some contexts, a resident of Belgium.

We saw in 4.1 that words in *-logie* cannot be considered to be constructed by analogy with any clear leader word. Rather, their semantic organization has to be considered "distributed". This kind of semantic organization shows the same type of phenomena that we observed with words having a clear leader word. In this case, it is not the range of semantic readings that is modified in newly constructed words but the number of possible bases. The ironical dimension plays a crucial role here. Thus, the 'study of a disease' like *cancérologie* 'research on cancer', *paludologie* 'research on paludism' becomes the 'study of a false or a minor disease': *ampoulogie* 'the study of blisters'. The 'study of a part of the body' (*dermatologie*, *cardiologie*) is turned into the 'study of a part of the body which is normally not studied': *moustachologie* 'the study of moustaches'. In addition, a lexicalized compound may be the source of a new sub-paradigm. We can observe it with *œnologie* 'oenology', which generated *vinologie* (same meaning but base of different origin), *biérogie* 'the study of beer' or *cidrologie* 'the study of cider'. Another example is *démonologie* 'demonology' which gave birth to *elficologie* 'the study of elves' or *vampirologie* 'the study of vampires'.

All these examples show that the observation of register variation allows a more detailed analysis of semantic variation. The lexicalized occurrences in our corpus gave us a global vision of the semantic behavior of each neoclassical element: we saw for instance that *-logue* and *-logie* can receive different analyses since we cannot sort out clear leader word or attracting poles, but only identify a general range of connected meanings. However, regardless of the behavior of each individual element, nonlexicalized compounds contribute to the lexical dynamics of the language, contributing to refine the original semantic models. Today neoclassical elements are not confined to the scientific, technical or learned registers of the language; their use has extended to common language. Below we will see to what extent these refined semantic models are productive.

### 4.3 Semantic models and productivity

Let us look at the semantic models attested for words containing the element *-phage*. We distinguished four different semantic types, exemplified, respectively, by *anthropophage*, *chronophage*, *téléphage* and *turcophage* (see above). Table 6 (see end of the paper) gives the number of occurrences belonging to each type and the distribution of lexicalized / non lexicalized compounds in our corpus. Note that some lexemes have been classified into different types. For instance, *bibliophage* can refer to a paper-eating insect or to someone who reads a lot. Very few words cannot be easily classified in any of the four types: *mélophage* ('parasite living in the fleece'), *mallophage* ('living on mammals and biting without stinging') and *métrophage* ('who regularly travels by subway'). It is clear from Table 6 that some semantic types are more productive than others. Thus, the *anthropophage* type is very predominant among the lexicalized compounds of our corpus (84.81%) whereas it represents less than a half (47.92%) of nonlexicalized compounds. The *chronophage* and *téléphage* types show the contrary trend. Most of the occurrences

of these types are not found in dictionaries (only 13.33% et 6.90%, respectively, of the occurrences of these types are present in the TLFi and/or in Le Grand Robert). If we consider that nonlexicalized compounds tend to be the most recent, we can suppose that these two semantic types are more recent than the first one. This corroborates the fact that it is nonlexicalized words that refine semantic models and generate new sub-models.

In addition, we can observe that some semantic patterns are not productive anymore. It is the case for the ‘discourse characterized by’ interpretation of *-logie* (e.g. *dyslogie* ‘language disorder’, *tachylogie* ‘speech disorder characterized by the accelerating pace of emission of words’) which is not found in new lexemes while the interpretation ‘set / collection of’ seems to occur in coined words (*critérologie* ‘set of criteria’, *discologie* ‘set of disks’, *bouletologie* ‘set of drags’).

## 5 Conclusion

What does language register variation tell us in the study of neoclassical compounds? First, formally, even compounds that have been coined recently with native material (not only from Greek or Latin bases) and are not lexicalized respect the same constraints than compounds using Greek or Latin bases, which have been in the lexicon for a long time, used in scientific, scholar domains. Second, from a semantic point of view, nonlexicalized compounds allow us to refine the global analysis elaborated on the basis of lexicalized compounds: they provide shades which may be then analyzed in sub-models. Some semantic patterns lost their productivity, while some others got a new life in the common language. In this respect, these elements are representative of the lexical dynamics of the language.

## Acknowledgements

I would like to thank Fabio Montermini for reviewing this article and Nabil Hathout for helping me in collecting data with Google Ngram.

## References

- AMIOT D. & DAL G. (2007). Integrating neoclassical combining forms into a lexeme-based morphology. *G. Booij et al*, p. 323–336.
- BAUER L. (1998). Is there a class of neoclassical compounds, and if so is it productive? *Linguistics*, **36**(3), 403–422.
- BOYÉ G., HATHOUT N., LIGNON S., PLÉNAT M. & ROCHÉ M. (2011). *Des unités morphologiques au lexique*. Hermes-Lavoisier.
- IACOBINI C. (2004). Composizione con elementi neoclassici. *M. Grossmann/F. Rainer (a c. di) La formazione delle parole in italiano, Tübingen, Niemeyer*.
- PLENAT M. (1996). De l’interaction des contraintes: une étude de cas. *Current Trends in Phonology: Models and Methods, ESRI, University of Salford, Salford*, p. 585–615.

PRINCE A. & SMOLENSKY P. (2008). *Optimality Theory: Constraint Interaction in Generative Grammar*. Wiley.

ROCHÉ M. (2011). Quel traitement unifié pour les dérivations en-isme et en-iste. *M. Roché, G. Boyé, N. Hathout, S. Lignon, et M. Plénat, Des unités morphologiques au lexique. Hermes-Lavoisier*.

VILLOING F. (2012). French compounds. *Probus, International Journal of Latin and Romance Linguistics. Mouton de Gruyter*, p. 26–60.

	TLFi/Grand Robert		Google	
	Number of occurrences	Percentage with respect of total number of occurrences of the resource	Number of occurrences	Percentage with respect of total number of occurrences of the resource
Type 1 "anthropophage"	67	84.81%	46	47.92%
Type 2 "chronophage"	2	2.53%	13	13.54%
Type 3 "téléphage"	2	2.53%	27	28.13%
Type 4 "turcophage"	6	7.59%	9	9.38%
Others	2	2.53%	1	1.04%
				40.71%
				86.67%
				93.10%
				60%
				33.33%

Table 6: Distribution of the four semantic types of -phage derivatives in dictionaries and Google

## **La variation morphosyntaxique du français radiophonique algérien**

Nabil Sadi<sup>1</sup>

(1) LAILEMM & Université Abderrahmane Mira – Bejaia –  
sadinabil1@yahoo.fr

**Résumé** Ce travail a pour objectif d'étudier le français parlé par des Algériens dans des discours radiophoniques et en situations d'interaction. Il vise à décrire cette pratique chez des professionnels et des non-professionnels de la radio. Nous avons choisi les émissions radiophoniques interactives d'Alger-Chaîne III qui sont considérées comme le miroir de la société et de son public. Notre objet d'étude porte sur l'analyse de la variation morphosyntaxique du français parlé interactionnel dans le but de voir si les contraintes liées à l'interaction verbale peuvent affecter la manière dont les règles grammaticales sont implémentées dans la parole, ou si elles peuvent conditionner la forme même de la grammaire du français parlé en Algérie. Cela permet d'examiner les différentes caractéristiques de ce discours en construction ainsi que les spécificités morphosyntaxiques des différents énoncés qui en découlent.

**Abstract** The aim of this study is to analyse the French variety that Algerians speak in interactive radio discourses. It aims at describing this discursive interaction among radio professionals and non-professionals. To this end, we have chosen the interactive radio broadcasts of "Alger Chaîne III" which mirror our society and the audience. Our goal is to analyse the morphosyntactic variations in interactive spoken French in order to understand whether the oral interaction constraints could affect the way grammatical rules are implemented in oral performance, or could even shape the form of French grammar spoken in Algeria. This allows for the analysis of the different characteristics of this orally constructed discourse as well as for the morphosyntactic details of the different utterances that are generated.

**Mots-clés :** Français parlé, discours radiophonique, variation, morphosyntaxe, contact de langues.

**Keywords:** Spoken French, radiophonic discourse, variation, morphosyntax, language contact.

## 1 Positionnement du problème de recherche

En Algérie, « les sujets parlants sont régulièrement exposés à la langue française. Cette exposition résulte de la présence effective de cette langue depuis la période coloniale jusqu'à ce jour dans toutes les situations de communication vécues par le locuteur » (Derradji, 1998 : 131). Le français fait donc partie d'un plurilinguisme qui règne en Algérie. Son utilisation diffère d'un locuteur à un autre, selon les compétences et la maîtrise de chacun, et d'un groupe générationnel à un autre. D'ailleurs, l'observation des pratiques langagières en situation des locuteurs algériens révèle une « transgression relative » (Derradji, 1999 : 71) du code de la langue française aussi bien au niveau de l'écrit que de l'oral.

C'est pour cette raison que nous nous sommes intéressé au français parlé par les Algériens dans un discours radiophonique et en situation d'interaction. Notre travail vise à décrire cette pratique par des professionnels et des non-professionnels de la parole sur le plan morphosyntaxique. Notre objectif est de déterminer ce qui différencie l'usage du français dans les médias du français parlé ordinaire (Gadet, 1997).

### 1.1 Questionnements

En parcourant les travaux sur le français parlé radiophonique, nous constatons qu'il a été démontré que les professionnels de la radio française utilisent des usages de la langue française caractérisés par certaines propriétés (prosodiques, lexicales, syntaxiques et pragmatiques) qui ne se trouvent pas dans les discours ordinaires et il semble même que les chaînes radiophoniques aient développé des usages propres (Blanche-Benveniste, 2007).

Il nous semble alors très intéressant d'étudier le français parlé par des professionnels et des non-professionnels de la radio algérienne, dans une situation d'interaction et de coexistence de plus de deux langues et dans un discours très particulier, l'oral spontané. Dès lors, la problématique des particularités morphosyntaxiques du français parlé radiophonique s'impose et plusieurs questions nous viennent à l'esprit :

L'utilisation du langage est le produit de la socialisation de l'individu ; lorsqu'il entre en interaction avec d'autres locuteurs, il use de toutes ses compétences langagières afin de mieux faire passer son message. Dans ce cas, quels sont les différents aspects de ce français parlé par les Algériens ? S'agit-il d'un français standard ou d'une variété spécifique qui reflète la culture algérienne ?

### 1.2 Le corpus

L'Algérie dispose de trois chaînes radiophoniques publiques et nationales : la chaîne I qui émet ses programmes en arabe, la chaîne II en kabyle et la chaîne III en français, à côté de nombreuses chaînes locales. L'objectif de notre travail étant de faire une analyse d'un discours oral (radiophonique) et d'étudier les caractéristiques du français parlé dans un contexte interactionnel et dans une situation de contact de langues, nous avons choisi la chaîne III afin de disposer d'un discours en langue française. Le choix des émissions sur lesquelles nous envisageons de travailler repose sur des critères que nous avons délimités au départ et qui correspondent aux caractéristiques de notre sujet, telles que les particularités des thèmes des émissions et des interactants.

Nous proposons d'étudier en premier lieu *Le rendez-vous de l'économie* (05 numéros), une émission qui traite de plusieurs aspects de l'économie, qu'elle soit nationale ou internationale,

tels que les activités industrielles, financières et commerciales du monde en général et de l'Algérie en particulier.

En deuxième lieu, nous avons choisi *Conseils et vous* (04 numéros), une émission qui représente un espace d'expression pour les gens en détresse, ayant des problèmes de cœur, des problèmes professionnels, personnels, familiaux, etc.

En dernier lieu, nous nous sommes intéressé à *Micro ondes* (05 numéros), une émission placée sous le signe de la musique de tous genres, de divertissement, de jeu et de plaisir,

La transcription<sup>1</sup> de la totalité des enregistrements (quatorze numéros) a donné un total de 255 pages, soit 104284 mots. *Le rendez-vous* de l'économie en contient 40293 mots sur 79 pages, *Conseils et vous* en compte 30369 mots sur 71 pages et enfin *Micro ondes* avec 33622 mots sur 105 pages.

### **1.3 Hypothèses de recherche**

Avant de pouvoir effectuer une description du français parlé en Algérie, il importe d'établir quel français d'Algérie nous allons décrire. Nous nous retrouverons en effet face à des réalités linguistiques très différentes. Nous pouvons aborder le français utilisé par les politiques, celui des médias (la radio, la télévision, etc.), celui des jeunes scolarisés en arabe avec un contact irrégulier avec le français ou celui utilisé par les intellectuels (universitaires et écrivains). L'analyse du français, parlé par des locuteurs différents et dans des situations différentes, nous amènerait à des interprétations différentes liées aux différents paramètres de son usage.

Nous voulons déterminer ce qu'est le système (l'appareil) du français tel qu'il est parlé à la radio, dans un contexte d'interaction et de coexistence de plusieurs langues (français, arabe, berbère...), cerner les procédés d'assemblage des mots, les formes que ceux-ci revêtent dans leurs différents emplois et les combinaisons plus ou moins complexes où ils entrent pour former des locutions. Nous formulons trois hypothèses afin de mieux mesurer/comprendre l'impact des variations morphosyntaxiques sur le système linguistique :

1. *Ces procédés sont différents de ceux de l'écrit oralisé et ils sont conditionnés par le contexte d'interaction et les contraintes liées à la situation de communication ;*
2. *A l'oral et dans une situation de contact de langues, chaque locuteur s'accommode des matériaux et des codes mis à sa disposition. Il s'agit de mécanismes dont il peut se servir pour communiquer ses messages ainsi que pour soutenir sa propre pensée discursive ;*
3. *Ce parler tire son efficacité des effets que le locuteur peut obtenir dans la réalisation des élocutions ainsi que des langues utilisées, qui ressortent sur sa manière de structurer ses énoncés, sur l'ordre des propositions et sur le sens qu'elles peuvent véhiculer.*

---

<sup>1</sup> Notre corpus étant long (255 pages), nous avons opté pour une transcription orthographique plutôt qu'une transcription phonétique. De plus, « il n'existe pas aujourd'hui de système de transcription unifié. Chacun forge son système du moment que la transcription répond aux contraintes de précision, de fidélité et de lisibilité. D'une manière générale, on n'utilise pas de transcription phonétique, trop difficile à lire, mais des transcriptions orthographiques, plus ou moins standard ou adaptées. L'orthographe adaptée cherche à rendre compte de certains phénomènes de prononciation » (Traverso, 1999 : 24).

## 2 La variation du français parlé

« Travailler sur des prises de parole longues permet de constater que le locuteur peut varier son expression à l'intérieur d'une seule et même production (Miller & Weinert, 1998), en fonction de divers facteurs. Entre autres, cela peut être lié à la représentation qu'il se fait de la situation de production » (Bilger & Cappeau, 2004 : 18). Nous proposons dans ce point de décrire les productions langagières des différents locuteurs des trois émissions radiophoniques et de déterminer les composantes qui régissent leurs réalisations du français parlé.

« La description de la variabilité des productions langagières telle qu'elle est proposée par la sociolinguistique variationniste s'appuie sur l'identification d'un trait structural qui donne lieu à des réalisations multiples, conditionnées par des facteurs internes et/ou externes que l'on cherche à déterminer dans un corpus » (Ploog, 2002 : 77). Les situations de contact de langues, comme c'est le cas en Algérie, génèrent chez les locuteurs des représentations spécifiques des langues impliquées, représentations qui se répercutent dans des pratiques linguistiques différentielles qui peuvent, dans certaines conditions, affecter le système lui-même. Que ce soit à l'oral ou à l'écrit, tout locuteur francophone maîtrise donc plusieurs registres de langue. Il est capable d'ajuster ses productions verbales en fonction de l'image qu'il se fait de son interlocuteur, de celle qu'il veut donner de lui et des circonstances de production de son discours.

L'analyse des émissions radiophoniques qui constituent notre corpus a permis de dégager une variation touchant certaines formes grammaticales, des cas précis de variation morphosyntaxique du français parlé par des professionnels et des non-professionnels de la radio algérienne. Nous donnerons une illustration à travers les exemples les plus récurrents et nous classerons les différentes tournures de celles attestées à l'oral en général à celles qui caractérisent le français parlé radiophonique algérien.

### 2.1 Tournures formelles/ informelles

Nous avons remarqué une simultanéité dans l'utilisation de tournures formelles et informelles de la part de nos locuteurs, avec une très grande fréquence dans *Le rendez-vous de l'économie*. La fonction et le niveau de ces différents spécialistes de l'économie ainsi que la nature du thème dont ils débattent influencent la forme des prises de parole.

- *A : difficile a été euh la façon dont on devait calibrer euh cette émission pour illustrer le monde du travail*

(Emission. 1, numéro 1, page 1, ligne 5-6)

L'animateur utilise dans ce passage une tournure fort soutenue et qui est habituelle dans les moments de lancement. L'inversion du sujet groupe nominal et du verbe d'état suivi de l'adjectif attribut est dite stylistique.

Un autre passage présente aussi une tournure très formelle qui est l'utilisation de deux participes présents :

- *Il : aujourd'hui sachant que le rôle de la banque étant très important*

(Emission 1, numéro 1, page 15, ligne 544-545)

D'autres tournures formelles sont à noter, à savoir l'utilisation d'un sujet nominal pluriel avec un présentatif formel « ce sont » :

- *I<sub>1</sub> : ce sont des experts de l'université*  
(Emission 1, numéro 1, page 3, ligne 9)
- *A<sub>U2</sub> : ce sont les personnes dans les banques qui se sont accaparé à leur profit les personnels*  
(Emission 1, numéro 1, page 14, ligne 506-507)
- *I<sub>3</sub> : ce sont pas des havres de compétences dans le marché*  
(Emission 1, numéro 1, page 15, ligne 536-537)

Dans ce passage, « ce sont » coïncide avec une absence de ne de négation, ce qui évoque les « séquences préformées » de Moreau qui « fonctionneraient comme des entités globales et pourraient être stockées comme telles dans notre lexique interne » (1986 :154).

Concernant la deuxième émission, deux exemples sont à noter :

- *A : je ne sais pas comment vous le dire parce que c'est vrai c'est euh ce sont des mots/ qui euh qui vont peut-être aider les autres*  
(Emission 2, numéro 2, page 11, ligne 376-577)

Dans ce passage, l'animatrice emploie un présentatif sous sa forme formelle avec un sujet nominal pluriel, avec une hésitation qui lui a permis de relancer avec « ce sont » plutôt que « c'est ». La psychologue use du même procédé dans le quatrième numéro où elle présente les auditeurs et les auditrices :

- *I<sub>A</sub> : je voudrais quand même rappeler que ce sont les auditeurs et les auditrices qui ont fait l'émission/ et je leur dis à septembre*  
(Emission 2, numéro 4, page 15, ligne 521)

Quant à la troisième émission, un seul exemple apparaît où l'animateur qualifie les musiciens de précurseurs du son électro-techno-pop avec le présentatif « ce sont ».

- *A : ce sont les précurseurs du son électro-techno-pop*  
(Emission 3, numéro 1, page 2, ligne 63)

Nous pouvons attribuer cette différence de fréquence entre les émissions à deux composantes de la situation de communication. Le thème et le cadre de la première émission diffèrent des deux autres émissions, dans la mesure où les débats sur l'économie amènent les locuteurs à adopter un registre plutôt soutenu. Tandis que dans *Conseils et vous* pèse une atmosphère assez lourde, vu que les interventions expriment des sentiments tristes et des expériences douloureuses des auditeurs. Ce sont des facteurs conduisant ces derniers vers un discours relâché. Quant à *Micro ondes*, le divertissement prime et une atmosphère chaleureuse règne tout au long de l'émission, ce qui rend les différentes interventions informelles.

Notre corpus regorge également de plusieurs tournures informelles. Les différents locuteurs emploient le présentatif « c'est » avec un sujet nominal pluriel :

- *I<sub>1</sub> : c'est les confédérations marquantes si je puisse m'exprimer ainsi de l'économie nationale*  
(Emission 1, numéro 1, page 2, ligne 71-72)

- *I<sub>2</sub> : c'est eux qui viennent frapper à votre porte*  
(Emission 1, numéro 2, page 5, ligne 175)
- *I<sub>A</sub> : est-ce que c'est toutes les constructions qui sont hors normes ?*  
*I<sub>4</sub> : c'est pas toutes les constructions qui sont hors normes*  
(Emission 1, numéro 2, page 10, ligne 366-367)
- *I<sub>4</sub> : donc c'est les actionnaires*  
(Emission 1, numéro 2, page 12, ligne 438)

Nous remarquons dans les deux derniers passages que l'animateur assistant influence l'invité sur le choix de la forme du présentatif dans sa réplique immédiate (le premier passage). Plus loin dans la conversation, l'invité utilise la même forme (le deuxième passage).

Un locuteur peut également influencer ces interlocuteurs par effet en mémoire. Ces derniers adoptent une forme déjà utilisée, entre plusieurs tours de parole. Il s'agit d'une influence éloignée inter-locuteurs :

- *I<sub>1</sub> : je pense que la construction de la nouvelle ville de Boubezhoune de Hassi Messaoud euh c'est c'est des projets très proches*  
(Emission 1, numéro 3, page 2, ligne 54)
- *I<sub>2</sub> : c'est des grands projets de transfert d'eau/ c'est des grands projets// c'est le programme du président de la république/ il faut le reconnaître il faut le dire/ c'est des méga projets/*  
(Emission 1, numéro 3, page 4-5, ligne 169-171)

Dans les deux exemples ci-dessus, le deuxième invité aborde le même sujet évoqué par le premier invité en calquant la forme et en la reformulant avec de nouveaux adjectifs.

Nous avons également recensé les mêmes caractéristiques dans les deux dernières émissions mais avec une faible fréquence :

- *I<sub>A</sub> : c'est des échanges sympathique*  
(Emission 2, numéro 1, page 15, ligne 542)
- *A<sub>U4</sub> : c'est des ennemis et les deux dernières c'est les chouchous de la famille*  
(Emission 2, numéro 2, page 2, ligne 39-40)

## 2.2 L'hypercorrection

« Croire qu'il y a une façon prestigieuse de parler sa langue implique, si l'on ne pense pas posséder cette façon de parler, qu'on tente de l'acquérir (...) Or ce mouvement tendanciel vers la norme peut engendrer une restitution exagérée des formes prestigieuses : l'hypercorrection » (Calvet, 1996 : 55). Nous avons relevé quelques exemples d'hypercorrection où les locuteurs, en voulant bien s'exprimer, usent de formes incorrectes :

- *I<sub>1</sub> : c'est les confédérations marquantes si je puisse m'exprimer ainsi de l'économie*

*nationale*

(Emission 1, numéro 1, page 2, ligne 71-72)

- *I<sub>2</sub> : nous sommes l'une des professions qui a le **taux d'emploi féminine** le plus élevé*  
(Emission 1, numéro 1, page 7, ligne 252-253)

Dans le premier passage, *puisse* coïncide avec le subjonctif *puisse*. Par contre, dans le deuxième extrait, l'invité emploie l'adjectif *féminin* au féminin.

D'autres exemples témoignent aussi d'une hypercorrection, dans la mesure où les locuteurs, en confondant des catégories grammaticales, tombent dans des incorrections :

- *A : c'est ça le véritable classique/ normalement les véritables classiques de la vieille école qui vous ouvre **grandes** leurs portes*  
(Emission 3, numéro 2, page 4, ligne 121-122)
- *A<sub>U6</sub> : ben↑ je suis **financière***  
(Emission 3, numéro 3, page 8, ligne 265)

Dans le premier extrait, l'animateur accorde l'adverbe invariable *grand*, comme un adjectif. Tandis que dans le deuxième extrait, il y a confusion entre le substantif et l'adjectif avec l'emploi de *financière*.

- *A<sub>U14</sub> : euh je pense que c'est une émission vraiment **réussite***  
*A : d'accord*  
*A<sub>U14</sub> : elle est **réussie**/ tu vois*  
(Emission 3, numéro 2, page 20-21, ligne 653-655)

Par ailleurs, l'emploi de *réussite* dans le passage ci-dessus relève d'un télescopage avec la forme adjectivale qui se termine avec *ant/ent*. L'animateur se rend compte de son erreur et la corrige ensuite.

Une confusion entre une charge sémantique d'un nom et sa flexion conduit le locuteur vers une hypercorrection. Dans l'exemple suivant, le locuteur utilise l'auxiliaire *avoir* à la troisième personne du pluriel *ont*, en référence au *monde* qui possède une charge sémantique du pluriel.

- *I<sub>2</sub> : il faut dire merci à **tout le monde**/ qui **ont contribué**/ soit PTT/ soit Sonelgaz/ soit Gaz/ soit Electricité*  
(Emission 1, numéro 3, page 5, ligne 186-187)

### 2.3 Le télescopage

Plusieurs locuteurs confondent des expressions ou des formes plus ou moins proches, que ce soit syntaxiquement ou lexicalement, et fournissent ainsi des passages plus ou moins complexes :

- *I<sub>2</sub> : on veut ouvrir ce pays mais malheureusement les étrangers **ils croivent** que ce pays il est ouvert*  
(Emission 1, numéro 2, page 4, ligne 160-161)

Dans cet extrait, l'invité utilise une forme populaire ancienne du verbe *croire* ; on peut penser

aussi qu'il a confondu avec le verbe *devoir* (ou le verbe *recevoir*) et il l'a conjugué de la même façon.

- *A<sub>U1</sub> : personnellement moi je suis une personne très déséquilibrée très euh très mal à l'aise dans ma peau// je euh je suis très malheureuse dans ma peau*  
(Emission 2, numéro 2, page 3, ligne 74-75)

Un autre exemple de télescopage est à noter dans ce passage. L'auditrice combine les deux expressions « être mal à l'aise » et « être mal dans sa peau » pour en faire « être mal à l'aise dans sa peau ». Même combinaison avec « être malheureuse » pour donner « être malheureuse dans sa peau ».

Un autre exemple de télescopage apparaît dans l'exemple suivant, mais cette fois-ci cela relève d'une part d'un mauvais choix de verbe et d'autre part d'une redondance de sens :

- *A<sub>U1</sub> : j'aimerais bien voulu savoir la position de monsieur Benkhelfa↑*  
(Emission 1, numéro 1, page 11, ligne 389)

En effet, l'auditeur conjugue le verbe *aimer* au conditionnel qui, logiquement, sera suivi d'un verbe à l'infinitif qui est *savoir*, mais il insère le participe passé du verbe *vouloir*, en télescopant avec la forme du futur antérieur *j'aurais voulu*.

- *I<sub>4</sub> : les petits achats sont euh se font par chèques certifiés :: donc il y a une idée derrière la tête*  
(Emission 1, numéro 2, page 19, ligne 709-711)

Dans ce dernier passage, l'invité essaye d'expliquer pourquoi l'Etat a suspendu un décret qui aurait été très bénéfique aux banques, en termes de flux financiers, en prétendant que ces dernières n'étaient pas prêtes. En donnant un contre-argument, il voulait démontrer que cela dissimulait intentionnellement quelque chose, et il a lui-même télescopé *il a une idée derrière la tête* avec une expression comme « il y a anguille sous roche/... ».

Nous avons noté une tournure récurrente chez une auditrice de *Conseils et vous* qui relève aussi d'un télescopage :

- *A<sub>U5</sub> : j'ai euh j'ai jamais donné confiance à une personne/ c'est impossible*  
(Emission 2, numéro 4, page 14, ligne 483)
- *A<sub>U5</sub> : je suis consciente de tout ça/ mais je ne peux pas donner confiance à une personne*  
(Emission 2, numéro 4, page 14, ligne 497)

En effet, l'auditrice utilise la même expression *donner confiance* pour *faire confiance*, télescopée avec *donner ma confiance*.

## 2.4 Traits syntaxiques et éléments interférentiels

En parcourant nos pages de transcriptions, nous avons remarqué que plusieurs traits syntaxiques et éléments interférentiels se dessinent :

- *I<sub>A</sub> : parfois les banques craignent de **prendre des actions***  
(Emission 1, numéro 1, page 11, ligne 424-425)
- *A<sub>U3</sub> : quand je fais confiance à personne tout s'envole*  
(Emission 2, numéro 1, page 12, ligne 411)

L'emploi de *prendre* dans le premier passage et de *personne* dans le deuxième relève d'une interférence avec l'arabe dialectal et/ou le kabyle, une traduction littérale des équivalents lexicaux des termes respectifs. Ainsi, nous aurons respectivement en arabe dialectal et en kabyle les tournures suivantes :

- *El bonouk tkhaf teddi des actions*  
*El bankath tagadhent adhawint des actions*  
(**Les banques craignent de prendre des actions**)
- *Ki n'dir ethiqa f'wahed koulech yrouh*  
*Koul madhamnegh hed koulech adhirouh*  
(**Quand je fais confiance à quelqu'un tout s'envole**)

Deux autres passages montrent également une interférence avec l'arabe et/ou le kabyle :

- *A<sub>U1</sub> : j'ai vraiment peur aussi de **parler sur mon enfant***  
(Emission 2, numéro 2, page 3, ligne 105-106)
- *I<sub>4</sub> : si on était venu sur ce marché il y a **cinq ans derrière***  
(Emission 1, numéro 5, page 4, ligne 161-162)

Dans le premier passage, le remplacement par l'auditrice de la préposition *de* (sémantiquement peu précise et très fréquente) par *sur* (qui est plus précise) est dû à une interférence de l'arabe et aussi du kabyle. En arabe et en kabyle, nous aurons :

- *N'khaft nahder aala oulidi*  
*Ougadhegh adhadhregh femmi*  
(**J'ai peur de parler sur mon fils**)

Par ailleurs, ce phénomène de remplacement d'une préposition sémantiquement vague par une plus précise est très répandu dans les situations de contact. Il s'agit d'une distribution différente des prépositions en arabe et en français. Les prépositions en arabe « appartiennent à une catégorie plus grande appelée *hourouf al maâni* (*outils de sens* en français) qui englobe des éléments linguistiques pouvant être rendus en français aussi bien par des prépositions que par des conjonctions, des pronoms ou encore des adverbes. Ces *outils de sens* sont classés non pas suivant leur nature grammaticale, mais plutôt en fonction des différents sens qu'ils peuvent apporter dans le discours. Ainsi, le même terme peut apparaître dans différentes sous-catégories en fonction des différents sens qu'il peut véhiculer » (Bendaoud, 2008 : 8).

La préposition *de* en arabe indique la provenance, l'origine relative ou la manière et s'applique aux idées et pas seulement aux origines matérielles ou géographiques. Tandis que *sur* indique ce qui dirige ou gouverne une situation, ou un endroit sur lequel une chose est posée et signifie « prendre de la hauteur ». La correspondance imparfaite entre les deux systèmes prépositionnels est une source d'erreur vu que le recours au calque et à la traduction littérale est ce qui domine le plus les productions des locuteurs algériens.

Nous rencontrons le même cas dans le deuxième passage où le verbe *réagir* semble pris comme synonyme de *parler* en arabe et en kabyle par la psychologue. Cette tournure est aussi attestée en français ordinaire, dans un sens de *à propos de*. En voici les équivalents en arabe et en kabyle :

- *Rayha nahder aala zaaf taa Linda*  
*Thoura adhhadregħ f'lefkiaath n'Linda*  
(*Je vais parler de la colère de Linda/ Je vais réagir à propos de la colère de Linda*)

Dans le troisième passage, un trait interférentiel apparaît, dans la mesure où *derrière* coïncide avec *en arrière* et vient redoubler *il y a X ans*. Le locuteur traduit littéralement l'équivalent lexical en arabe et/ou en kabyle :

- *Loukane jina lhad el marcher khems snine mellour*  
*Loukane noussad ghou marcher khems snine aghdhefir*  
(*Si on était venu sur ce marché cinq ans derrière*)

## 2.5 Du côté de la conjugaison

Nous avons relevé des exemples en référence à l'utilisation de quelques temps de conjugaison et à l'absence de quelques modes.

- *A : l'émir du Golfe s'arrête à un moment euh et pose la question au bureau d'étude euh à euh l'architecte anglais de nationalité/ Britannique de nationalité/ et lui **demanda** où est-ce que nous sommes actuellement ? le Britannique **répondit**/ nous sommes dans un désert et complètement désert/ le euh l'émir euh du haut de sa sagesse/ de ses millions de dollars **répondu** euh **répondit** euh que penserais-tu si on décidait de remplir ce désert et le transformer en euh en une contrée riche/ et le Britannique lui **répondit**/ pourquoi pas/ euh ils l'ont fait/ et c'est Dubaï aujourd'hui*  
(Emission 1, numéro 3, page 1, ligne 17-21)

L'animateur de la première émission relate des faits qui se sont passés à Dubaï entre un émir du Golfe et un architecte anglais. Il s'agissait d'une rencontre qui a révolutionné l'histoire du désert de Dubaï. L'animateur fait usage du passé simple tout au long de son récit et s'efforce de le maintenir en corrigeant sa conjugaison.

Une autre caractéristique se dégage de notre corpus, à travers l'emploi de l'hypothétique avec le conditionnel :

- *A<sub>U1</sub> : voilà↑/ **si j'aurais** l'occasion je vous ferais un petit mot euh pour vous écrire euh on peut vous écrire par euh enfin un courrier normal ?*  
(Emission 2, numéro 3, page 2, ligne 71-72)
- *A<sub>U2</sub> : **si je pourrais** faire plus je suis toujours là/ je suis toujours là euh y a pas de problème/ on m'appelle y a aucun aucun souci*  
(Emission 2, numéro 4, page 4, ligne 119-120)

Comme nous pouvons le constater, les auditrices utilisent *Si* suivi du conditionnel.

À côté de l'usage du passé simple qui tend à disparaître même à l'écrit, nous notons une absence du subjonctif dans quelques passages. En ce qui concerne les modes, nous pouvons

noter, selon les études, des divergences quant au degré de représentation du subjonctif à l'oral. Son absence est rarement attestée pour certaines études alors que pour d'autres, elle l'est plus fréquemment, tel que c'est le cas dans notre corpus, où c'est toujours l'indicatif qui apparaît :

- $I_2$  : *il faut que le consommateur choisit et les bons tarifs*  
(Emission 1, numéro 2, page 4, ligne 134)
- $I_2$  : *c'est un grand projet/ que le gouvernement algérien est en train de reprendre/ pour qu'elle sera une autoroute vers le Nigeria*  
(Emission 1, numéro 3, page 5, ligne 168-169)

Dans ces passages, l'indicatif remplace le subjonctif dans la mesure où le verbe *choisir* est au présent (premier passage) et le verbe *être* est au futur simple (deuxième passage),

A ce sujet, Gadet (2000) souligne qu'« aux plans significatifs il n'y a pas spécificité radicale des phénomènes relevant de l'oral et de l'écrit : on ne saurait évoquer une seule forme grammaticale qui soit l'apanage de l'un ou de l'autre. Il y a cependant des différences dans l'exploitation des ressources linguistiques, et des préférences tendancielle. (...) On peut citer le passé simple ou l'imparfait du subjonctif que l'on ne rencontre plus guère dans la majorité des oraux ordinaires (sauf dans quelques variétés *conservatrices*) ».

En conclusion, nous pouvons dire que le français parlé interactionnel dans un contexte radiophonique algérien se caractérise par une variation d'ordre morphosyntaxique liée à des spécificités situationnelles. D'une part, il se soustrait à certaines règles grammaticales qui dépendraient étroitement des situations de communication, dans la mesure où ces dernières génèrent chez les locuteurs des représentations spécifiques des langues impliquées. Ces représentations qui se répercutent dans leurs pratiques linguistiques différentielles peuvent, dans certaines conditions, affecter le système linguistique normatif. D'autre part, il se caractérise par la diversité de registres de langue qui sont fortement influencés par le type de sujet abordé et par le rôle incombé aux locuteurs qui semblent passer alternativement par des phases de langage familier, standard selon leurs interlocuteurs, et presque simultanément par d'autres phases très soutenues.

## Références

- BENDAOU M.-L. (2008). « La confusion dans l'usage des prépositions du français chez les locuteurs non natifs : cas des élèves de 3.A.S. du lycée Ziri B'en Menad à Bou-Saâda ». Mémoire de magistère, université de M'sila.
- BILGER M., CAPPEAU P. (2004). « L'oral ou la multiplication des styles ». *Langage et Société* 109, 13-30.
- BLANCHE-BENVENISTE C. (2007). « Normes anciennes et nouvelles dans le langage des médias ». Actes de *Le français parlé des médias*, 31-48.
- CALVET L.-J. (1996). *La sociolinguistique*. Que sais-je ? 2<sup>ème</sup> édition. Paris : P.U.F.
- DERRADJI Y. (1998). « Remarques sur l'alternance codique conversationnelle en Algérie », in Queffélec, A. (éd.). *Alternances codiques et français parlé en Afrique*, 131-141.

DERRADJI Y. (1999). « Le français en Algérie : langue empruntée et emprunteuse ». *Le français en Afrique* 13, 71-82.

GADET F. (1996-1997). *Le français ordinaire*. Paris : Armand Colin.

GADET F. (2000). « L'oral : quelles modalités de production pour quelles significations ». Extrait des actes *du séminaire national, Perspectives actuelles de l'enseignement du Français*. Paris les 23, 24 et 25 octobre 2000.

MILLER J., WEINERT R. (1998). *Spontaneous spoken language: Syntax and discourse*, Oxford: Clarendon Press.

MOREAU M-L. (1986). « Les séquences préformées : entre les combinaisons libres et les idiomatismes. Le cas de la négation avec et sans *ne* ». *Le Français moderne* 54, 137-60.

PLOOG K. (2002). « Questions de syntaxe : L'approche syntaxique des dynamiques langagières : non-standard et variation ». *Cahiers de Grammaire* 27, 77-96.

TRAVERSO V. (1999). *L'analyse des conversations*. Paris : Editions Nathan.

## La variation et les corpus : une entrée méthodologique, avec l'exemple du style

Françoise Gadet<sup>1</sup>

(1) MoDyCo & Université de Paris Ouest

Les corpus sont aujourd'hui incontournables en sciences du langage, avec pour effet la production d'une masse potentiellement illimitée de données. Mais tout est-il bon pour "faire corpus"? Habert avait plaisanté à ce sujet, avec l'heureuse formule "gros, c'est beau" (2000). Si une masse importante de données présente de l'intérêt pour certains objectifs, il en est d'autres qui exigent plutôt qualité et fiabilité (Cappeau & Gadet 2007), ce qui soulève la question des critères pour évaluer la qualité: qu'est ce que de "bonnes données"? Nous défendrons ici l'idée que c'est en amont, lors de la constitution d'un corpus, que se joue son traitement, et non *post hoc* sur des données recueillies à l'aveugle: c'est parce qu'on se pose certaines questions que l'on adopte telle ou telle démarche méthodologique, ce qui plaide pour une diversification des modalités de recueil.

### 1 Le style

Je ne parlerai ici que d'objectifs de corpus sur la *variation*, car si la documentation pour les faits de langue proches du standard est assez abondante, c'est loin d'être le cas pour les données non ou peu standard, qui soulèvent des questions sur la façon de traiter/situer les données.

#### 1.1 Des données pour étudier le style

Parmi les données les plus difficiles à recueillir, il y a ce qui relève du *style*, ou produit de l'instabilité des productions d'un même locuteur (variabilité stylistique, intra-locuteur, diaphasique). Cette variabilité donne lieu à un répertoire dont dispose tout locuteur, mis à contribution dans la vie sociale par son adaptabilité aux interactions diversifiées qu'il traverse (Goffman 1979 parle d'*ajustement*).

Le *style* n'est plus aujourd'hui la « dimension négligée » que pouvait dénoncer Bell 1984, il a au contraire donné lieu à des travaux, qui soulignent tous le caractère polysémique du terme (Coupland 2007, Auer 2007). Ce terme renvoie en effet, selon une opposition entre produits et processus, à des conceptions divergentes de la variabilité dans la langue :

- d'un côté, on localise le style en des points précis de la chaîne parlée, à partir de phénomènes linguistiques répertoriés, quantifiables, se manifestant dans la linéarité (c'est la notion de variable), autorisant l'hypothèse d'investissement social de phénomènes linguistiques;
- de façon radicalement opposée, une conception que l'on dira sémiotique regarde le style comme plus diffus, concernant tout ce qui contribue à la "signification sociale" des

énoncés.

On fait l'hypothèse que des données d'origines différentes ne documentent pas les mêmes faits linguistiques, qu'il y a donc des effets de la méthodologie de recueil des corpus, dont il faut dès lors diversifier les objectifs, les contextes, et les modalités de recueil.

Le style est souvent donné comme une évidence, maladroitement formulée dans la notion scolaire de "niveaux de langue" qui présente, parmi divers inconvénients, celui de supposer que c'est la même chose qui se dit en des styles différents. A partir de Labov, la sociolinguistique a renouvelé la problématique du style, en mettant en avant deux acquis décisifs: 1) il n'y a pas d'énoncé sans style, 2) il n'y a pas de locuteur monostyle, contrairement aux représentations de la doxa. La conception labovienne a deux effets concernant les données: encourager le recours à l'interview (voir *infra*), et favoriser la localisation du style dans des variables, lieux ponctuels dans le fil du discours où est censé se dire le style.

On se demandera dans quelle mesure les corpus peuvent contribuer à comprendre le style, difficile à documenter par son caractère fugace et instable. Il a surtout été étudié à partir d'interviews<sup>1</sup>, en particulier dans la tradition labovienne qui alterne différents moments afin de provoquer différents styles; avec pour limites une certaine monotonie et le formatage par le format de l'interview (qui n'est souvent qu'un faux dialogue, donnant au contraire lieu à un certain monologisme - voir Thibault & Vincent 1990).

## 1.2 Le style en français, et chez les jeunes

Le français est une langue très standardisée qui connaît une vaste amplitude et souplesse stylistique, en même temps qu'une certaine rigidité de ses genres. Les styles ont été bien étudiés à l'écrit, moins à l'oral (ce qui peut être un effet d'idéologie du standard). Ce sont surtout la phonologie et le lexique qui ont été étudiés, la prosodie un peu moins, le discursif, le pragmatique et le syntaxique encore moins. Quand le style est étudié à l'oral, c'est en général chez des locuteurs des classes moyennes censés disposer d'une palette plus large, ce qui répond à la conception traditionnelle en standard/non-standard et formel/informel. Il est cependant possible d'étudier le style partout, puisque tous les locuteurs sont *pluristyle*, disposant d'une amplitude de *répertoire* plus ou moins vaste, mais jamais monocorde. On peut en particulier l'étudier chez les jeunes, qui offrent un effet de loupe sur le style, par leur rapport direct au style et par leur croyance en la spécificité de leurs façons de parler.

L'ANR-09-FRBR-037-01 (*Multicultural London English/Multicultural Paris French*) a constitué un corpus en région parisienne (voir Gadet/Guerin 2012), étudiant les façons de parler des jeunes en train d'advenir en région parisienne, en rapport avec les contacts de langues. La population enquêtée concerne des urbains, jeunes (ou se voulant tels - voir Bulot 2004), de milieux sociaux modestes, issus de l'immigration. Nous élaborons une tentative de caractérisation "interne" non dominée par des critères socio-démographiques, qui impose un passage au crible des données recueillies. Nous disposons ainsi d'entretiens réalisés sur une base d'exploitation de réseaux, reposant pour partie sur une qualité de la relation antérieure entre l'enquêteur et l'enquêté. C'est pourquoi le corpus MPF est l'œuvre conjointe de 17 enquêteurs différents, qui à l'heure actuelle ont recueilli plus de 37 heures aujourd'hui transcrites/revues (437.000 mots).

---

<sup>1</sup> En préparant une banque de données sur les corpus de français hors de France pour la DGLFLF, on a pu constater sur 145 réponses répertoriées qu'environ 80% des corpus ont été constitués sur la base d'interviews.

Le focus porte donc sur le pôle "informel", celui du *vernaculaire*, défini par les effets sur les façons de parler de contextes relevant de l'immédiat communicatif (Koch & Oesterreicher 2001), et non sur l'échantillonnage socio-démographique, en cherchant à dépasser le mythe d'un degré zéro de surveillance sociale, auto- ou hétéro-, puisque c'est plutôt la qualité de l'interaction qui est en jeu. On suppose ainsi une conception du style non comme production solitaire, mais comme se jouant dans l'interaction (*recipient design* de Goffman, repris par Bell 1984 dans sa conception d'*audience design*).

Nous nous arrêterons maintenant à deux aspects des réflexions méthodologiques soulevées à l'occasion de la constitution de ce corpus, concernant un type de population (à tort) tenue pour ne disposer que d'un répertoire stylistique limité, peu susceptible d'être prise comme terrain pour illustrer la variabilité stylistique.

## **2 La "qualité" des interviews : se forger l'oreille par des écoutes réitérées**

Nous avons recueilli des données en gros caractérisables comme relevant de trois types<sup>2</sup>: entretiens "traditionnels" (relation préalable faible ou inexistante entre les protagonistes, script préparé à l'avance); entretiens "de proximité", avec des protagonistes sélectionnés par réseau; données "écologiques", événements discursifs qui prennent place indépendamment des sollicitations d'un enquêteur. Ces derniers sont les plus difficiles à recueillir, car ils ne peuvent l'être sans l'aide active d'enquêtés, qui n'ont pas toujours l'envie, la capacité ou le savoir pour répondre aux objectifs du chercheur<sup>3</sup> - et la plupart des gens ont du mal à comprendre pourquoi l'ordinaire peut être digne d'intérêt.

### **2.1 La diversité des données : comment**

L'hypothèse derrière la diversification recherchée ne recoupe ni la diversité en genres, ni la recherche de différents moments stylistiques qui était à l'œuvre dans les interviews laboviennes (moments du plus au moins formels, obtenus par la diversification des sujets traités).

Des entretiens peuvent-ils atteindre l'ordinaire ? Ils peuvent en tous cas y tendre, plus ou moins, car dès lors qu'on privilégie l'interaction, on est dans le qualitatif donc le continu. Faute de terme mieux adapté, on parlera ici d'*authenticité*, en étendant légèrement le raisonnement de Coupland 2003 ou de Palomarès & Tersigni 2001, car ces auteurs situent l'authenticité dans le fond, alors que le (socio)linguiste a aussi affaire à la forme.

La qualité d'un enregistrement n'est garantie ni par un quadrillage externe, ni par la reconduction de conditions d'interviews (script semi-directif, même lieu, même enquêteur)... Notre choix méthodologique est de ne pas compter sur des facteurs externes pour déterminer la validité d'un entretien, mais de chercher la perception de qualité (on a ainsi écarté environ 25% des enregistrements, pour des raisons qui vont de la qualité acoustique au sentiment d'inauthentique, de forçage, ou d'excès de "distance"). Les critères sur lesquels on fait reposer

---

<sup>2</sup> "En gros", car la frontière entre les trois types est souvent fragile, comme on peut s'y attendre. Voir Gadet/Guerin (2012) pour des précisions.

<sup>3</sup> Nous avons tout vu de ce point de vue, depuis les enquêtés qui ont d'abord accepté pour ensuite manquer tous les rendez-vous, jusqu'à l'enquêté qui se prend au sérieux et joue à l'enquêteur.

un tel jugement, qui fait l'objet d'évaluations et de discussions dans l'équipe, ne sont pas facilement généralisables (c'est pourquoi j'ai parlé de "se forger l'oreille")<sup>4</sup>.

Une certaine "qualité des enregistrements" peut être reliée (certainement pas de façon mécanique) à la familiarité avec un terrain/une personne. On peut ainsi s'appuyer sur des notions comme l'*histoire conversationnelle partagée* (Golopentia-Erescu 1988), la *communauté de pratique* (Eckert 2012) et leurs effets sur la co-construction du sens social, ou encore l'*immédiat communicatif* (Koch & Cesterreicher 2001). En voici deux exemples: Wajih qui a été 5 ans surveillant dans un collège de Mantes-la-Jolie, exploite ses relations avec des adolescents qu'il connaît depuis la 6<sup>e</sup>; Nacer, à l'association Zy'va de Nanterre, fait de l'aide aux devoirs et du soutien en arabe, ce qui facilite ses demandes d'entretien. Ces cas diffèrent d'enregistrements dans un même lieu, par un enquêteur qui n'a pas de connaissance préalable des enquêtés et qui suit un script<sup>5</sup>. Mais ces pratiques sont exigeantes en temps et en disponibilité personnelle, et elles limitent le rayon d'action.

L'enquêteur est donc confronté à la complexité et à l'intrication des facteurs en jeu, relevant de paramètres divers, peut-être non répertoriés et certainement non quantifiables. Ainsi, pour Nacer: il est Algérien, et n'appartient pas à la communauté des beurs qu'il interviewe. Mais son extériorité a à voir avec une identité "imaginée" des jeunes enquêtés, ce qui d'ailleurs peut jouer dans des directions opposées, l'extériorité pouvant susciter l'empathie favorisant le naturel, ou être perçue comme une distance qui entraîne surjeu ou absence de collaboration (en l'occurrence, cela a joué favorablement).

Pour mesurer l'impact de l'enquêteur, l'histoire de la sociolinguistique a retenu l'expérience de Rickford & McNair-Knox 1994, qui fait interviewer une jeune fille afro-américaine successivement par une enquêtrice de la même communauté, puis par une enquêtrice hors communauté, avec des résultats probants sur le plan quantitatif (taux de variables). Il reste néanmoins à s'interroger sur le sens social que l'enquêté peut prêter au renouvellement d'une interview, même à long terme (avec probablement plus ou moins les mêmes questions). Car les relations humaines ne rejouent jamais la même pièce, et les interactions véritables ne sont pas répliquables.

## 2.2 La diversité des données : pourquoi

Nous nous interrogeons sur ces aspects des données qui, pour la syntaxe, ne sont que rarement, ou insuffisamment, ou jamais, documentés dans les corpus construits sur la seule base d'entretiens, même dits informels. On fait ainsi l'hypothèse que des données obtenues dans des conditions différentes ne montreront pas tout à fait les mêmes faits linguistiques, ni selon les mêmes fréquences, saillances ou contraintes (voir Gadet 2008), en supposant des effets de la différence des modes d'obtention, en particulier entre données sollicitées et données écologiques. Mais de quel ordre seraient-elles ? Les phénomènes vont-ils se présenter en dichotomie, ou selon une gradation fine ?

Se poser de telles questions constitue une invitation à dépasser la notion de variable,

---

<sup>4</sup> L'entretien étant une relation intersubjective, ce n'est pas l'enquêteur qui est plus ou moins bon, et un même enquêteur peut "réussir" un entretien et en "rater" d'autres, sans parler des maladroites, par exemple quand un enquêteur prend à son compte une expression comme *de la beuh*, sans attendre que ce soit l'enquêté qui introduise ses propres mots.

<sup>5</sup> Entendons-nous : ces pratiques ne sont pas à rejeter en bloc, car elles donnent des résultats d'un intérêt incontestable. Elles sont seulement à compléter par d'autres modalités de recueil.

davantage concernée par la place structurelle des phénomènes que par la réactivité sociale des enquêtés. Le quantitatif répond en effet à une conception du style qui, tout en imageant bien quelque chose, demeure d'un fort degré d'abstraction, et qui ne suit pas forcément (ou pas du tout) les cheminements des interactants pour produire/interpréter du style.

La linéarité est-elle en effet une bonne hypothèse? Elle est certes conforme au mode de production/réception de l'oral, mais les interactants évaluent-ils le style selon des processus linéaires? Tenter ainsi d'adopter le point de vue du locuteur conduit à se demander ce qu'il "fait" avec le style. Comment en joue-t-il pour son propre compte? A-t-il une maîtrise sur sa capacité à moduler le style dans sa langue? Comment perçoit-il le style de l'interlocuteur? Est-ce que tout contribue à l'évaluation? Ou bien est-ce qu'au contraire il y a saillance de certains phénomènes? Et y a-t-il toujours cohérence? Ou une sorte de pondération entre traits, valorisants, discriminants, neutres?

Il y a là plus de questions que de réponses, mais ne pas se les poser reviendrait à adopter implicitement les réponses apportées par les prédécesseurs.

### **3 La palette stylistique d'un même locuteur**

Nous avons voulu écarter la tentation d'essentialisme, prenant un aspect d'une personne ou d'un groupe comme explication d'un comportement linguistique, en posant un lien entre identité collective et langue (Mendoza-Denton 2002).

#### **3.1 Des enregistrements en doublets**

MPF offre pour le moment 13 enregistrements dédoublés, offrant pour un même locuteur un enregistrement en interview de proximité et un enregistrement en situation écologique choisie par l'enquêté lui-même. Le recueil de ces situations écologiques s'avère extrêmement difficile, que ce soit en pratique d'auto-enregistrement ou dans la tentative pour faire comprendre à des non-linguistes ce que l'on attend d'eux<sup>6</sup>.

Il s'agit donc de deux situations d'immédiat, l'une des deux étant sollicitée et l'autre pas. Ce n'est pas un schéma méthodologique classique, car la pratique habituelle est plutôt de rechercher un écart maximal (formel *vs* informel), ce qui renvoie en fin de compte à une certaine conception de la langue où le tranché l'emporte sur un continuum subtil et réversible.

On peut s'interroger sur la "comparabilité" de données ainsi constituées, sachant que le style doit avant tout être conçu comme distinctif (ou contrastif). La comparabilité est en effet en général considérée comme assurée par des facteurs externes, alors que nous cherchons à construire cette notion de façon interne, sur des considérations d'authenticité linguistique et langagière.

#### **3.2 Stéphane, en entretien et en écologique**

Stéphane (pseudo) n'a rien de jeune, il a 36 ans (et aurait donc échappé à un quadrillage socio-démographique). Pourtant, son mode de vie et ses options fondamentales font qu'il se

---

<sup>6</sup> Il faudrait être naïf pour se persuader que l'auto-enregistrement obtenu dans ces conditions sera exempt de toute conscience. Pourtant, les productions des interactants sont encore une fois en continuum, et les réflexions sociologiques sur les groupes ont montré que le poids de la situation en groupe pouvait l'emporter sur la situation *in absentia*, celle de l'engagement envers le chercheur. La seule alternative est le micro caché (voir à ce sujet les réflexions de Gasquet-Cyrus à paraître).

vit comme "jeune". Nous disposons de deux enregistrements transcrits, "Stéphane entretien" (enregistrement pratiqué par un ami d'une amie d'enfance) et "Stéphane écologique" (conversation ordinaire, chez lui, avec un intime). Les deux enregistrements ont été retenus pour la qualité interne et leur authenticité.

*(1) Stéphane entretien*

Les french pour euh pour être international en fait c'est un délire de sape moi je reviens dans la musique en solo avec un délire de sapes de fringues tu vois où c'est vintage de 50 à 80 tu vois et en fait euh je réinvente le style je vais essayer de réinventer un style hip hop parisien tu vois le délire parce qu'on en a marre d'être habillés comme des blaireaux d'Américains là non mais ça allait ça allait quand c'était pas populaire maintenant c'est trop populaire ici n'importe quel blaireau ici il porte un il porte tout un jordan il porte un baggy tu vois avant nous ça avait une signification aujourd'hui ça en a plus non c'est c'est devenu un peu la la l'habit des beaufs tu sais ce que c'est un beauf<sup>7</sup>

Un beauf c'est le français moyen c'est euh c'est le français lambda non un bobo c'est un bourgeois bohème c'est un fils de pute qui a plein d'argent et qui va faire genre il en a pas et il s'habille avec des guenilles tu vois moi je suis le contraire moi je suis euh j'appelle ça boto bourgeois du ghetto on a pas d'argent mais faut que tout brille

*(2) Stéphane écologique (récit ordinaire fait à son ami Farid)*

St je sais pas qui vous êtes tu vois ce que je veux dire je leur ai fait comme ça (.) genre je parfois il y a des jeunes ils ont la haine sur nous hein

F Bien sûr

St Donc moi je vois vous vous me faites des zephas j'accélère hein je me casse hein après euh il me dit vous savez qu'on peut vous enlever le permis ben je fais bien sûr vous avez tous les droits je je connais vos droits monsieur et après bon que ça se reproduise pas

F Tu lui dis ouais peut-être que les jeunes ils voulaient me faire un quetru hein tu sais pas

St Non mais c'était eux les nejeu c'était eux les nejeu en <vrai>

F <Ah> ouais

St Moi je croyais que c'était des des des

F Tu lui dis peut-être que c'est oit hein tu voulais me faire un faire un quetru tu vois

St Je voyais ils étaient <despee>

F <Attends> tu t'annonces pas tu es un keuf

St Ben ouais je dis attends euh vous êtes en vilci vous jouez les cailleras

F Ben ouais

St Non mais trop de fois man hein franchement je me serais fait autant arrêté mais comme ça en privé en civil je me serais chié dessus vingt cinq fois (.) mais quand tu sais quand tu as le truc Securitas tu sais ils savent tu vailletras tu vois

### 3.3 La comparabilité des énoncés

Les deux enregistrements diffèrent de plusieurs manières. Une caractérisation en termes de variables<sup>8</sup> n'est certes pas incongrue ici: une seule occasion de *ne*, absent en (1), davantage en (2) (mais c'est surtout l'absence d'énoncés négatifs en (1) que cela désigne - différence de genres discursifs?); trois absences de *que*, en (2) (*je voyais ils étaient despee; tu t'annonces pas tu es un keuf; ils savent tu vailletras*), aucune en (1) (mais même remarque)...

<sup>7</sup> L'enquêteur, Roberto, est italien, et il a adopté une attitude de type "aide-moi à mieux comprendre les banlieues françaises".

<sup>8</sup> Il est un peu malhonnête de recourir à la notion de variable sur des exemples brefs, parce qu'elle est justement destinée à traiter des corpus longs, avec l'idée que le quantitatif aplanit d'éventuelles idiosyncrasies.

Est-ce toutefois ce qui est décisif? Ces remarques n'épuisent pas les différences, devant des facteurs plus difficiles à repérer, comme la connivence et les allusions, opaques à un observateur extérieur, qui concernent peu des points précis mais plutôt des facteurs comme le débit et le rythme, la courbe intonative et plus particulièrement la présence de certains contours; le lexique, en particulier la présence de beaucoup de verlan (*nejeu, quetru, oit, keuf, cailleras, despee, vailletras, vilci, zephas*), plus spécifique que l'argot que l'on rencontre aussi en (1)<sup>9</sup>.

On peut donc se demander en quoi la syntaxe et le discours contribuent au jugement de différences entre (1) et (2). Peut-on l'interpréter en termes de saillance? (comme *genre* ou *même*), de cumul? (accumulation de traits non-standard)? Est-ce interprétable perceptivement par une oreille extérieure au contexte?

Il y a là de vraies questions.

## Conclusion

Qu'a-t-on appris sur le style? On a opposé la pratique de la variable au tissage de phénomènes, convergents ou divergents, cohérents ou non, qui contribuent à produire une évaluation pour les interactants (dont le linguiste ne fait que mimer l'attitude), et jouent un rôle dans la constitution du sens social. Le style, ou le stylage (*styling*) est une activité dynamique, un processus de co-construction, plus qu'un produit stabilisé.

Et qu'a-t-on appris sur les corpus? Que faudrait-il que soient les corpus pour mieux servir une réflexion sur le style? Il n'y a toujours pas assez de corpus de français (de tous les français), et ils ne sont pas suffisamment variés. Mais il n'y a pas non plus suffisamment d'exigences méthodologiques, pas suffisamment de sensibilité aux conditions d'enquêtes, auxquels pourtant les sociologues ou les ethnologues sont si sensibles (Bizeul 2007); et les linguistes se laissent trop souvent aller à la facilité de reconduire ce qu'ont fait les prédécesseurs. Plusieurs sociolinguistes ont d'ailleurs fait remarquer récemment qu'il était de moins en moins de bon ton d'afficher ses problèmes méthodologiques (contrairement à ce que font les ethnologues).

Les considérations sur le terrain ne doivent certes pas conduire à considérer qu'il faut empiler n'importe quoi comme données, mais au contraire, qu'il faut prêter la plus grande attention à chacun des gestes méthodologiques pratiqués par le linguiste, chaque choix pouvant avoir des conséquences dont l'impact est difficile à mesurer avant la phase d'exploitation, moment où il sera trop tard pour s'apercevoir qu'on a laissé passer quelque chose.

## Références

AUER P. (2007 Ed.). *Style and Social Identities. Alternative Approaches to Linguistic Heterogeneity*. Berlin: Mouton de Gruyter.

BELL A. (1984). Language Style as Audience Design, *Language in Society* 13, 145-204.

BIZEUL D. (2007). Que faire des expériences d'enquête? Apports et fragilité de l'observation directe. *Revue française de science politique*, 57, 69-89.

---

<sup>9</sup> Le verlan, omni-présent en (2), est-il évité en (1) parce que Stéphane pense que Roberto ne comprendra pas ? Impossible de répondre à cette question. Pour le lexique, voir le *Dictionnaire de la zone*.

BULOT T. (2004). Les parlers jeunes et la mémoire sociolinguistique. Questionnements sur l'urbanité langagière. *Cahiers de Sociolinguistique* 9, 133-47. Rennes : Presses Universitaires de Rennes

CAPPEAU P., GADET F. (2007). L'exploitation sociolinguistique des grands corpus. Maître-mot et pierre philosophale, *Revue Française de Linguistique Appliquée* XII-1, 99-110.

COUPLAND N. (2003). Sociolinguistic authenticities, *Journal of Sociolinguistics* 7, 416-31.

COUPLAND N. (2007). *Style. Language variation and identity*, Cambridge University Press.

*Dictionnaire de la zone*. [www.dictionnairedelazone.fr](http://www.dictionnairedelazone.fr)

ECKERT P. (2012). Three Waves of Variation Study: the Emergence of Meaning in the Study of Sociolinguistic Variation. *Annual Review of Anthropology* 41, 87-100.

GADET F. (2010 [2008]). Les corpus oraux et la diversité des productions langagières, *Verbum*, Tome XXX n°4, 261-73.

GADET F., GUERIN E. (2012). Des données pour étudier la variation : petits gestes méthodologiques, gros effets. *Cahiers de linguistique* 38-1. 41-65.

GASQUET-CYRUS M. (à paraître). Cachez ce micro (caché) que je ne saurais voir.

GOFFMAN E. (1979). Footing, *Semiotica* 25, 1-29, tr. fr. 1981, La position, *Façons de parler*, Editions de Minuit.

GOLOPENTIA-ERESCU S. (1988). Interaction et histoire conversationnelle, in *Échanges sur la conversation*, J. Cosnier, N. Gelas et C. Kerbrat-Orecchioni (Dir.), Paris : Éditions du CNRS. 69-81.

HABERT B. (2000). Des corpus représentatifs : de quoi, pour quoi, comment ? In Bilger M. (éd.), *Linguistique sur corpus - Etudes et réflexions*, Perpignan, Presses Universitaires de Perpignan.

KOCH P. & CESTERREICHER W. (2001). Langage parlé et langage écrit, in *Lexikon der romanistischen Linguistik*, tome 1, 584-627, Tübingen, Max Niemeyer Verlag.

MENDOZA-DENTON N. (2002). Language and Identity, in Chambers *et al*, *The Handbook of Language Variation and Change*, Oxford, Blackwell Publishing, 475-99

PALOMARES E. & TERSIGNI S. (2001). Les rapports de place dans l'enquête: les ressources du malentendu. *Langage & Société* 97, 5-26.

RICKFORD J. & MCNAIR-KNOX F. (1994). Addressee- and topic-influenced style shift: a quantitative sociolinguistic study, in D. Biber and E. Finegan (Eds) *Sociolinguistic Perspectives on Register*, New York: Oxford University Press, 235-76.

THIBAUT P. & VINCENT D. (1990). *Un corpus de français parlé*, Québec, Recherches sociolinguistiques 1.

## **Methodological considerations in the study of sociophonetic variation in an underdocumented minority language: Somali Bantu Kizigua as a case study**

Holman Tse<sup>1</sup>

(1) University of Pittsburgh  
hbt3@pitt.edu

**Abstract** This paper presents an example of how research can be done on sociophonetic variation within a community of speakers of an underdocumented minority language. The paper begins with a discussion and critique of the Labovian sociolinguistic interview, which has formed the basis of a large body of work on variation. Current research on Somali Bantu Kizigua, a minority language from East Africa, is discussed. A modified Labovian sociolinguistic interview was implemented for collecting data on the language. Results from this study reveal evidence of a sound change in which a historically voiceless prenasalized velar stop, \*/nk/, becomes a voiceless glottal fricative, /h/, with uvular consonants as intermediate stages. This change appears to confirm speaker's intuitions about age as a major factor in variation within the community. The paper concludes with a discussion of opportunities for increased collaboration between documentary linguists and sociolinguists.

**Keywords:** sociolinguistics, documentary linguistics, obsolescence, sound change, fieldwork, Zigula, Mushunguli

## 1 Introduction

Variation and variability have long been central concerns of sociolinguistics. Yet, as much as variation has defined the particular branch of sociolinguistics known as variationist sociolinguistics, there has ironically been little variation in the range of languages and in the range of communities studied. Stanford and Preston (2009), for example, discuss the lack of quantitative variationist sociolinguistics research in indigenous minority language communities even as there is increasing interest among researchers in other areas of linguistics in working on them. Meyerhoff and Nagy (2008) have also highlighted a monolingual bias in variationist sociolinguistics. In their survey of two leading journals, they found that only 11% of articles in *Language Variation and Change* and 28% of articles in the *Journal of Sociolinguistics* focus on more than one language. While Stanford and Preston (2009) as well as Meyerhoff and Nagy (2008) discuss different but not necessarily mutually exclusive types of communities, indigenous minority communities in the former and multilingual communities in the latter, both volumes highlight the need for more variationist sociolinguistics research in a greater variety of communities. Meyerhoff and Nagy (2008) suggest that one reason why there is not more diversity in the communities researched by sociolinguists is hegemonic monolingualism in North America, where a majority of sociolinguists are trained. Even if this really does account for the research gap, however, hegemonic monolingualism has clearly not been a hurdle in the increasing interest in work on endangered languages among documentary and theoretical linguists. Nagy (2012) highlights this disparity by showing that 83% of articles in the journal *Phonology* are focused on languages other than English. In contrast, this is the case for 47% and 38% of articles in *Language Variation and Change* and in the *Journal of Sociolinguistics* respectively.

In this paper I argue that research on basic variationist questions in a minority language does not require fluency in the language, although it would be highly desirable. I present an example of how this can be accomplished by discussing current research on Kizigua<sup>1</sup>, an underdocumented minority language spoken within a multilingual community of Somali Bantu refugees in the US. While sociophonetic variation is the main focus of the current project, documentation of the speech of an individual consultant still proved essential in gaining background knowledge of the language and in identifying parts of the phonological system worth further study. A modified Labovian sociolinguistic interview was implemented that revealed evidence of community-wide sound changes in progress. By discussing the results, I argue that variation can and should be studied at an early stage of language documentation. Doing so provides opportunities for greater collaboration between documentary linguists<sup>2</sup> and sociolinguists with mutual benefits for both parties and opportunities to fill in a major research gap.

---

<sup>1</sup> Various names and spellings for Kizigua have been used in the literature including Zigula, Seguha, Wayombo, Wazegua, Zeguha, Zegura, Zigoua, Zigwa, and Chizigula. In addition, Mushungulu and Mushunguli are alternate names for the language used in Somalia. The dialect discussed in this paper is referred to as Somali Bantu Kizigua to distinguish it from more conservative varieties of Somali Kizigua.

<sup>2</sup> Many terms have been used to describe work on undocumented languages. Aside from “documentary linguistics”, other terms include “field linguistics”, “descriptive linguistics”, and “anthropological linguistics”. In this paper, all of these terms are treated as synonymous. What is most crucial is that sociolinguistics is widely viewed as a field of inquiry distinct from documentary linguistics as Munro (2008: 130) mentions.

## **2 Paradoxes in Sociolinguistic Research Methodology**

The basis of sociophonetic data collection has long been the Labovian sociolinguistic interview, which generally consists of at least two parts: a casual conversation with the subject and structured elicitation tasks such as word lists and reading passages. These different parts were designed to elicit different styles of speech with the assumption that speech varies along a single dimension. On one end is “vernacular” style, which involves the least conscious attention paid to speech. On the other end is “formal” style, which involves the most attention paid to how one speaks. As Labov (1984) discusses, “vernacular” style is the most desirable for sociolinguists because it is the most natural way of speaking. Formal speech, on the other hand, is described as less natural because it is easily affected by conscious attention paid to speech. Labov used the term “Observer’s Paradox” (1972) to describe the fact that the goal of the sociolinguist to get data on how speakers speak when they are not being observed runs directly at odds with the fact that the researcher can only get this data by observing the speaker. The casual conversation part of the interview was, thus, designed to make subjects feel as comfortable as possible so that they can speak with a more vernacular style and not think about the fact that they are being recorded. For sociophonetic analysis, however, casual speech does not always produce enough samples of the variable of interest. So, word lists and reading tasks are included to ensure that at least a few tokens of the variables of interest are collected even if read in a more formal style.

In the context of working on an underdocumented minority language, a major problem with variationist research is the fact that the sociolinguistic interview was not designed for use in working on a minority language or in a multilingual setting. For example, having a conversation with subjects in an attempt to gather naturalistic phonetic data is not possible unless the researcher is fluent in the language. Thus, there is an inherent problem in the design of the sociolinguistic interview that makes it difficult to increase the number of studies of underrepresented languages unless there is an increase in the number of researchers proficient in these languages. While one could argue that the best research methodology should always depend on the given context and the specific questions, some methodological consistency is also desirable for ensuring comparability between different studies. Does this mean that sociophonetic research on a language can only proceed if the researcher or interviewer is proficient in the language? Are language documentation and the study of sociophonetic variation two incompatible goals?

First of all, assuming that language documentation and the study of sociophonetic variation need to be two separate projects and that complete documentation must occur first assumes that language documentation can only involve describing a language as a static entity without any internal variation. This, however, would be a view of language that would be untenable for an increasing number of researchers across all the language sciences. Although it would be desirable for the researcher to be as proficient as possible in the language of interest, there is the question of how proficient the researcher needs to be to carry out such research. Even if the researcher were fluent in the language, however, there would still be a certain degree of unnaturalness in the context of an interview. Minority languages are spoken by a small community, which means that there is often a tighter link between the social meaning of speaking in the language and one’s ethnic identity. The researcher’s identity as an outsider could be virtually impossible to overcome in many situations. Furthermore, in a multilingual context, the most “natural” way for a minority language speaker of communicating with a researcher would be to use a lingua franca rather than the language of interest.

Although minority languages have been researched using sociolinguistic techniques, very few of these studies have focused on sociophoentic variation within the context of questions about language variation and change. There have been some studies that have examined variation and change in the context of language obsolescence, however, such as Schmidt's (1985) study of Young People's Dyrbal and Dorian's (1978) study of East Sutherland Gaelic. Both of these studies focused on the type of structural changes that occur when languages are in a near death situation. The research presented in this paper, however, differs from these studies in several ways. First of all, obsolescence is not the main focus of the present study. All of the speakers interviewed for this study are fully proficient in the language. Although obsolescence is occurring in the youngest generation of speakers of the Somali Bantu dialect, the language is still generally much more robust than Dyrbal and East Sutherland Gaelic. Secondly, much less prior documentation on Kizigua exists than what was available for Dorian (1978) and Schmidt (1985). In the next section, I present relevant background information on the little that is known about the sociolinguistic context of Kizigua.

### **3 Background on Somali Bantu Kizigua**

The history of the Kizigua language as well as the social history of its speakers is a complex one that involves two major periods of geographical displacement in the past two centuries. As a Bantu language related to Swahili, Kizigua traces its origins to Northeastern Tanzania where it is still spoken today by the Zigua tribe. Ethnologue reports an estimated 355,000 speakers present in Tanzania (Lewis, 2009). Beginning in the 1840's and continuing through the end of the 19<sup>th</sup> Century, however, Arab-Omani traders brought people from various East African tribes including the Zigua to Somalia where they were forced to work as slaves. Between 1865 and 1895, an estimated 20,000 slaves managed to escape (Eno and Eno, 2007). Without the practical means to return to their homelands, however, fugitive slave communities were formed primarily in the Lower Juba River Valley area of Southern Somalia. These communities have remained in Somalia ever since.

The second major displacement of the Zigua people occurred in the 1990's with the outbreak of the Somali Civil War. At this time, the Zigua as well as members of other oppressed minority groups, fled across the border to Kenya where they lived in UN refugee camps for many years. These refugees became collectively known as the "Somali Bantu". Some of them were subsequently able to resettle in other countries including the US, which accepted 12,000 Somali Bantu refugees in 1999 (Besteman, 2012). Major Somali Bantu communities are now present in various American cities including San Diego (California), Boise (Idaho), Columbus (Ohio), Buffalo (New York), and Pittsburgh (Pennsylvania).

With all of this history of movement in less than two centuries, it should be no surprise to find the Kizigua spoken by the Somali Bantu refugees and the Kizigua spoken in present-day Tanzania to be different. The uncertain question is how different and whether or not they are still mutually intelligible. In Somalia, the Zigua came in contact with speakers of various Cushitic languages including Somali, Maay, and Oromo. Many Somali Bantu refugees also report having learned Swahili when they lived in Kenyan refugee camps. All of these languages may have contributed in varying degrees to divergence from Tanzanian Kizigua although the extent to which this occurred is currently unknown. Furthermore, given the lack of documentation on the language and the lack of face-to-face contact between Somali Bantu refugees and the Zigua in Tanzania, the degree of mutual intelligibility is also unknown. Crevatin (1993), however, describes the Kizigua in Somalia as generally well preserved from Tanzanian Kizigua based on fieldwork in Somalia conducted between 1986 and 1988. The

few differences identified were lexical and phonological. Since Crevatin (1993), the only other known studies of Somali Kizigua are works in progress by Temkin Martinez (2011) and Odden (2013). Both of these studies involve speakers in the US. The literature on Tanzanian Kizigua is only a little bit more plentiful. This includes a recently published dictionary (Mochiwa, 2008), a dictionary published at the turn of the 20<sup>th</sup> Century (Kisbey, 1906), and various articles on the tonal phonology of the language (Kenstowicz, 1988; Kenstowicz, Kisseberth 1990; Kisseberth, 1992).

## 4 The Data

### 4.1 Field Methods Work and Motivation for Current Study

The current project developed as part of a Field Methods course at the University of Pittsburgh, which took place between January and April 2012. This course met with a native speaker of Kizigua, who was compensated for a total of 3 hours per week and for meeting with smaller groups for a total of three additional hours per week outside of the regularly scheduled class time. The consultant was a 21-year old female native-speaker of Kizigua who has a personal history quite similar to that of many other Somali Bantu refugees. She was born in Somalia and fled across the border to Kenya when she was little to escape the Somali Civil War. After spending some time in Kenyan refugee camps, she was able to resettle in the US in 2004. The consultant has continued using Kizigua as her primary language at home although she has also been able to gain a good level of proficiency in English. She also learned a little bit of Maay, Somali, and Swahili while growing up in East Africa.

During the course of the four months of working with the consultant, a corpus of recorded materials was created. These recordings include pronunciations of lexical items, sentences, stories, and other samples of connected speech. All of them were made using a Marantz solid state recorder and a headset with an attached Shure Microphone. The sampling frequency used for all recordings was 10,000 Hz. A total lexicon of about 700 words was also compiled.

Upon consulting sources of Tanzanian Kizigua, one of the major differences noted between the consultant's speech and the language described by Kisbey (1906) was the pronunciation of the reflexes of the voiceless prenasalized stops. Crevatin's (1993) study of Somali Kizigua in the 1980's also noted a sound change described as follows in which all of the voiceless prenasalized stops become voiceless aspirated stops:

$$(1) *mp, *nt, *nk > p^h, t^h, k^h$$

For the historic prenasalized labial stop, there were too few examples elicited to be able to make any broad generalizations. Examples of the reflexes of coronal prenasalized stops, though, are more plentiful. Unlike what Crevatin (1993) reported, however, the coronal prenasalized stops appear to be more retroflex than alveolar with a trill-like aspiration. The phonetic and phonological properties of these stops are worth further investigation. Of more crucial relevance for this paper are the reflexes of the voiceless prenasalized velar stop. As shown in Table 1 below, the reflex of \*nk is not /k<sup>h</sup>/ as reported by Crevatin (1993) but either a voiceless uvular stop, /q/, or a voiceless glottal fricative, /h/. While /h/ seems restricted to word-initial position, /q/ can be found both word-initially and word-medially. Word-initially, however, the historic nasal is lost while it is retained with the presence of a uvular stop in word-medial position as in [kuinqa] and [cinqo].

Consultant's Speech	Tanzanian Kizigua	English Gloss
kuinqa	kuinka	'to give'
kununqa	kununka	'to smell'
cinqo	kinko	'elbow'
qumbito	nkumbito	'eyebrows'
hande	nkande	'food'
honde	nkonde	'action of planting'
-hundu	-nkundu	'red'
hundu	nkunde	'bean'

Table 1: Examples of words with historic \*/nk/ and their reflexes

The question that arises from this data is whether the consultant's distribution of /q/ and /h/ is unique or whether it can also be found more broadly across all present-day Somali Bantu Kizigua speakers. Could there be variation based on demographic factors such as age, gender, or proficiency in Cushitic languages?

## 4.2 Additional Subjects

Following completion of the Field Methods course, additional speakers of Kizigua were sought to determine whether the prenasalized stop reflexes have the same realization among all speakers of Somali Bantu Kizigua or whether they are unique to the speech of the consultant. Three volunteers were eventually recruited for interviews lasting between one and two hours. The recruitment method used was to contact a local organization involved with the Somali Bantu community for assistance. All three of these subjects were male and ranged in age from 23 to 30 at the time of interview. All of them also have very similar histories of migration. Like the consultant, all of them were born in Somalia and fled across the border to Kenya where they settled in refugee camps until they were able to move to the US in 2004. All subjects are also proficient in the same set of languages including Maay, Somali, and Swahili as well as English.

## 5 Methodology

The methodology of this study was designed to address the question of whether or not there is variation in the phonetic production of the historic prenasalized stops in Somali Bantu Kizigua and whether or not this variation correlates with demographic factors. It involved a three-part interview that included the following: (1) an elicitation task, (2) a word-list task, and (3) a semi-structured interview to obtain sociological information. All recordings were made using the same equipment used for the Field Methods consultant discussed in Section 4.1. The researcher served as the interviewer.

### 5.1 The Pear Film Elicitation Task

The first part involved showing each interviewee *The Pear Film* (Chafe, 1980) and asking the interviewee to re-tell the story in Kizigua. *The Pear Film* is a 6-minute video created by a team of linguists interested in cross-linguistic similarities and differences in the way people talk about things that they have experienced. Its original purpose was to create a stimulus that

can be used for speakers of any language so that it would be possible to study how people with different linguistic backgrounds talk about the same events. For the purpose of this study, the video was used simply to elicit a sample of naturalistic speech. The recorded speech sample can later be used for acoustic analysis. The version of the film used was downloaded from the Chinese Pear Stories website (Erbaugh, 2001). The film was played on a laptop and shown in its entirety to each subject. The subject was then recorded re-telling the entire story from memory.

## **5.2 Word Lists**

As in the Labovian sociolinguistic interview, a word list task was also included. Presenting the word list in written form as in the Labovian sociolinguistic interview, however, would not work since there is no official orthography for Kizigua. Instead, the words were elicited using pictures as prompts. Although English could have been used, pictures were used instead to ensure a consistent methodology for cases in which the subject does not speak English. Subjects were asked to say each word three times followed by a carrier phrase containing the word of interest. The carrier phrase was as follows:

(2) Silonga naho ‘X’. (‘I say X again’).

The word list included all the words shown in Table 1 as well as other words that historically have a prenasalized stop for a total of 77 words. An attempt was made to select words that can easily and unambiguously be represented in visual form. This, of course, meant that the word list was generally limited to nouns. Speakers, however, sometimes varied in their interpretation of the pictures. In these cases, the researcher made an effort to elicit alternate words but if it was clear a particular word is not present in an individual speaker’s lexicon, the researcher would continue on to the next word. Due to these individual differences, the total number of words recorded for each speaker was not the same.

## **5.3 Semi-structured Interview**

Finally, the third and final part involved a semi-structured interview that consisted of questions about language use within their families, contact with speakers of other languages, communication with relatives, and their thoughts and observations about variation in Kizigua. This interview was conducted in English. This made it possible to collect information that would be relevant for understanding the possible sociological factors that may be at play in accounting for variation without requiring the researcher to be fluent in Kizigua. Unlike for the Labovian sociolinguistic interview, this conversation did not need to be acoustically analyzed and was designed for information collection purposes only. Concerns about the Observer’s Paradox were, thus, not relevant since speech content was more important than speech form.

# **6 Results and Evaluation of Methodology**

## **6.1 The Pear Film Elicitation**

*The Pear Film* proved to be quite successful in eliciting naturalistic speech as evidenced in the frequent reduction of consonant and vowel segments for all speakers. Interestingly, all speakers took about the same amount of time to re-tell the entire story. The average length of about 2 minutes, however, was too short to elicit an adequate number of tokens for statistical

analysis. While there were plenty of tokens of words containing reflexes of \*/nt/, primarily due to the frequent use of the word [nɲangulu] ('basket'), there were hardly any words containing reflexes of \*/nk/. Getting enough tokens of the variables of interest is, of course, the same problem with the casual conversation part of the Labovian sociolinguistic interview. For this reason, the word list as supplemental data proved to be crucial for this study.

## 6.2 Word List

While the word list task was successful at getting a small sample of words that contain reflexes of historic \*/nk/, one shortfall of this technique was that subjects did not always produce the intended target word since the pictures sometimes opened room for multiple interpretations. Getting stable pronunciations in the carrier phrase was also a challenge. Subjects sometimes lost track of the number of repetitions and occasionally changed their rate of speech to compensate for time lost once they realized they made a mistake. Subjects also occasionally varied the form of the carrier phrase suggesting that they attempted to think about the task in a communicative way rather than as simply a phrase that they were supposed to fill in as a consistent frame. In spite of these challenges, the Word List Task still proved successful at revealing evidence for a change in the historic \*/nk/ as shown in Table 2 below.

Subject 2 (b. 1982)	Subject 1 (b. 1984)	Subject 3 (b. 1989)	Consultant (b. 1991)	English Gloss
not recorded	kuinqa	kuinqa /kuinha	kuinqa	'to give'
not recorded	cinqo	cinxo	cinqo	'elbow'
not recorded	humbito / qumbito	humbito / qumbito	humbito / qumbito	'eyebrows'
qande / xande	qande	qande / hande	hande	'food'
qonde	qonde	honde / qonde	honde	'action of planting'
-qundu	-hundu / qundu	-hundu	-hundu	'red'
qunde / xunde	qunde / hunde	hunde	hunde	'bean'

Table 2: Pronunciation of reflexes of \*/nk/ for all subjects

The table indicates how each subject pronounced the reflex of historic \*/nk/ in a selection of words that is also included in Table 1. The columns are ordered from oldest speaker on the left to the youngest speaker on the right. If a speaker had two different pronunciations during the Word List Task, both variants are listed. Otherwise, if the speaker was consistent, only one form is listed. As can be seen, the consultant was the most stable. This could be due to the fact that she had more experience with working with linguists and made a more conscious effort to accommodate. Even so, she pronounced 4 of the 7 words consistently with the glottal fricative /h/ while the oldest speaker consistently produced these same words with a voiceless uvular, although the exact realization varied between a stop, /q/, and a fricative, /χ/. The second oldest speaker did not have any words that were consistently pronounced with /h/ while the second youngest did. Thus, what we see in the Table above is evidence for a sound change in progress that appears to be proceeding as in (3) below. Also included in (3), but not attested in the speech of any of the subjects is /k<sup>h</sup>/, which was reported by Crevatin (1993).

(3) \*nk > k<sup>h</sup> > q > χ > h

### 6.3 Semi-structured Interview

The interview revealed useful sociological information to contextualize the use of the language. One question that was asked of all subjects that seems most relevant for this paper was whether or not they recognize different ways of speaking Kizigua. All subjects consistently mentioned age as the biggest source of variation. They discussed a high rate of language obsolescence among the youngest members of the community. This would include children who were born in the US or who came to the US at a very young age. Some of the subjects also mentioned anecdotes of children who have completely shifted to English and are unable to communicate with their non-English speaking grandparents. The subjects interviewed for this study appear to be part of a generation sandwiched between a younger generation that is quickly losing the language and an older generation that generally has less proficiency in English and speaks what is described as a more “pure” form of Kizigua. They consciously recognize that some of the words they use are borrowed from Somali, Swahili, and other languages. They also recognize that their parents would sometimes use words that are different from the borrowed words that they would use.

## 7 Discussion

What seems most remarkable about the results of this study is evidence of a rather abrupt sound change. While it seems phonetically plausible for /q/ to weaken to /h/, the speed at which this occurs seems rather quick given the fact that there is only a nine-year age difference between the oldest and youngest speakers. Assuming a similar pattern holds for other speakers, the best possible explanation is perhaps the Somali Civil War, which began in the 1990’s. This was when the two youngest speakers were born and when the the oldest speakers entered their teenage years. It seems quite likely that acquiring language during this politically unstable time may account for why the two youngest speakers exhibit a different pattern from the two oldest speakers. For all 4 speakers, the 1990’s were a time during which they and their families moved around frequently to escape the violence of the Civil War.

One question to ask about the change from /q/ to /h/ is whether or not this leads to a loss in phonological contrast in word-initial position. /h/ was historically present in Tanzanian Kizigua (Kisbey, 1906) and is retained in Somali Bantu Kizigua so it seems possible. Actual minimal pairs, however, are completely lacking in the data. There is also one piece of evidence that suggests that a phonological merger has not taken place. The word for ‘damage’ in Tanzanian Kizigua is [hasala] (Kisbey, 1906; Mochiwa, 2008) and is pronounced by the youngest speaker as [qasara]. This is the only example of a word that historically began with /h/ that is pronounced as /q/. Data on this word from the other three speakers is unavailable. Other words that historically began with /h/ are largely function words and morphemes that seem to retain their pronunciation as /h/ for all of the four speakers interviewed. A possible direction of change that seems to be emerging is that nouns that historically began with \*/nk/ (or /q/ for the oldest speakers interviewed) are being pronounced with /h/ while nouns that historically began with \*/h/ are being pronounced as /q/. Thus, this appears to be a reversal of /q/ and /h/ leading to the maintenance of a phonological contrast rather than a merger. There may also be some interaction with the morphology, which would account for different patterns of change that seem to vary based on lexical category. More data is needed before any further claims can be made. What is clear from the data collected so far is that there is change occurring in the reflexes of historic \*/nk/ that may involve various factors.

## 8 Conclusion and Future Research Directions

One of the goals of this paper was to illustrate an example of how research on variation can take place on an underdocumented language at an early stage of the documentation process. In this case, a significant pattern worth further investigation was identified through consultant work. The results of this study show evidence of a sound change in progress in Somali Bantu Kizigua. The sample size is small, but it proved adequate in showing an age-based pattern that matches the intuitions that subjects have about age being the most salient factor in accounting for variation. The next step would be to collect more data from speakers of different ages to see if this pattern holds. More samples of connected speech would also be helpful. Recruiting native speaker interviewers or simply recording two speakers having a conversation with each other is also a possibility. This would allow for collection of longer speech samples.

While communicative competence for the researcher was clearly not needed to identify a pattern, some background knowledge was still essential. This background was gained through consultant work with a native speaker. This highlights an opportunity for more sociolinguists to work with documentary linguists, who, of course, have already become the experts in learning about minority languages from scratch. Documentary linguists have already built up a large repertoire of techniques used to elicit linguistic data. Sociolinguists can benefit from adopting some of these techniques in their study of variation, which would also be helpful in learning more about the structure of languages of interest. Although variation has not been completely ignored by documentary linguists, it has also not been a major focus. Thus, documentary linguists can also benefit from working with sociolinguists by doing more to study variation as part of the documentation process. Such collaboration can work towards filling in a major research gap.

## Acknowledgements

I would like to thank the University of Pittsburgh Dietrich School of Arts and Sciences for support, the Spring 2012 Field Methods class at the University of Pittsburgh, David Mortensen, Scott Kiesling, Monica Duffell, Dave Odden, an anonymous reviewer for *JéTou* 2013, and most importantly of all, the Somali Bantu Community Organization of Pittsburgh as well as all subjects who participated in this study.

## References

- BESTEMAN C. (2012). Translating race across time and space: The creation of somali bantu ethnicity. *Identities-Global Studies in Culture and Power* 19, 285-302.
- CHAFE W.L. (1980). *The pear stories: cognitive, cultural, and linguistic aspects of narrative production* Norwood, N.J: Ablex Pub. Corp.
- CREVATIN F. (1993). Incontri con l’afrika. *Studi italiana di linguistica teorica ed applicata*, 22:1, 11-26.
- DORIAN, N. (1978). The fate of morphological complexity in language death: Evidence from east sutherland gaelic. *Language* 54.3: 590-609.
- ENO, O., ENO, M. (2007). The journey back to the ancestral homeland: the return of the somali bantu (wazigwa) to modern tanzania. In Kusow, Abdi & Stephanie R. Bjork (eds.) *From mogadishu to dixon: the somali diaspora in a global context*. Trenton, NJ: Red Sea Press.

- Methodological considerations in the study of sociophonetic variation in an underdocumented minority language: Somali Bantu Kizigua as a case study*
- ERBAUGH M. (2001). *The chinese pear stories : Narratives across seven chinese dialects* Website. <http://www.pearstories.org>
- KENSTOWICZ M. (1988). Tone and accent in kizigua-a bantu Language. In Bertinetto P.M., Loporcaro M., *Certamen phonologicum*, 177-88. Turin: Rosenberg & Sellier.
- KENSTOWICZ M., KISSEBERTH C. (1990). Chizigula tonology: the word and beyond. In Inkelas S., Zec D. (eds) *The phonology-syntax connection*, 163-194. Chicago, IL: University of Chicago Press.
- KISBEY W.H. (1906). *Zigula-English dictionary. Compiled for the universities' mission to central africa*. London: Society for Promoting Christian Knowledge.
- KISSEBERTH C. (1992). Metrical structure in zigula tonology. In Gowlett D. (ed.) *African linguistic contributions*, 227-259. Pretoria: Via Afrika.
- LABOV W. (1972). *Sociolinguistic patterns*. Philadelphia : University of Pennsylvania Press.
- LABOV W. (1984). Field methods of the project in linguistic change and variation. In Baugh J., Sherzer J. (eds) *Language in use: Readings in sociolinguistics*. 28-53 Englewood Cliffs, NJ : Prentice-Hall.
- LEWIS M. P. (2009). *Ethnologue: Languages of the World. 16<sup>th</sup> Edition*. Dallas, TX : SIL International.
- MEYERHOFF M., NAGY N. (2008). *Social lives in language – Sociolinguistics and multilingual speech communities: Celebrating the work of Gillian Sankoff*. Amsterdam : John Benjamins Publishing Company.
- MOCHIWA Z.S.M. (2008) *Kizigula: msamiati wa Kizigula-Kiswahili-Kiingereza = Zigula-Swahili-English lexicon* Dar-es-Salaam: Languages of Tanzania Project University of Dar-es-Salaam.
- MUNRO P. (2008). Field linguistics. In Aronoff M., Rees-Miller J. (eds). *The handbook of linguistics*. John Wiley & Sons.
- NAGY N. (2012) Sociolinguistics and phonology. In Bayley R., Lucas C., Cameron R. *Oxford Handbook of Sociolinguistics*. Oxford University Press.
- ODDEN D. (2013). The Mushunguli – Chizigua language of Somalia. Website <http://www.ling.ohio-state.edu/~odden/mushunguli/>
- SCHMIDT A. (1985). *Young people's dyirbal: An example of language death from australia*. Cambridge: Cambridge University Press.
- STANFORD J., PRESTON D. (2009). *Variation in indigenous minority languages*. Amsterdam : John Benjamins Publishing Company.
- TEMKIN MARTINEZ M. (2011). The use of mobile devices for language documentation. *Mobile learning symposium*. Boise State University.

## **L'alternance entre [β] et [w] dans les terminaisons de l'imparfait de l'indicatif en catalan dans la région de La Selva**

Eva Bosch i Roura

Departament de Filologia Catalana, Universitat de Barcelona  
eva.bosch.roura@ub.edu

**Résumé** Cette communication porte sur l'alternance entre [β] et [w] dans les terminaisons de l'imparfait de l'indicatif de la première conjugaison produites par des informateurs catalanophones de deux générations de la région de La Selva, au nord-est de la Catalogne. L'étude est structurée en une description des données et une analyse de la variation en relation avec la distribution géographique du phénomène dans la région autant qu'en relation avec la distribution générationnelle de l'alternance.

**Abstract** This communication deals with the alternation between [β] and [w] in the endings of the 1<sup>st</sup> Conjugation Imperfect Indicative verb forms produced by Catalan-speaking informants belonging to two group ages in the region of La Selva, in North-Eastern Catalonia. The paper describes the data and proposes an analysis of the variation, both in terms of the geographical distribution of the phenomenon within the region and in terms of the generational distribution of the alternation.

**Mots-clés :** Phonologie, Phonétique, Dialectologie, Catalan, Variation, Sociolinguistique

**Keywords:** Phonology, Phonetics, Dialectology, Catalan, Variation, Sociolinguistics

# 1 Introduction<sup>1</sup>

Les enquêtes linguistiques effectuées auprès de 34 informateurs de la *comarca* de La Selva, une région au nord-est de la Catalogne (v. Figure 1), ont permis de découvrir, déjà dans une première écoute, de nombreuses alternances entre [β]<sup>2</sup> – la forme standard – et [w] dans le segment intervocalique des terminaisons de l'imparfait de l'indicatif (ImpInd) de la 1<sup>re</sup> conjugaison verbale (CV I)<sup>3</sup> : nous entendons, par exemple, [əs'ta.βə] autant que [əs'ta.wə] pour la forme verbale *estava* (« j'étais / il était »). Les références à ce phénomène dans la littérature portant sur le Catalan sont très minces (cf. §2) et, pourtant, sa prééminence dans nos enquêtes est remarquable. De plus, lors des écoutes postérieures des enregistrements, nous avons relevé, en sus de [w], deux autres réalisations de l'élément intervocalique : d'un côté, [β],<sup>4</sup> et, de l'autre, l'élision du segment (c'est-à-dire, [əs'ta.βə] et [əs'ta.ə]). Ces réalisations sont moins fréquentes et plus difficiles à percevoir, mais elles sont toutefois intéressantes pour notre recherche.

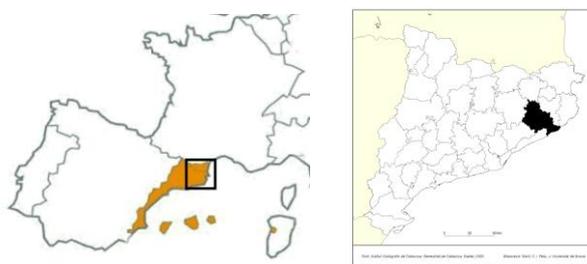


Figure 1 : Cartes de situation de la région de La Selva

Le but principal de cette communication est donc de documenter la variation du segment intervocalique des terminaisons de l'ImpInd CV I dans la région de La Selva pour en évaluer la portée. Outre cela, nous visons à constater que cette variation affecte différemment les locuteurs appartenant aux deux groupes d'âge enquêtés ainsi que les locuteurs provenant de différentes communes dans la région. Concrètement, nos hypothèses sont les suivantes : a) la réalisation [w] est employée majoritairement chez les locuteurs les plus âgés, tandis que [β] et Ø sont plus fréquentes chez les informateurs jeunes ; b) la réalisation [w] n'est réalisée ni dans les communes qui appartiennent au diocèse de Vic ni dans celles qui sont sur la frontière avec le diocèse de Barcelone, et les réalisations [β] et Ø sont également moins fréquentes dans ces secteurs. Finalement, nous tentons une explication de cette variation.

Dans la section 2, nous décrivons les terminaisons de l'ImpInd CV I du catalan et nous faisons un bref état de la question au sujet de la variation géolectale qui nous occupe. La méthodologie de recherche est présentée dans le §3, et les résultats, dans le §4. La section 5 permet de discuter des données et les conclusions sont présentées dans la section 6.

<sup>1</sup> Recherche financée par MINECO et FEDER (bourse pré-doctorale FPI2011, projet FFI2010-22181-C03-02).

<sup>2</sup> Nous utilisons le symbole [β], en accord avec la tradition linguistique catalane, pour représenter la consonne spirante bilabiale voisée, qui dans l'Alphabet Phonétique International (API) est représentée plus précisément par le symbole [β̞]. Ce symbole [β̞], à son tour, est utilisé ici pour représenter une réalisation considérablement plus abaissée – c'est-à-dire, plus ouverte – de la consonne spirante bilabiale voisée.

<sup>3</sup> Nous incluons ici des verbes comme *estar* (« être ») et *anar* (« aller »), qui sont en fait des verbes irréguliers terminés par *-ar* mais qui suivent le paradigme régulier de la CV I pour la formation de l'ImpInd.

<sup>4</sup> V. note 2, *supra*.

## 2 Considérations préliminaires

L'imparfait du latin classique présentait, pour la CV I, les désinences -ABAM, -ABAS, -ABAT, -ABAMUS, -ABATIS et -ABANT. L'évolution de la consonne labiale B dans les différentes conjugaisons est hétérogène dans les diverses langues romanes, mais elle est généralement maintenue – sous la forme phonétique pertinente – dans la CV I en catalan, de même qu'en occitan, en espagnol et en portugais (Pérez, 1998 : 186). Ainsi, l'ImpInd CV I du catalan standard central se conjugue avec les terminaisons -ava, -aves, -ava, -àvem, -àveu et -aven, dont la consonne labiale est réalisée comme une spirante bilabiale voisée, [β].

L'alternance entre [β] et [w] dans la position intervocalique des terminaisons d'ImpInd CV I est signalée par des auteurs comme Gelada (1906 : 93), Alcover (1909 : 243), Julià (1986 : 135, 137-138), Godoy et Comas (1989 : 3), Veny et Pons (1998 : §65), Adam (2006 : 256-257), Bibiloni (2008 : 286) ou Dorca (2010 : 155), qui la situent dans certaines communes du diocèse de Gérone, dont Anglès, Santa Coloma de Farners et Caldes de Malavella, à La Selva. Recasens (1996 : 203) parle des terminaisons en -a[w]a dans les régions du Gironès – qui limite avec La Selva –, le Camp de Tarragona et l'île d'Eivissa, et dans le *DCVB* d'Alcover et Moll (2002) nous trouvons des références à ce phénomène aux Îles Baléares. Cependant, dans tous les cas, l'alternance n'est mentionnée que très vaguement et les données documentées à ce sujet sont virtuellement inexistantes.

L'élision du segment intervocalique, de son côté, est constatée dans les régions méridionales du catalan nord-occidental et du valencien selon Recasens (1996 : 201), Alcover et Moll (dans Perea, 1999), Perea (2008 : 629), et dans le *DCVB*, entre autres, mais on n'a pas de références à ce phénomène dans les variétés du catalan oriental.

## 3 Méthodologie

### 3.1 Points d'enquête

La recherche s'est développée à partir de 34 enquêtes linguistiques menées dans 17 des 26 communes de la région de La Selva (v. Figure 2). Même si cette unité administrative appartient, d'abord, au catalan central (v. Figure 3), elle se trouve dans la région de Gérone, où plusieurs auteurs ont défendu l'existence d'un sous-dialecte distinct (Luna, 1982 : 158 ; Sala, 1983 : 21 ; Julià, 1986 : 130 ; Monturiol et Domínguez, 2001 : 25-26 ; Adam, 2006 : 421-422). Ce facteur est considérable puisque le sous-dialecte géronais coïnciderait avec les limites diocésaines et non administratives (Coromines, 1953 : 209 ; Adam, 2006 : 421-422) : la frontière entre le diocèse de Gérone et celui de Vic croise La Selva, et celle entre les diocèses de Gérone et Barcelone coïncide partiellement avec celle de La Selva (v. Figure 4).

Outre cela, La Selva, contrairement à ce qui se passe dans la majorité des autres *comarques* catalanes, n'est qu'un ensemble essentiellement administratif de communes. Elle est marquée par de nombreux et profonds contrastes concernant la géographie, l'économie, la démographie, les infrastructures et les ressources disponibles (v. Consell Comarcal de la Selva, 2008). Ces différences affectent également la conception de la région comme unité (ou plutôt le manque d'une telle conception) et ont des conséquences évidentes du point de vue de l'évolution linguistique au sein de La Selva (v. Bosch, 2012).

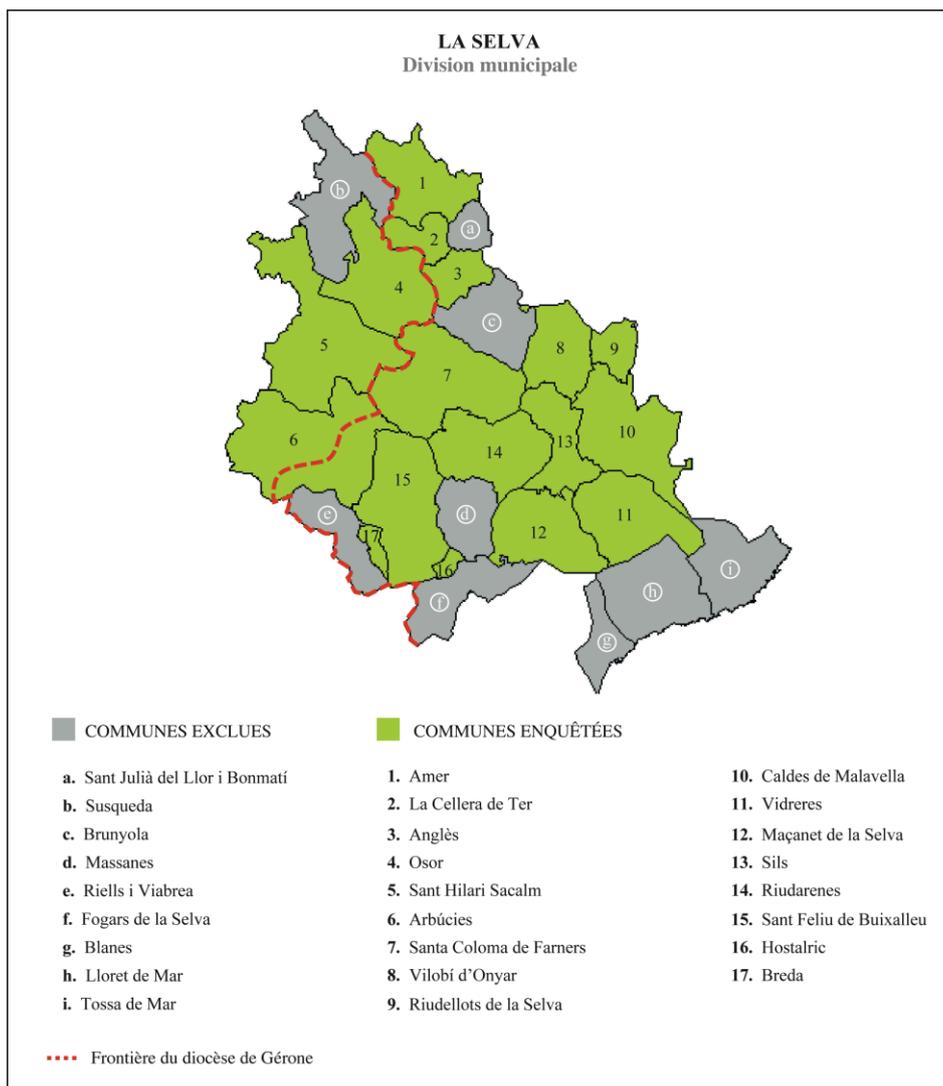


Figure 2 : Division municipale de La Selva



Figure 3 : Division géolectale du Catalan



Figure 4 : Diocèse de Gérone

(Source : [www.encyclopedia.cat](http://www.encyclopedia.cat))

### **3.2 Informateurs**

Les données utilisées pour notre recherche ont été obtenues auprès de 34 informateurs, dont un informateur entre 20 et 30 ans (Groupe 1, désormais G1) et un informateur de 60 ans ou plus (Groupe 2, désormais G2) dans chacune des 17 communes. Ces informateurs ont les caractéristiques suivantes : ce sont des locuteurs catalanophones appartenant à des familles catalanophones ; ils sont nés entre 1982 et 1992, pour le G1, et en 1952 ou avant, pour le G2 ; ils sont nés dans la commune où s'est déroulée l'enquête et ils y ont habité pendant presque toute leur vie ; ce sont les enfants d'au moins un parent né dans la même commune, tandis que l'autre parent est né soit aussi dans la même commune, soit dans la région de La Selva où dans des communes proches ; ils n'ont pas fait d'études universitaires. Parmi les 34 participants, 16 sont des femmes et 18 sont des hommes.

### **3.3 Enquête linguistique**

Les enquêtes linguistiques qui ont permis l'obtention des données ont été conduites entre avril et juillet de 2012. Chaque enquête, d'une durée d'entre 40 et 90 minutes, a été enregistrée à l'audio et sur vidéo.

L'enquête était structurée en deux parties : la première était une entrevue semi-dirigée, à travers laquelle nous avons essayé d'obtenir des fragments de discours le plus spontané possible ; la deuxième était un questionnaire visant à recueillir des segments phoniques précis. Cette recherche porte sur des données obtenues lors de l'entrevue semi-dirigée.

### **3.4 Extraction et traitement des données**

Les enregistrements ont été analysés afin d'extraire toutes les occurrences de l'ImpInd CV I produites par les informateurs lors des enquêtes. Nous avons segmenté les fragments de vidéo pertinents et chaque forme verbale a été transcrite en API au moyen du programme Phon (<http://childes.psy.cmu.edu/phon/>, version 1.5.2). Ensuite, nous avons fait le décompte des occurrences totales proférées par chacun des informateurs, ainsi que le décompte des occurrences de chaque réalisation de l'élément intervocalique de la terminaison de l'ImpInd CV I ([β], [w], [β̞] ou Ø), et nous en avons calculé les pourcentages.

À ce sujet, il faut remarquer l'importance et les conséquences du fait que les verbes analysés ici sont extraits des fragments de discours plutôt spontané. Cette caractéristique entraîne, d'un côté, que les imparfaits enregistrés auprès de chaque informateur diffèrent, quant au nombre d'occurrences autant qu'aux items lexicaux proférés, et, donc, que l'analyse comparative des données est nécessairement limitée. D'un autre côté, pourtant, en accord avec ce que nous avons observé lors de l'étude des données obtenues à partir du questionnaire de la deuxième partie des enquêtes (v. Bosch, 2012), la variation trouvée dans les terminaisons de l'ImpInd CV I aurait été sans doute plus difficile à obtenir au moyen d'un questionnaire, en raison du caractère plus formel du contexte.

## **4 Résultats**

Les enregistrements ont permis d'extraire un total de 931 occurrences d'ImpInd CV I, dont 73,58 % des terminaisons sont prononcées avec [β] en position intervocalique, 17,83 % avec

[w], 5,69 % avec [β] et 2,90 % sont produites avec une élision du segment intervocalique (Ø). Toutefois, comme on le prévoyait (§3.4), ces chiffres décrivent à peine la situation globale. Afin de mieux illustrer la variation nous présentons dans le Tableau 1 les résultats individuels.

COMMUNE	GROUPE	NOMBRE DE REALISATIONS				Total
		[β]	[β̃]	[w]	Ø	
1. AMER	G1	10 (71,43 %)	4 (28,57 %)	0 (0 %)	0 (0 %)	14 (100 %)
	G2	39 (81,25 %)	4 (8,33 %)	5 (10,42 %)	0 (0 %)	48 (100 %)
2. LA CELLERA DE TER	G1	15 (60 %)	9 (36 %)	0 (0 %)	1 (4 %)	25 (100 %)
	G2	12 (16,67 %)	3 (4,17 %)	57 (79,17 %)	0 (0 %)	72 (100 %)
3. ANGLES	G1	2 (100 %)	0 (0 %)	0 (0 %)	0 (0 %)	2 (100 %)
	G2	11 (25 %)	3 (6,82 %)	28 (63,64 %)	2 (4,55 %)	44 (100 %)
4. OSOR	G1	8 (66,67 %)	3 (25 %)	0 (0 %)	1 (8,33 %)	12 (100 %)
	G2	161 (96,41 %)	4 (2,40 %)	2 (1,20 %)	0 (0 %)	167 (100 %)
5. SANT HILARI SACALM	G1	8 (53,33 %)	5 (33,33 %)	0 (0 %)	2 (13,33 %)	15 (100 %)
	G2	41 (100 %)	0 (0 %)	0 (0 %)	0 (0 %)	41 (100 %)
6. ARBUCIES	G1	5 (100 %)	0 (0 %)	0 (0 %)	0 (0 %)	5 (100 %)
	G2	54 (100 %)	0 (0 %)	0 (0 %)	0 (0 %)	54 (100 %)
7. SANTA COLOMA DE FARNERS	G1	13 (59,09 %)	1 (4,55 %)	0 (0 %)	8 (36,36 %)	22 (100 %)
	G2	16 (94,12 %)	0 (0 %)	1 (5,88 %)	0 (0 %)	17 (100 %)
8. VILOBI D'ONYAR	G1	12 (54,55 %)	2 (9,09 %)	3 (13,64 %)	5 (22,73 %)	22 (100 %)
	G2	15 (68,18 %)	0 (0 %)	7 (31,82 %)	0 (0 %)	22 (100 %)
9. RIUDELLOTS DE LA SELVA	G1	11 (47,83 %)	1 (4,35 %)	11 (47,83 %)	0 (0 %)	23 (100 %)
	G2	49 (98 %)	1 (2 %)	0 (0 %)	0 (0 %)	50 (100 %)
10. CALDES DE MALAVELLA	G1	6 (75 %)	2 (25 %)	0 (0 %)	0 (0 %)	8 (100 %)
	G2	9 (40,91 %)	0 (0 %)	13 (59,09 %)	0 (0 %)	22 (100 %)
11. VIDRERES	G1	6 (50 %)	1 (8,33 %)	1 (8,33 %)	4 (33,33 %)	12 (100 %)
	G2	25 (78,13 %)	0 (0 %)	7 (21,88 %)	0 (0 %)	32 (100 %)
12. MAÇANET DE LA SELVA	G1	10 (100 %)	0 (0 %)	0 (0 %)	0 (0 %)	10 (100 %)
	G2	1 (16,67 %)	0 (0 %)	5 (83,33 %)	0 (0 %)	6 (100 %)
13. SILS	G1	6 (37,50 %)	6 (37,50 %)	4 (25 %)	0 (0 %)	16 (100 %)
	G2	15 (93,75 %)	0 (0 %)	0 (0 %)	1 (6,25 %)	16 (100 %)
14. RIUDARENES	G1	18 (85,71 %)	1 (4,76 %)	0 (0 %)	2 (9,52 %)	21 (100 %)
	G2	2 (8,33 %)	0 (0 %)	22 (91,67 %)	0 (0 %)	24 (100 %)
15. SANT FELIU DE BUIXALLEU	G1	10 (76,92 %)	2 (15,38 %)	0 (0 %)	1 (7,69 %)	13 (100 %)
	G2	25 (96,15 %)	1 (3,85 %)	0 (0 %)	0 (0 %)	26 (100 %)
16. HOSTALRIC	G1	31 (100 %)	0 (0 %)	0 (0 %)	0 (0 %)	31 (100 %)
	G2	26 (100 %)	0 (0 %)	0 (0 %)	0 (0 %)	26 (100 %)
17. BREDÀ	G1	9 (100 %)	0 (0 %)	0 (0 %)	0 (0 %)	9 (100 %)
	G2	4 (100 %)	0 (0 %)	0 (0 %)	0 (0 %)	4 (100 %)

Tableau 1 : Nombre d'occurrences et pourcentage d'utilisation de chacun des 4 segments possibles dans la terminaison de l'ImpInd CV I. Les couleurs indiquent l'intensité d'utilisation de chaque variable :



Le tableau<sup>5</sup> montre une variation du segment intervocalique des terminaisons de l'ImpInd CV I très élevée. Nous y observons, ainsi, qu'il n'y a que 9 participants qui utilisent la forme avec [β] pour 100 % de leurs occurrences de l'ImpInd CV I. Cependant, ce segment est préféré (≥50 % des occurrences) par 27 des 34 informateurs. Quant à la réalisation [w], elle est utilisée au moins dans l'une des occurrences enregistrées auprès de 14 de nos participants (41,18 %), mais il s'agit de la solution de préférence seulement pour 5 d'eux et le pourcentage d'utilisation maximum est de 91,67 % (G2, Riudarenes). Cette réalisation n'est utilisée que par 4 locuteurs du G1 et, du point de vue géographique, nous ne la trouvons pas dans les communes d'Arbúcies et Sant Hilari Sacalm, appartenant au diocèse de Vic<sup>6</sup>, ni dans celles de Breda, Hostalric et Sant Feliu de Buixalleu, près de la limite du diocèse et de la province de Barcelone. Les variables [β] et Ø sont utilisées par 18 et 10 des informateurs, respectivement, mais toujours de façon marginale, donc le pourcentage d'utilisation maximum ne dépasse pas 37,50 % des occurrences pour le premier, et 36,36 % pour le deuxième. Ces segments sont aussi utilisés principalement par des informateurs du G1, et ils se trouvent spécialement dans des communes où nous remarquons des occurrences avec [w] auprès du G2. Toutefois, il y a aussi des occurrences avec [β] et/ou Ø dans des communes où [w] n'est pas utilisé.

## 5 Discussion

Les données présentées dans le Tableau 1 confirment, au moins de façon partielle, nos hypothèses de travail. Premièrement, la variation de l'élément intervocalique des terminaisons de l'ImpInd CV I existe et elle a une incidence considérable, puisque nous y observons 4 variables différentes ([β], [w], [β], Ø), dont 3 qui ne coïncident pas avec la forme standard ([w], [β], Ø) sont utilisées dans 26,42 % des occurrences registrées. Ce qui est plus remarquable c'est que 25 de nos 34 informateurs, soit 73,53 %, utilisent 2 ou plus de ces variables.

Quant à la variation générationnelle, les données montrent que 88,55 % des occurrences de [w] sont proférées par des informateurs du G2, contre 11,45 % d'utilisation des informateurs du G1. L'abandon de cette forme de la part des informateurs jeunes corrobore l'idée généralisée que le catalan a actuellement une forte tendance au nivellement dialectal en faveur des formes préconisées pour la variété standard (Perea, 2007 : 87-88 ; Viaplana, 1999 : 115). Cette propension, qui évidemment n'est pas une particularité exclusive de la langue catalane, est néanmoins renforcée par des facteurs spécifiques, tels que le processus de normalisation linguistique et l'introduction du catalan dans le système éducatif et les médias, phénomènes relativement récents (Viaplana, 2002 : 222). Toutefois, il faut considérer encore deux questions. D'un côté, la variable [w] constatée auprès de 4 locuteurs du G1, dont 2

---

<sup>5</sup> Les occurrences enregistrées correspondent à 142 items lexicaux distincts. Les verbes *anar* (« aller », 216 occurrences), *estar* (« être », 89 occ.), *treballar* (« travailler », 41 occ.) et *passar* (« passer », 32 occ.) sont les plus fréquents, suivis de 138 autres verbes employés entre 1 et 29 fois chacun. Mais ni ces chiffres ni les informations relatives à chaque item sont très informatives (v. §3.4) : nous observons, par exemple, que 17 des 26 occurrences du verbe *posar* (« mettre ») ont été proférées par l'informateur du G2 d'Osor, ou qu'entre les verbes avec une seule occurrence il y a des exemples de toutes les variables. Il faut souligner, toutefois, que nous n'avons enregistré aucune occurrence de 2<sup>e</sup> personne du pluriel.

<sup>6</sup> La commune d'Arbúcies est en fait divisée entre les diocèses de Gérone et Vic, mais les données obtenues lors de nos enquêtes, tout comme celles qu'on trouve dans la littérature portant sur le catalan, montrent que le comportement linguistique municipale tend à ne pas suivre les caractéristiques considérées *géronaises*, probablement à conséquence de facteurs qui favorisent une diffusion spatiale (*spatial diffusion*) par accommodation des traits linguistiques (v. Trudgill, 2003 : *accommodation* ; Britain, 2008, §6.2).

appartiennent à des communes où l'informateur du G2 utilise également cette forme (Vidreres et Vilobí d'Onyar), mais les 2 autres l'emploient bien que leurs concitoyens âgés ne le font pas (Riudellots de la Selva et Sils) – et celui de Riudellots de la Selva se sert de [w] aussi fréquemment que de [β] (47,83 %). Et de l'autre côté, l'apparition dans le même contexte de [β] et Ø, utilisées majoritairement auprès du G1 (69,81 % et 88,89 % des occurrences, respectivement).

Conformément à ce que prévoiaient nos hypothèses, il y a une relation entre la distribution géographique du phénomène et les frontières diocésaines ainsi que les caractéristiques socioéconomiques de la région. Ainsi, les réalisations de l'ImpInd CV I avec [w] surviennent toujours dans le diocèse de Gérone, à l'exception de 2 occurrences (sur 167) auprès de l'informateur du G2 d'Osor. Cette commune appartenant au diocèse de Vic a un caractère particulier, lié à son relatif isolement géographique, et l'utilisation de certaines formes d'abord *géronaises* n'y est pas surprenant (v. Dorca, 2010). De façon similaire, le fait que nous ne trouvons pas cette variable auprès des informateurs de Breda, Hostalric et Sant Feliu de Buixalleu, qui appartiennent au diocèse de Gérone, est cohérent avec les caractéristiques socio-économiques de ce secteur et avec l'analyse présentée dans Bosch (2012 : §6.2) : la vitalité économique de ces communes, ainsi que leur proximité et facilité de communications avec Barcelone, a entraîné une grande mobilité de leurs habitants, ce qui favorise l'abandon des formes traditionnelles.<sup>7</sup> En ce qui concerne [β] et Ø, nous observons ces réalisations surtout dans les villages où il y a des occurrences de [w]. Cependant, le fait que nous trouvons des exemples de ces deux variables dans des villages sans formes avec [w] enregistrées pose des questions additionnelles.

Pour expliquer l'alternance entre [β] et [w] intervocaliques dans les terminaisons d'ImpInd CV I, nous nous référons à Adam (2006 : 257-259), qui offre deux réponses possibles. La première, morphologique, considère que l'alternance serait une conséquence analogique des formes adjectivales féminines telles que *bla[w]a* au lieu de *bla[β]a* (*blava*, « bleue »), qui à son tour se produirait par analogie à la forme *me[w]a* du possessif féminin *me[β]a* (*meva*, « ma »),<sup>8</sup> analogique du masculin *me[w]* (*meu*, « mon »).<sup>9</sup> Cette concaténation d'analogies serait motivée par la basse fréquence, en catalan, de mots paroxytons terminés par [βə]. La deuxième possibilité est phonético-phonologique, et répondrait à une tendance à produire le segment /b/ intervocalique comme [w]. À ce propos, Recasens (1996 : 203) affirme que le passage de consonne bilabiale à [w] se produirait soit du fait d'une tendance à l'affaiblissement articulaire du son originel ([β]), soit par confusion perceptuelle, en raison de la ressemblance entre les fréquences de F2 de [β] et [w].

De notre côté, eu égard aux données, la solution phonétique est la plus plausible. Bien que l'alternance entre [β] et [w] puisse être également expliquée tant du point de vue morphologique que du point de vue phonologique – lors de nos enquêtes nous avons

<sup>7</sup> D'après Milroy (2008 : 566), « if a close-knit community network loosens and members become mobile, the social prerequisites for supporting highly localized norms disappear, and dialect leveling takes place ».

<sup>8</sup> Aussi : *te[w]a* au lieu de *te[β]a* (*teva*, « ta ») et *se[w]a* au lieu de *se[β]a* (*seva*, « sa »).

<sup>9</sup> Nous trouvons des possessifs féminins avec [w] (*me[w]a* / *te[w]a* / *se[w]a* au lieu de *meva* / *teva* / *seva*), dans la plupart des aires catalanophones – pourtant, cette forme n'est pas la plus habituelle dans le catalan central. Par contre, les réalisations avec [w] des formes féminines de certains adjectifs oxytons masculins terminés par [w] (par exemple, *bla[w]a*, *vi[w]a* et *to[w]a* au lieu de *bla[β]a* – *blava*, « bleue » –, *vi[β]a* – *viva*, « vivante » – et *to[β]a* – *tova*, « molle » –, féminins de *blau*, *viu* et *tou*) n'ont été localisées que dans des régions appartenant au diocèse de Gérone (v. Adam, 2006 : *mapa MI-8*, 159 et *mapa MI-15*, 265).

enregistré des réalisations avec [w] de *blava* et *meva*, mais aussi de substantifs comme *cove* (« panier ») –, il faut tenir compte des deux autres variables, [β] et l'élision du segment. Ainsi, si nous considérons que le passage de [β] à [w] est entièrement un phénomène de caractère phonique, nous pouvons accepter que la réalisation [β] soit un pas intermédiaire entre les deux. Ceci pourrait expliquer son utilisation majoritairement de la part du G1, dont les locuteurs *tendent* à substituer les formes de la variété standard aux formes traditionnelles (v. Bosch, 2012). Quant à l'élision du segment, cette variable serait une évolution de la réalisation [β], qui est à peine perceptible et, donc, est spécialement susceptible à l'élision. Cette explication fondée sur la phonétique permettrait aussi de justifier la variation notable enregistrée auprès de chaque informateur. Il s'agit pourtant d'une solution qui n'explique pas forcément l'apparition des variables [β] et Ø dans des communes où [w] n'est pas employée.

## 6 Conclusions

Lors de cette recherche, nous avons constaté que l'élément intervocalique de la terminaison de l'imparfait de l'indicatif de la première conjugaison présente, dans la plupart des communes de la *comarca* de La Selva, une forte variation : nous en avons enregistré quatre formes distinctes ([β], [w], [β̥], Ø) et près de 3 sur 4 de nos informateurs utilisent au moins deux de ces variables. Tous les informateurs emploient la forme standard ([β]). La variable [w] est enregistrée pour la plupart dans le G2 et seulement dans des communes du diocèse de Gérone, tandis que [β̥] et Ø sont quantitativement marginaux mais plus fréquents auprès des informateurs du G1. Nous avons considéré que ce phénomène a une motivation phonétique.

## Remerciements

Je remercie Maria-Rosa Lloret pour ses commentaires et annotations, Montserrat Adam pour ses remarques et Annabelle Caillou pour ses corrections grammaticales.

## Références

- [=DCVB] ALCOVER A. M., MOLL F. de B. (2002). *Diccionari català-valencià-balear*. Barcelone: Institut d'Estudis Catalans. <<http://dcvb.iecat.net>> [05-01-2013]
- ADAM I AULINAS M. (2006). *El Català Septentrional De Transició: Nova Visió Des De La Morfologia*. Biblioteca filològica. Barcelone: Institut d'Estudis Catalans.
- ALCOVER A. M. (1909). Una mica de dialectologia. BDLC, vol. IV (1908-1909), 194-303.
- BIBILONI G. (2008). Elisió de *-n* i *-r*, distribució de les ròtiques i altres fenòmens consonàntics en el mot. Dans SOLA J., LLORET M.R., MASCARO J., PEREZ SILDANYA M. (dir.), *Gramàtica del català contemporani*, Chapitre 7, Phonétique et Phonologie (4<sup>e</sup> éd. définitive), vol. 1, 271-285. Barcelone: Empúries.
- BOSCH I ROURA E. (2012). *Aspectes fonològics del català de la comarca de la Selva*. Thèse de maîtrise. Universitat de Barcelona.

BRITAIN D. (2008). Space and Spatial Diffusion. Dans CHAMBERS J. K., TRUDGILL P. et SCHILLING-ESTES N. (éd.), *The Handbook of Language Variation and Change*, 603-637. Oxford: Blackwell.

CONSELL COMARCAL DE LA SELVA (2008). La Selva. Agenda 21. Memòria descriptiva de la situació actual de la comarca. <[http://www.selva.cat/agenda21/pla\\_comarcal.php](http://www.selva.cat/agenda21/pla_comarcal.php)> [02-01-2013]

COROMINES J. (1953). Algunes lleis fonètiques catalanes no observades fins ara. *Estudis Romànics* 3, 201-230. <<http://revistes.iec.cat/index.php/ER/article/viewFile/37651/48276>> [19-12-2012]

DORCA J. (2010). Aproximació geolingüística als parlars de les Guillerries: notes de presentació. *Quaderns de la Selva* 22, 147-160. <<http://www.raco.cat/index.php/QuadernsSelva/article/view/225276/334739>> [20-12-2012]

GELADA G. (1906). Particularitats idiomàtiques d'Olot. BDLC, vol. III (1906-1907), 91-94.

GODOY I TOMÀS L., COMAS I PUJOL M. T. (1989). La parla de Tapis. *El Pedrís (Setmanari de l'Alt Empordà)* 26, 3-5.

JULIA I MUNE J. (1986). Algunes notes sobre el parlar de Cornellà de Terri. Dans CALLIS M (éd.), *Cornellà de Terri: pregons i estudis*, 129-139. Cornellà de Terri: Ajuntament.

LUNA X. (1982). *El vocalisme del català a la zona de transició central-septentrional (Alt Empordà)*. Thèse de licence, Universitat Autònoma de Barcelona.

MILROY L. (2008). Social Networks. Dans CHAMBERS J. K., TRUDGILL P. et SCHILLING-ESTES N. (éd.), *The Handbook of Language Variation and Change*, 549-572. Oxford: Blackwell.

MONTURIOL J., DOMINGUEZ E. (2001). *El Parlar De La Garrotxa*. Olot: Ràdio Olot.

PEREA M. P. (1999). *Compleció i ordenació de «La flexió verbal en els dialectes catalans» d'A. M. Alcover i F. de B. Moll*. Barcelone: Institut d'Estudis Catalans.

PEREA M. P. (2007). Dedialectalization or the Death of a Dialect: The Case of the Catalan Subdialect Spoken in the Costa Brava. *Dialectologia et Geolinguística* 15, 77-89. <doi:10.1515/DIALECT.2007.006>

PEREA M.P. (2008). Flexió verbal regular. Dans SOLÀ J., LLORET M.R., MASCARÓ J., PÉREZ SALDANYA M. (dir.), *Gramàtica del català contemporani*, Chapitre 4, Morfologia (4<sup>e</sup> éd. définitive), vol. 1, 583-646. Barcelone: Empúries.

PÉREZ SALDANYA M. (1998). *Del llatí al català: morfosintaxi verbal històrica*. Valence: Universitat de València.

RECASENS D. (1996). *Fonètica descriptiva del català: assaig de caracterització de la pronúncia del vocalisme i consonantisme del català al segle XX* (2<sup>e</sup> éd.). Barcelone: Institut d'Estudis Catalans.

SALA E. (1983). *El Parlar de Cadaqués*. Gérone: Diputació de Girona.

*L'alternance entre [β] et [w] dans les terminaisons de l'imparfait de l'indicatif en Catalan dans la région de La Selva*

TRUDGILL P. (2003). *A Glossary of Sociolinguistics*. Oxford University Press.

VENY J., PONS I GRIERA L. (1998). *Atles lingüístic del domini català. Enotextos del català oriental*. Barcelone : Institut d'Estudis Catalans.

VIAPLANA J. (1999). *Entre la dialectologia i la lingüística: la distància lingüística entre les varietats del català nord-occidental*. Barcelone: Publicacions de l'Abadia de Montserrat.

VIAPLANA J. (2002). *Dialectologia* (2a éd., corr.). Valence: Universitat de València.

## **La variation entre l’hiatus et les diphtongues croissantes en portugais brésilien**

Vanessa Meireles<sup>1</sup>

(1) SFL & Université de Paris 8  
vanessameireless@gmail.com

**Résumé** Le portugais, comme d’autres langues, présente une tendance à éviter l’hiatus, c’est-à-dire à rejeter une séquence de deux voyelles appartenant à des syllabes différentes. Nous décrivons un mécanisme mis en jeu dans la résolution des séquences vocaliques en hiatus en portugais brésilien : la diphtongaison. On examinera la variation entre l’hiatus et les diphtongues croissantes et sa conséquence pour la représentation phonologique des diphtongues croissantes en portugais dans le cadre théorique du CVCV.

**Abstract** Portuguese, like other languages, exhibits a tendency to avoid hiatus, i.e., to reject a sequence of two vowels belonging to different syllables. We will describe one possibility to solve sequences in hiatus in Brazilian Portuguese, namely diphthongization. We will examine the variation between hiatus and rising diphthongs and its consequence for the phonological representation of this kind of diphthongs in Portuguese within CVCV framework.

**Mots-clés :** Hiatus, diphtongues, portugais, syllabe, théorie CVCV.

**Keywords:** Hiatus, diphthongs, Portuguese, syllable, CVCV theory.

## 1 Introduction

Dans les grammaires traditionnelles du portugais, les rencontres vocaliques sont normalement classées en trois types : il existe des hiatus, des diphtongues et des triphthongues. Selon la façon dont les voyelles du portugais peuvent se combiner, deux types de rencontres vocaliques sont à distinguer : celui des diphtongues décroissantes, et celui des hiatus, en alternance avec des diphtongues croissantes. À partir des faits de variation, on discutera le statut phonologique de la diphtongue (croissante) en portugais. Nous verrons différents traitements qui en ont été proposés et le principal point de divergence entre eux, qui touche à la structure de la syllabe. Cette divergence concerne la position syllabique qu'occupent les semi-voyelles ou glides en portugais (attaque ou première partie d'un noyau complexe dans les cas des diphtongues croissantes ; deuxième partie d'un noyau complexe ou coda dans les diphtongues décroissantes). Nous proposerons enfin une description des groupes vocaliques du portugais (les hiatus et les diphtongues) dans le cadre de la phonologie CVCV ou « CV strict », qui rend compte des problèmes ci-dessus.

## 2 Les hiatus et les diphtongues en portugais

Selon Cunha & Cintra (1985), un hiatus est la rencontre de deux voyelles syllabiques (*pais* [pa'ĩ]). Même si plusieurs séquences en hiatus ont été défaites diachroniquement, l'observation de l'état actuel de la langue montre que quelques hiatus ont subsisté. Selon Rodrigues (2007 : 11), ces hiatus se maintiennent « seulement en contexte tonique, si la voyelle accentuée est la deuxième et qu'elle est une voyelle haute, comme par exemple dans 'saúde', 'baú' et 'saí', entre autres ». Mis à part ce contexte où l'hiatus est stable, les autres rencontres vocaliques en hiatus auront tendance à être défaites, moyennant plusieurs mécanismes, parmi lesquels la diphtongaison nous intéresse particulièrement ici.

Cunha & Cintra (1985) définissent la diphtongue comme la rencontre d'une voyelle avec une semi-voyelle, ou d'une semi-voyelle et une voyelle. Selon l'ordre et la nature des éléments qui les constituent, les diphtongues peuvent être décroissantes (quand la voyelle vient en première place) ou croissantes (quand la semi-voyelle précède la voyelle), plus ou moins stables, orales ou nasales. La classification des diphtongues en deux catégories, « décroissantes » et « croissantes », selon l'ordre d'apparition du glide par rapport à la voyelle, suscite cependant beaucoup de controverses. Plusieurs auteurs affirment que la langue portugaise n'a pas de diphtongue croissante phonologique, seulement des diphtongues décroissantes.

Telle est la thèse défendue par Câmara Jr. (1970), pour qui les « vraies » diphtongues du portugais sont les diphtongues décroissantes ; les diphtongues croissantes seraient en variation libre avec l'hiatus, et n'auraient donc pas de statut phonologique en portugais. Les seuls cas, dans le parler courant, de diphtongues croissantes stables sont ceux où [w] est précédée d'une consonne vélaire [k] ou [g] : *quase*, *igual*, *equestre*, *tranquilo*, etc. Il est alors possible d'identifier des paires minimales : par exemple, *quando* / *coando* ([kwã] vs [kuã]).

Outre les diphtongues listées auparavant, d'autres diphtongues peuvent surgir en certains contextes. Il s'agirait d'un processus strictement phonétique, déclenché par des facteurs euphoniques, liés au contexte phonétique précédent ou suivant, à la vitesse de prononciation et au caractère formel ou informel du registre. Par exemple, l'hiatus entre une voyelle haute devant une autre voyelle haute ou devant une voyelle basse peut être défait par la formation d'une diphtongue croissante (*ruim* : ru-'im > 'rwim ; *luar* : lu-'ar > 'lwar ; *piada* : pi'ada >

'pjada, où les voyelles hautes [i] et [u] non accentuées alternent avec les glides correspondants [j] et [w]).

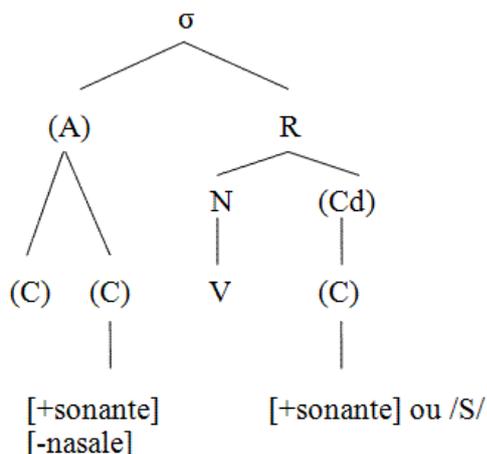
Câmara Jr. (1970) en est venu à penser que la description des diphtongues portugaises fait partie de la description des types de syllabe en portugais. Le tout est de déterminer, pour une syllabe qui contient une diphtongue décroissante, comme dans *pai*, s'il s'agit d'une syllabe ouverte avec deux éléments vocaliques (CVV) ou d'une syllabe fermée avec une coda (CVC). De la même façon, dans le cas des diphtongues croissantes, il faut déterminer s'il s'agit d'une syllabe CCV ou CVV, c'est-à-dire si les glides sont des consonnes ou des allophones des voyelles.

Les glides se comportent en partie comme des consonnes, mais ont pourtant certaines caractéristiques phonétiques des voyelles [i] et [u]. Ce caractère hybride des glides a donné lieu à beaucoup de controverses sur l'interprétation de leur statut au sein de la syllabe. Il y a deux possibilités pour la position du glide prévocalique dans les diphtongues croissantes : soit le glide occupe l'attaque, soit il partage le noyau avec la voyelle qu'il précède. Deux interprétations sont également possibles pour les glides postvocaliques dans les diphtongues décroissantes : soit ils occupent le noyau, soit ils occupent la coda d'une syllabe fermée. Nous présenterons les interprétations des glides de Câmara Jr. (1970) dans le cadre de la phonologie structuraliste, de Bisol (1999) et de Mateus & D'Andrade (2000) dans le cadre de la phonologie multilinéaire, pour enfin proposer à notre tour une analyse des glides du portugais à la lumière de la phonologie du CVCV ou « CV strict » (Lowenstamm 1996, Scheer 2004).

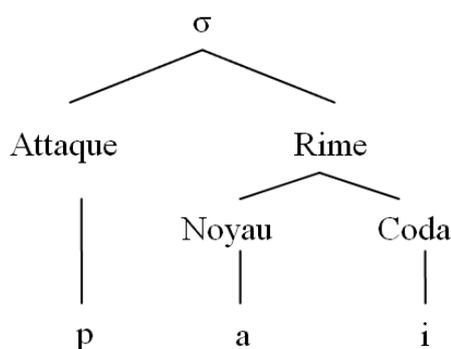
Câmara Jr. (1970) pose que la diphtongue, étant le résultat de la modification d'une voyelle par une autre voyelle dans la même syllabe, est composée de deux segments vocaliques, la semi-voyelle faisant partie du noyau, où [j] et [w] sont des allophones assyllabiques des voyelles hautes /i/ et /u/. L'argument à l'appui de cette position est que l'on trouve un [r] après une diphtongue (*Laura, eira, Europeu*), car cette consonne n'existe en portugais, en attaque, qu'entre deux voyelles : *era* ['ɛ.ra] / *erra* ['ɛ.ha] ; après une syllabe fermée, il ne peut y avoir que [h] : *Israel* [iʒ.ha.'ɛw]. D'autres arguments sont la possibilité de monophthongaison ([ou] > [o], par exemple) et la syllabation variable lorsque la diphtongue se trouve en position non accentuée (*vaidade* : *vai-'dade* ou *va-i-'da-de*), ou encore la facilité avec laquelle on passe, sous certaines conditions, d'un segment assyllabique à une voyelle moyenne (*papai* : *papa*[j] ~ *papa*[e]). Câmara Jr. ne discute pas le statut des glides prévocaliques dans les diphtongues croissantes, car pour cet auteur les seuls cas de diphtongue croissante phonologique sont, comme il a déjà été dit, les séquences formées avec les consonnes vélaires /k/ et /g/.

Bisol (1999) affirme que la variation entre hiatus et diphtongues mène à croire que le système sous-jacent n'a pas de diphtongues croissantes, mais seulement des hiatus qui, en surface, alternent avec les diphtongues. Selon Bisol, les diphtongues décroissantes sont définies dès le niveau lexical, alors que les diphtongues croissantes se formeraient au niveau post-lexical.

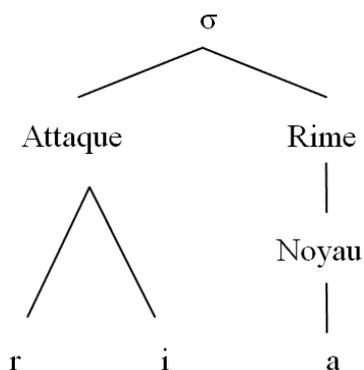
Dans les diphtongues décroissantes, Bisol soutient que les segments [j] et [w] font partie de la coda, en avançant deux arguments. Le premier est le fait qu'en portugais il n'y a pas de séquence voyelle + glide + consonne sonante : le glide ne peut pas coexister dans la même syllabe avec les consonnes -n, -l, -r car il occuperait la même position qu'elles dans la coda (d'où \**boyl*, \**sayr*). L'autre argument est l'inexistence de voyelles longues en portugais, qui occuperaient deux positions dans le noyau. Selon Bisol, en portugais le patron syllabique est CCVC(C) :



Bisol affirme que, dans le système sous-jacent, il n'y a pas de glides. Dans la représentation sous-jacente, il y aurait deux voyelles, dont une haute, avant la formation de la diphtongue décroissante dans le lexique. Lorsque le processus de syllabation a lieu, la voyelle de plus grande sonorité est placée dans le noyau et la voyelle haute est placée dans la coda. Une voyelle en position de coda se convertit automatiquement en glide d'après Bisol.



Dans les diphtongues croissantes, il y aurait de la même façon deux voyelles à la sous-jacence, dont une haute qui se maintient comme telle jusqu'au niveau post-lexical. Au niveau post-lexical, la voyelle haute peut se transformer en glide occupant l'attaque de la syllabe : ri.a.cho → rja.cho. La diphtongue croissante serait ainsi le résultat d'une resyllabation et se trouve en variation libre avec l'hiatus.

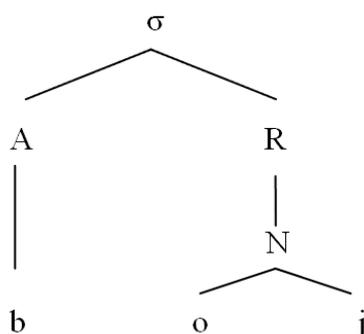


Selon Bisol, l'attaque en portugais serait binaire au niveau lexical. Si le glide occupe l'attaque dans des mots comme *criado*, *criança*, on a une attaque ternaire, mais comme la formation des diphtongues croissantes aurait lieu au niveau post-lexical, nulle contrainte n'est violée selon Bisol.

Il y a cependant des diphtongues croissantes qui n'alternent pas avec l'hiatus, les seuls cas de diphtongues croissantes stables comme il a déjà été dit. Il s'agit des diphtongues précédées d'une consonne vélaire [k] ou [g] et comprenant le glide [w] suivi de /a/ ou /o/. Bisol présente deux possibilités d'analyse pour ce cas spécifique. La première possibilité est que la diphtongue soit déjà dans la représentation sous-jacente, vu le nombre réduit de ces cas de figure. Une autre possibilité est que [kw] représente une consonne complexe /k<sup>w</sup>/, avec une première articulation dorsale et une articulation secondaire labiale. Au niveau post-lexical le trait labial se convertirait en glide, générant la diphtongue croissante.

Mateus & D'Andrade (2000) se basent sur le portugais européen (parler de Lisbonne), même si le portugais brésilien est également évoqué. Les auteurs postulent que tous les glides en surface sont issus de voyelles hautes sous-jacentes, ceux des diphtongues décroissantes comme ceux des diphtongues croissantes. L'argument majeur est le fait qu'il n'existe pas d'opposition phonologique entre une voyelle et un glide en portugais, c'est-à-dire, entre [i], [u] et [j], [w], comme par exemple \*[<sup>h</sup>paɨ] vs [<sup>h</sup>paɨ], où la commutation de la voyelle avec le glide changerait le sens<sup>1</sup>.

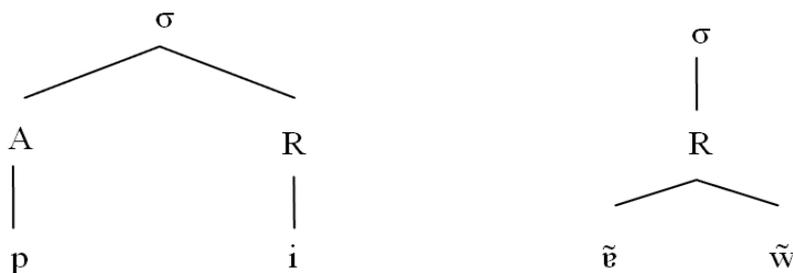
Dans le cas des diphtongues décroissantes, une voyelle haute deviendrait glide car elle recevrait une marque dans sa représentation lexicale qui indique l'impossibilité de porter l'accent si elle est précédée d'une voyelle. Les auteurs postulent que le glide des diphtongues décroissantes appartient au noyau syllabique (Mateus & D'Andrade, 2000, p. 46) :



Une raison en est, selon les auteurs, que, dans les diphtongues nasales, les voyelles et les glides sont nasalisés comme un tout (*mãõ* [<sup>h</sup>mãõ̃]).

Les glides prévocaliques sont également dérivés d'un hiatus au niveau sous-jacent :

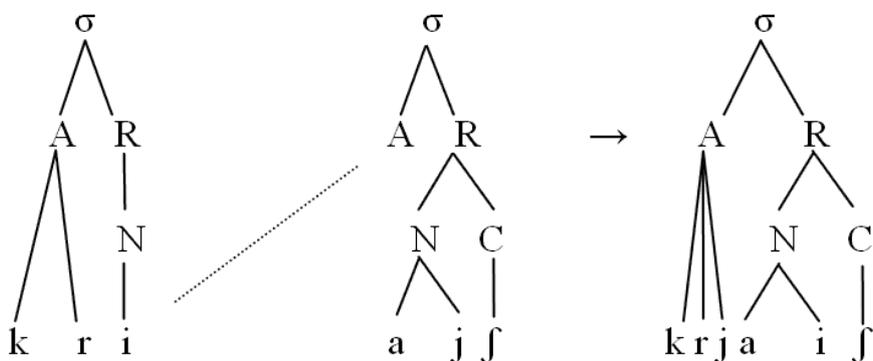
<sup>1</sup> *pais* (« parents ») et *páis* (« pays ») ne constituent pas une paire minimale, puisque l'accent ne tombe pas sur la même syllabe (*pais* [<sup>h</sup>paɨ̃] ; *páis* [pa'ɨ̃]).



À la surface, ces hiatus peuvent être interprétés comme des diphtongues croissantes. Mais les glides des diphtongues croissantes, contrairement à ceux des diphtongues décroissantes, ne se nasalisent pas : *pião* ['pjãĩ]. Ceci indiquerait qu'ils appartiennent à l'attaque de la syllabe. S'ils faisaient partie de la rime, ils seraient également nasalisés comme il arrive aux glides des diphtongues décroissantes.

Mateus & D'Andrade évoquent l'analyse de Harris (1983) pour qui les glides prévocaliques en espagnol occuperaient la rime syllabique. Selon Harris, en espagnol la rime supporterait jusqu'à trois segments comprenant le noyau et la coda (*buei* : *bwej*). Dans des mots comme *limpiais*, le /s/ est extramétrique car il s'agit d'un morphème. En revanche, en portugais on trouve des exemples où un /s/ en coda n'est pas un morphème : *reauscultar*<sup>2</sup> : *rjauscultar*. Or, considérer qu'en portugais le glide prévocalique appartient à la rime reviendrait à accepter quatre segments dans ce constituant. Pour cette raison, en plus du fait que les glides prévocaliques ne se nasalisent pas s'ils sont suivis d'une voyelle nasale, Mateus & D'Andrade postulent que les glides prévocaliques appartiennent à l'attaque syllabique.

Voici comment les auteurs illustrent l'application de la règle qui transforme une voyelle haute en un glide appartenant à l'attaque de la première syllabe du mot *criais* (Mateus & D'Andrade, 2000, p. 51) :



Nous pouvons constater qu'il y a des arguments pour considérer que les glides occupent une position consonantique ainsi que des arguments pour considérer qu'ils occupent plutôt une position vocalique, ce qui est dû au comportement hybride de ces entités. Nous avons donc décidé d'analyser cette question à la lumière des prémisses du cadre CVCV, un cadre plus

<sup>2</sup> Exemple de Mateus & D'Andrade (2000).

« radical » de la phonologie de gouvernement, car nous pensons qu'un cadre phonologique sans constituants comme celui-ci permet d'éliminer un des grands problèmes présents dans les autres analyses : celui de savoir si le glide appartient à une attaque branchante, à un noyau branchant, ou à une coda.

### 3 Les hiatus et les diphtongues dans la théorie du CVCV

Dans la version « classique » de la phonologie de gouvernement (Kaye, Lowenstamm & Vergnaud 1990), on reconnaît comme constituants syllabiques seulement l'attaque, le noyau et la rime ; la coda ne serait pas un constituant. Lowenstamm (1996) va plus loin et pose que la structure syllabique universelle est du type CV (il n'y a pas d'attaque branchante ni coda). Selon lui, les syllabes fermées, les consonnes géminées et les voyelles longues peuvent être réduites à une configuration simple CV en admettant l'existence de noyaux et d'attaques vides. Scheer (1998, 2004) développe la théorie dite du CVCV ou « CV strict ».

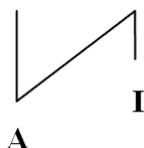
Dans un cadre sans constituants comme celui du CVCV, les propriétés affichées par les consonnes, voyelles et les glides découleront de leur architecture interne (de la façon dont les éléments qui les composent s'organisent) et de quelques principes fondamentaux à la théorie de gouvernement comme le gouvernement propre et le licenciement. Il n'y a pas d'algorithme de syllabation et de ce fait le comportement des glides peut être décrit plus aisément. La théorie CVCV peut concilier les aspects apparemment contradictoires du comportement des glides dans les diphtongues, ainsi que la variation existante entre les diphtongues croissantes et l'hiatus.

Nous avons vu que certains hiatus n'alternent jamais avec la diphtongue, et que certaines diphtongues n'alternent jamais avec l'hiatus. Par conséquent, il faut que, d'une part, les deux membres de l'hiatus stable soient clairement séparés ; d'autre part, il faut que dans les diphtongues décroissantes, les deux membres ne puissent pas se séparer. Enfin, il faudra aussi rendre compte des cas où la variation entre l'hiatus et la diphtongue croissante existe (lorsque la première voyelle est haute et non accentuée). Nous avons donc besoin d'une représentation qui rende compte de ces comportements. Afin de différencier l'hiatus des diphtongues, Caratini (2009) propose une représentation de la diphtongue dans le cadre de la théorie CVCV dans laquelle les positions vocaliques doivent *partager* du matériel mélodique :

a) *pai* [aj]

b) *pau* [aw]

C1 V1 C2 V2



C1 V1 C2 V2



Le fait que les deux positions aient quelque chose en commun différencie la diphtongue de l'hiatus. Si telle est la structure des diphtongues décroissantes, on s'explique pourquoi elles n'alternent pas avec l'hiatus : les deux segments sont liés, alors que l'hiatus présuppose l'indépendance complète des deux segments vocaliques l'un par rapport à l'autre :

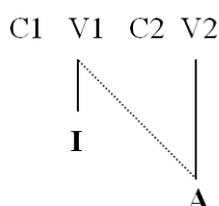
c) *saída* [ai]

d) *baú* [au]

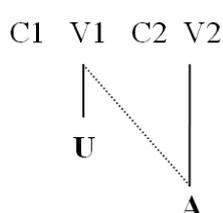


Pour expliquer que certaines séquences en hiatus admettent une prononciation en diphtongue croissante, on postule qu'au niveau phonologique il s'agit d'hiatus, qui, phonétiquement, alternent avec les diphtongues croissantes car il y a propagation des éléments **I** et **U** vers la deuxième position vocalique. Si cela est vrai, la représentation des diphtongues croissantes dans le cadre CVCV est celle-ci :

e) *piada* [ja]



f) *luar* [wa]



La controverse autour de la position syllabique occupée par les glides, c'est-à-dire s'ils appartiennent à l'attaque ou au noyau syllabique branchant ou à la coda, est ainsi résolue : il s'agit toujours des positions vocaliques (V) dans les deux types de diphtongues observées. Les diphtongues décroissantes sont phonologiques : le partage du matériel vocalique entre les deux positions vocaliques est déterminé par le système phonologique de la langue. Les diphtongues croissantes sont le résultat de la propagation du matériel vocalique vers la position vocalique adjacente à droite, un mécanisme strictement phonétique activé pour éliminer l'hiatus.

## 4 Conclusion

La stabilité des diphtongues décroissantes d'une part, et la variation entre l'hiatus et les diphtongues croissantes, d'autre part, portent à croire que seules les diphtongues décroissantes sont phonologiques en portugais, comme le soutiennent plusieurs analyses. Un point de désaccord entre les analyses présentées ici concerne la position syllabique occupée par les glides, ceux des diphtongues décroissantes comme ceux des diphtongues croissantes. Sur ce point, d'après les analyses exposées, le glide peut appartenir à l'attaque, à un noyau branchant ou à la coda syllabique. Ce problème touche donc la structure syllabique supposée. Dans un cadre théorique sans arborescence syllabique comme celui du CVCV, ce problème ne se pose même pas : les segments composés d'éléments consonantiques appartiennent à une position C et les segments ayant des éléments vocaliques dans leur composition appartiennent à une position V. Les rapports entre les segments s'établissent d'après leur complexité interne (nombre et organisation d'éléments) et aux mécanismes de gouvernement. Une explication plus simple en ce qui concerne le comportement des glides en portugais est donc possible dans le cadre du CV strict, où ces segments, dans les diphtongues décroissantes ou dans les diphtongues croissantes, seront toujours associés à une position V.

## **Références**

- BISOL, L. (1999). A sílaba e seus constituintes. In: NEVES, M. H. de M. (org.). *Gramática do português falado*. 2. ed. São Paulo: Humanitas; Campinas: Editora da UNICAMP, v. VII: Novos Estudos, 701-742.
- CÂMARA JR. J. M. (1970 41<sup>ed</sup>). *Estrutura da língua portuguesa*. Petrópolis : Vozes.
- CARATINI, E. (2009). *Vowel and consonantal quantity in German: synchronic and diachronic perspectives*. PhD dissertation, Université de Nice-Sophia Antipolis.
- CUNHA, C. & CINTRA, L. (2008 5<sup>ed</sup>). *Nova gramática do português contemporâneo*. Rio de Janeiro : Lexikon, 2008.
- HARRIS, J. W. (1983). *Syllable Structure and Stress in Spanish*. Cambridge, Massachussets : MIT Press.
- KAYE, J., LOWENSTAMM, J. & VERGNAUD, J.-R. (1990). Constituent structure and government in phonology. In: *Phonology* 7, 193-231.
- LOWENSTAMM, J. (1996). CV as the only syllable type. In: Durand, J. & Laks, B. (orgs.). *Current Trends in Phonology: Models and Methods*, CNRS, Paris X: ESRI, v. 2, 419-441.
- MATEUS, M. H. & D'ANDRADE, E. (2000). *The phonology of Portuguese*. Oxford : Oxford University Press.
- RODRIGUES, M. (2007). O hiato no português: a tese da conspiração. *Letras de Hoje*. Porto Alegre, v. 42, n. 3, 7-26.
- SCHEER, T. (1998). A unified model of Proper Government. In: *The linguistic Review* 15, 41-67.
- SCHEER, T. (2004). *A lateral theory of phonology; Volume 1: What is CVCV and why should it be?* Berlin : M. de Gruyter.

## **Étude de la variation écholalique chez un enfant autiste**

Carmen Lucia Florez Pulido<sup>1</sup>

(1) Laboratoire MoDyco, Université Paris- Descartes  
kappulletto@gmail.com

**Résumé** La conception classique d'écholalie fait la distinction entre écholalie immédiate et écholalie différée. Toutefois, la littérature n'opère pas de distinction de degré d'intégration linguistique entre ces deux pôles. A notre sentiment, dans le cadre des troubles envahissants du développement (autisme), les variations entre ces deux types d'écholalies méritent d'être mieux caractérisées. En effet, la description détaillée des énoncés écholaliques peut contribuer à comprendre plus finement le langage des enfants autistes et son développement. Ainsi, la présente étude consiste à mettre en évidence, à travers différents exemples recueillis auprès d'un même enfant autiste, les modifications qu'il peut apporter aux énoncés qui lui sont adressés. Nous pensons, en effet, qu'il existe différents types d'écholalies correspondant à des intentions de communication et de signification différentes. Pour étayer notre hypothèse, nous avons mené des analyses qualitatives fines et détaillées à partir d'un corpus de données vidéo filmées lors de consultations thérapeutiques.

**Abstract** Echolalia is normally divided into immediate and delayed. There is no place for variations between these two types of echolalia. When we talk about the language of autistic children, those variations are really important: they help us to understand how autistic children communicate. This paper shows some examples of autistic child echolalia. Our goal is to put in evidence the modifications that an autistic child may perform on the language directed to him. In order to do this we analyse the videos taken during therapeutic consultations.

**Mots-clés :** Echolalie immédiate, écholalie différée, autisme infantile, Récupération communicationnelle.

**Keywords:** Immediate echolalia, delayed echolalia, autism on childhood, communicational recovery.

# 1 Introduction

L'autisme fait partie de la catégorie des troubles envahissants du développement. Les limitations communicationnelles des enfants autistes sont aujourd'hui bien connues. Le DMS IV les présente ainsi: «*Les caractéristiques essentielles du trouble autistique sont un développement anormal ou déficient de l'interaction sociale et de la communication [...]L'altération de la communication est, elle aussi, marquée et durable, elle affecte à la fois les capacités verbales et les capacités non verbales. Chez les sujets qui savent parler, on peut observer un usage stéréotypé et répétitif du langage*».<sup>1</sup>

L'écholalie s'inscrit dans cet usage répétitif et stéréotypé du langage: «*dans le cas du retard du langage, en particulier pour l'autisme, l'imitation verbale ou vocale est souvent appelé écholalie, définie comme la répétition de ce qui vient d'être dit*».<sup>2</sup> L'écholalie est d'ordinaire considérée comme un refus de dialogue. Notre propos est de montrer que contrairement à cette conception banale, elle doit être envisagée comme une première tentative destinée à établir un échange avec l'interlocuteur.

On distingue traditionnellement deux types d'écholalie : *l'écholalie immédiate*, qui est produite tout de suite après la production du modèle par l'adulte ou un pair d'âge, et *l'écholalie différée*, dont la source est lointaine et parfois difficile à déterminer en amont de l'échange où intervient la reprise.

## 1.1 Objectif de l'étude

Notre étude a pour but deux objectifs différents. Le premier est de mettre en évidence le clivage existant entre les concepts d'écholalie immédiate et d'écholalie différée et l'insuffisance de cette dichotomie elle-même. Parfois l'énoncé écholalique différé, en l'absence d'un modèle directement identifiable, ne permet pas de savoir s'il s'agit d'une production originale, créée par l'enfant ou s'il s'agit d'une reprise à distance.

Le deuxième est la nécessité de prendre les énoncés écholaliques comme étant porteurs de sens dans la communication.

## 1.2 Méthodologie

Le corpus étudié se compose de vidéos recueillies auprès d'un enfant lors de séances de psychothérapie suivies au Centre Alfred Binet ASM de Paris 13. Les trois exemples présentés font partie de la première séance d'entretien familial, il s'agit donc d'un corpus homogène.

L'enfant a 5 ans, au moment de l'enregistrement. Il présente un autisme moyen selon les tests CARS qui lui ont été appliqués. Il possède un contact oculaire inconstant, une mimique faciale pauvre, des sourires non adressés et un langage modeste en termes de grammaire et de syntaxe. La présence et l'autorité de sa mère lui donnent des repères. Il parvient à être en interaction avec elle.

Afin de mener une première étude qualitative «pilote» des énoncés écholaliques, nous avons

---

<sup>1</sup> DSM- IV- *Diagnostic and Statistical Manual of Mental Disorders* (2005).

<sup>2</sup> NADEL J. (2011).

séquence l'enregistrement vidéo de cette première séance d'entretien familial pour illustrer la façon dont l'enfant se sert de ces énoncés. Nous avons trouvé 19 exemples d'énoncés écholaliques. Notre analyse se centrera sur 3 exemples typiques.

Notre but est :

- de montrer le sens des écholalies dans le dialogue
- de trouver leur modèle dans le discours qui a précédé la parole de l'enfant
- de distinguer différents types d'écholalies

Dans la pièce où se passe l'enregistrement se trouvent présents l'enfant en question, accompagné de sa maman, le médecin et l'assistante Sociale (qui filme) ainsi qu'un nouvel intervenant qui va être présenté à l'enfant.

## **2 L'écholalie et l'enfant autiste**

L'écholalie est un «reformatage» des productions orales faites par les adultes. L'enfant essaie d'accommoder, d'adapter le langage qu'on lui adresse. Le *copier-coller* du langage écholalique autiste nous montre comment l'enfant met en place les moyens dont il dispose pour associer deux situations: une situation antérieure et une situation actuelle. Ces deux situations sont liées chez l'enfant, nous ignorons de quelle façon. Nous cherchons des indices objectifs qui nous permettront de nous approcher de cette liaison étrange pour nous.

Si un enfant a recours à l'écholalie dans un contexte déterminé, c'est parce qu'il a déjà entendu l'énoncé qu'il reprend en écholalie mais aussi parce qu'il a compris cet énoncé. Par ailleurs, il l'associe à une situation particulière. Certains éléments de la situation où l'écholalie est produite lui évoquent ceux de la situation où l'énoncé initial a été prononcé, d'où sa production. Toutefois, contrairement aux enfants tout-venants, il ne peut pas restituer les seuls éléments qui seraient pertinents et construit un « bloc écholalique » dans lequel certains mots sont pertinents dans la situation d'emploi mais non tous. La difficulté pour l'enfant autiste réside apparemment dans la capacité à isoler certains mots de l'énoncé initial et à les réaménager. C'est un peu comme si l'enfant avait à produire une brique appartenant initialement à un ensemble, mais que, dans la manière qu'il a eu de la mémoriser, cette brique est collée à d'autres briques dont il n'arrive pas à la séparer. Quand il a besoin de désigner un point d'une situation, il doit donc sortir le bloc de briques dans sa totalité alors qu'une seule d'entre elles serait pertinente. L'énoncé écholalique est lié à un phénomène de non décontextualisation et de non recontextualisation. Ce n'est pas l'intention communicative qui est fautive, mais le savoir faire qui permet d'adapter les énoncés entendus à la situation de leur réemploi.

Le langage écholalique signe chez l'enfant autiste effort pour entrer en communication avec autrui. Nous allons donc l'aborder comme le point de départ de son développement langagier, comme une première étape lui permettant à terme d'entrer dans une communication banale.

Nous allons voir comment un même enfant autiste peut faire recours à différents types d'écholalies. Nous montrerons que les variations liées au contexte de dialogue que l'adulte parvient à établir avec l'enfant.

### 3 Analyse du corpus recueilli

Nous avons mis les éléments écholaliques en gras, et les éléments ajoutés par l'enfant en souligné.

#### 3.1 C'est une petite caméra...

L'enfant arrive à la consultation. Il se rend compte qu'il est filmé.

Enfant: (regarde la caméra)

Mère: C'est une petite caméra.

Enfant: **c'est une petite caméra, c'est une petite caméra**, (il entoure sa mère avec ses bras) ah ben oui, maman...maman bobo (il s'accroche au cou de sa mère).

Mère: (en asseyant l'enfant sur ses genoux) Il faut que tu descendes.

Docteur: (en regardant l'enfant). Elle te fait peur avec sa petite caméra?

Mère: une caméra et il y a un monsieur en plus.

Enfant: (Regarde le nouvel intervenant).

*C'est une petite caméra* est une écholalie immédiate, l'enfant a repris l'énoncé de sa mère et l'a reproduit tout de suite. Mais son intervention ne se limite pas à la reproduction fidèle de l'énoncé, il y apporte une variation. L'utilisation de «oui» et «ben oui» (certainement écholalies aussi) montrent comment l'enfant peut constater la réalité qu'il observe après l'intervention de sa maman. «Bah oui» peut être interprété comme «oui maman tu dis que c'est une petite caméra et je veux être sûr que je vois la même chose». Nous avons déjà un élément ajouté par l'enfant qui enrichit et complexifie l'énoncé, une variation à l'intérieur de l'écholalie, même si l'élément ajouté par l'enfant ne vise pas à relancer l'interlocuteur mais semble plus du registre d'un commentaire auto-adressé.

L'énoncé «Maman, maman bobo» pourrait révéler quelque chose d'écholalique. L'enfant ne le dit pas parce qu'il s'est fait mal, il le dit parce que la situation le déstabilise et il a besoin de sa mère pour se recentrer. En effet, il veut provoquer en elle la même réaction que lorsqu'il se blesse. Le fait de s'agripper à sa mère appuie notre hypothèse. L'enfant met en mots le fait de se sentir en difficulté face à la camera, en même temps qu'il cherche le réconfort maternel. Pour ce faire il a recours à un énoncé de type écholalique dont la source n'est pas présente dans l'échange, mais semble présente dans la mémoire de l'enfant.

Le lexique utilisé par l'enfant est en inadéquation avec la situation. Le mot bobo a été cristallisé, ainsi que la réaction de la mère à ce mot. La trace persistante de la coordination énonciative non pertinente nous laisse penser à un effort de décontextualisation.

### 3.2 C'est toi qui l'a défaite...

Le thérapeute regarde jouer l'enfant. Le jeu consiste à mettre des barrières en plastique l'une à côté de l'autre. Le thérapeute essaye de rythmer le jeu de l'enfant.

Thérapeute: Encore une? , encore une?

Enfant: Oui (il place soigneusement la barrière).

Thérapeute: Allez, encore. Ouais...Encore une ?.

Enfant: Oui (il place soigneusement la barrière). je ferais plus(incompréhensible).  
**C'est toi qui l'as défaite.**

Thérapeute: C'est toi moi qui l'ai défaite?

Enfant (en prenant une autre barrière) **c'est toi qui a défait ça, ça...**

Thérapeute: Ahhhh

Enfant: **Ahhhh** (avec la même intonation, il met la barrière à côté de l'autre).

Thérapeute: C'est cassé, tu as raison. Encore? Et encore? Ça fait comme un train.

Enfant: **Comme un train.**

Quand l'enfant a vu qu'une barrière était cassée, il n'a pas pu dire «ah c'est défait ». Il utilise à la place un autre énoncé qu'il doit déjà connaître, avec le mot *défaite*: «c'est toi qui l'a défaite». Il n'est pas capable d'isoler le mot «défaite» dont il a besoin pour construire une phrase qui puisse attirer l'attention de l'autre sur quelque chose qu'il vient de découvrir. Nous serions face à une écholalie différée par le biais de laquelle l'enfant arrive à donner du sens à ce qu'il observe et qu'il désire partager avec son interlocuteur. Dans les faits, il n'est évidemment pas possible que la personne ayant cassé la barrière ne pouvait pas être le thérapeute, interagissant avec l'enfant à ce moment-là, puisque c'est la première fois qu'il voit l'enfant et dispose du matériel de jeu. Force est alors d'envisager l'énoncé comme l'expression d'une appétence de l'enfant pour l'échange avec autrui.

L'enfant reprend l'expression de l'étonnement immédiatement après l'adulte. En fait c'était ça qu'il voulait faire quand il a constaté l'état de la barrière : exprimer son étonnement. Il ne peut que le faire une fois qu'il a entendu l'autre le faire.

L'adulte continue «c'est cassé, oui tu as raison». L'enfant ne répète pas les mots de l'adulte. Pour quoi il ne répète pas cet énoncé ? Peut être parce qu'il n'apporte rien de nouveau. L'adulte dit «encore?», l'enfant met la dernière barrière. L'adulte ajoute: «C'est comme un train» et l'enfant répète «Comme un train». Cette fois-ci, l'enfant reprend une partie de l'énoncé de l'adulte. Nous croyons bien qu'il s'agit d'un énoncé écholalique «confirmatif», comme ceux qu'on trouve chez les enfants tout venants. Il répète pour montrer à l'autre que l'information donnée lui convient, qu'il l'a assimilée, qu'il est d'accord avec son interlocuteur : l'enchaînement linéaire des barrières fait , en effet, penser aux wagons d'un train. Le *comme un train* est une marque d'un processus d'intégration de concepts, mais pas une relance du dialogue.

### 3.3 Tu veux l'autre?

L'enfant et le thérapeute sont en face à face. Le thérapeute lance une voiture, l'enfant la reçoit.

Thérapeute: tu veux une autre?

Enfant : Oui.

Thérapeute: ...(s'attarde à lancer une nouvelle fois la voiture).

Enfant: (en pointant vers son interlocuteur) **Tu veux une autre.**

Thérapeute: Tu es prêt?

Enfant: Oui.

L'enfant a repris l'énoncé de l'adulte sans pouvoir y apporter la modification qui était nécessaire; le changement du pronom. Il n'arrive pas encore à faire la distinction entre je /tu. C'est une écholalie traditionnellement immédiate, mais l'enfant a ajouté une variation gestuelle: le pointage, qui contribue à une meilleure compréhension de l'échange en restant tout à fait paradoxal. L'enfant parle et pointe l'autre, mais c'est lui-même qui veut l'autre voiture. Il pointe vers l'objet qu'il désire.

Il est utile de remarquer ici que l'enfant est capable de répondre «oui» à chaque proposition faite par l'adulte. C'est seulement quand l'enfant perçoit une variation temporelle dans le tour de parole de l'autre, qu'il va prendre l'énoncé écholalique pour manifester à l'adulte sa volonté de continuer le jeu, renforcée par le pointage. Comment rendre compte de la valeur communicationnelle de cet énoncé écholalique ?

## 4 Conclusion - bilan, discussion

La classification générale d'écholalie immédiate et d'écholalie différée est assez restreinte et ne tient pas compte des degrés des variations proposées par l'enfant. Il s'avère nécessaire de trouver une autre façon d'aborder l'écholalie, qui puisse rendre compte des nuances des productions des enfants autistes et de ses variations. C'est dans les modifications aux énoncés de départ, faites dans l'écholalie, que l'enfant autiste trouve le pont entre une communication stéréotypée et un langage plus dynamique. Nous proposons de définir ce comportement verbal comme étant le fruit d'une reprise partielle de l'énoncé de l'interlocuteur, c'est à dire une «écholalie partielle». Nous avons vu que l'enfant reprend une partie de l'énoncé de l'adulte, certes, mais en ajoutant un autre élément pertinent qui peut aussi être une écholalie, un pointage. En fait, chez un même enfant autiste les degrés d'écholalies qui vont de l'écholalie immédiate à l'écholalie différée s'avèrent révélateurs de l'évolution de ses capacités langagières.

Une fois que l'enfant autiste a pu mettre en place l'écholalie, principalement celle qualifiée d'écholalie différée, son langage tend à évoluer. Petit à petit, il arrivera à décortiquer de manière plus efficace le langage qui lui est adressé pour le transformer en une série d'énoncés «prêts à l'emploi» dans une interaction plus dynamique. L'écholalie, remaniée par l'enfant, donnera un peu plus de fluidité au dialogue et permettra l'utilisation des tours de parole. Ces variations verbales ou gestuelles sont, dans la plupart des cas, très significatives, bien que subtiles. Si l'enfant a recours à l'écholalie c'est qu'il a quelque chose à dire, et qu'il a compris

qu'on attend quelque chose de lui.

Il est assez difficile de cliver les interventions de l'enfant entre écholalie immédiate ou écholalie différée. Une conception plus flexible et plus intégratrice de l'écholalie est nécessaire. En outre d'un point de vue clinique, considérer l'écholalie comme une prise de parole de l'enfant autiste permet d'en faire le point de départ d'une progression linguistique en évitant de la disqualifier comme vide de sens.

Les trois exemples explorés nous donnent une ébauche de la façon dont fonctionne l'écholalie chez les enfants autistes et nous laissent voir comment les variations qui se présentent entre l'écholalie immédiate et l'écholalie différée doivent être prises en compte dans l'étude de la pathologie autistique. L'énoncé écholalique est donc un outil que le thérapeute possède pour soutenir la récupération langagière chez l'enfant. Plus l'enfant ressent que ce qu'il construit avec son langage est compris, plus il va s'efforcer de communiquer.

Une étude ultérieure nous permettra de mieux classifier les énoncés écholaliques repérés dans les vidéos prises à intervalle d'un an, de les comparer et d'établir une progression au fil du temps jusqu'à l'arrivée du langage dont dispose l'enfant dans l'actualité.

## **Références**

BERQUEZ. G. (1983). *L'autisme infantile. Introduction à une clinique relationnelle selon Kanner*. Paris : PUF.

DANON-BOILEAU L. (2002). *Des enfants sans langage*, Paris : Odile Jacob.

GEPNER B. (2006). Le monde va trop vite pour les personnes autistes! Hypothèses neurophysiopsychopathogéniques et implications rééducatives, In *Neuropsychiatrie de l'enfance et de l'adolescence* 54, Aix-en- Provence, pp 371-374.

Manuel diagnostique et statistique des troubles mentaux DSM- IV-TR. (2005). France : Masson Editeur.

TUSTIN F. (1977). *Autisme et psychose de l'enfant*. France : Éditions du Seuil.

NADEL J. (2011). *Imiter pour grandir, développement du bébé et de l'enfant avec autisme*. Paris : Dunod.

TOMASELLO M. (2007). A new look at infant pointing, In *Child Development*, Volume 78, Number 3, 705-722.

## **Analyse psychopathologique comparée des usages du signe dans la névrose et la psychose**

Arthur Mary<sup>1</sup>

(1) LIRCES & Université de Nice-Sophia Antipolis  
a.mary@gmx.com

**Résumé** L'auteur, psychologue clinicien, rend compte d'une construction théorique visant à distinguer différents usages possibles du signe linguistique selon les trois grandes structures éthiques (névrose, psychose et perversion) traditionnellement reconnues par la psychopathologie clinique d'orientation psychanalytique. Il propose la modélisation de deux principaux usages du signe (usage référentiel et usage sui-référentiel) ainsi que du rapport du locuteur à la langue qui se déduit de son usage du signe.

**Abstract** The author, a clinical psychologist, reports a theoretical construct to distinguish the possible uses of the linguistic sign according to the three major ethical structures (neurosis, psychosis and perversion) traditionally recognized by the psychoanalytically oriented clinical psychopathology. He proposes the modeling of two major uses of the sign (a referential use and a sui-referential use) and the relation between the speaker and the language system which is derived from his/her use of the sign.

**Mots-clés :** signe, discours, éthique, névrose, psychose, holophrase, signifiant

**Keywords:** sign, discourse, ethos, neurosis, psychosis, holophrase, signifier

## 1 Introduction

Les trois grandes structures psychopathologiques que sont la névrose, la psychose et la perversion, sont trois modalités différentes d'être-au-monde. Nous les considérons ici comme trois positions éthiques, c'est-à-dire, trois façons de se rapporter à autrui. Trois « mise[s] en question [différentes] de ma spontanéité par la présence d'autrui », pour reprendre une définition que donnait Emmanuel Levinas de l'éthique. Notre hypothèse est la suivante : dans sa façon de mobiliser la langue dans l'occasion d'un discours chacune de ces positions éthiques fait un usage spécifique du signe linguistique. Nous tâcherons de préciser cette spécificité ; ou pour le dire autrement, nous dégagerons ce qui demeure invariant dans l'inscription langagière de chacune de ces options éthiques. Nous présentons ici ce qui s'est déduit de rencontres cliniques, soit précisément le discours singulier d'un locuteur particulier avec autrui (en l'occurrence un clinicien) – et cela quel que soit le support sémiotique emprunté : verbal, écrit, graphique, gestuel... –, méthode du cas par cas donc, qui exige certaines précautions quant aux conditions et aux possibilités de généralisation de ce que la clinique rencontre de singulier.

Le recours à la logique et à la linguistique permet d'introduire certaines distinctions fines entre les différentes structurations éthiques que peuvent prendre les discours. Nous tenterons ainsi de tirer les conséquences pour chacun de la structure de la langue, et particulièrement de l'articulation différentielle des signes linguistiques entre eux. Et cette différence joue son rôle aussi bien dans les rapports syntagmatiques (rapports de successivité, de contiguïté des signes dans la chaîne parlée, rapports dont le signe tire sa *signification*) que dans les rapports paradigmatiques (rapports associatifs hors de la chaîne du discours, rapports dont le signe tire sa *valeur*). Ce qui paraît distinguer les différentes structures éthiques tient ainsi à l'*usage* qui est fait du signe linguistique et donc à l'attitude spécifique à l'égard des éléments langagiers. Saussure a proposé d'imaginer la langue comme une symphonie dont la « réalité est indépendante de la manière dont on l'exécute ; les fautes que peuvent commettre les musiciens qui la jouent ne compromettent nullement cette réalité » (Saussure, 1972, 36). Façon de souligner que la langue est structurée de la même façon pour tous et indépendamment de ce qu'on y fait mais qu'il existe néanmoins des manières singulières de l'exécuter (et on pourrait sans doute élargir cette perspective aux systèmes sémiotiques a-signifiants). Ainsi, quelle que soit la façon d'exécuter la langue, la propriété différentielle des signifiants n'est pas modifiable. Nous tentons de préciser la spécificité de l'usage du système-langue par chacune de ces trois éthiques. J. Kristeva emploie le terme d'*idéologème* pour cerner une valeur historiquement déterminée du signe linguistique, dont l'évolution et les transformations façonnèrent intrinsèquement l'histoire littéraire (Kristeva, 1974). Je propose d'entendre ce concept d'idéologème comme le système sous-jacent qui régule l'usage du signe et structure *a priori* son champ et sa modalité de déploiement *dans le discours*. Idéologème est le nom que je donne ainsi aux différents usages possibles du signe dont les propriétés intrinsèques, insistons, sont néanmoins les mêmes pour tous. Ces différents usages sont en somme des attitudes, des dispositions, des façons de se rapporter aux éléments d'un système symbolique et de les mobiliser singulièrement dans un discours. Nous en reconnaissons deux principaux : le *signifiant* (lacanien) dans ses usages névrotique ou pervers et l'*holophrase* psychotique.

Le tableau suivant présente le développement qui lui fera suite. Faute de place pour détailler suffisamment clairement l'usage pervers du signe, nous renonçons à le présenter ici. J'ajoute une dernière ligne ouverte, car reste à comprendre la spécificité de l'idéologème au cœur de l'autisme que j'écarte méthodiquement de ce travail inaugural. Il ne faut pas exclure non plus

que des idéologèmes particuliers puissent être à l'œuvre dans la schizophrénie, la mélancolie, ou les états limites. Prenons soin enfin de toujours maintenir ouvert ce tableau, ou mieux : de savoir nous affranchir des nosographies quelles qu'elles soient, afin de ne pas être sourds aux inventions jusqu'à présent inouïes que nous réservent les locuteurs qui prendront la parole, quitte à ce que ces « inventions de l'autre » (Derrida, 1984) échappent totalement à cette classification programmatique. C'est donc comme abstraction théorique qu'il faut entendre mon jeu de concepts. Concepts toujours formés « de l'oubli de ce qui différencie un objet d'un autre » (Nietzsche).

Ethique	Idéologème	Usage	Articulation
Névrose	Signifiant	Usage représentationnel	$S_2=\{S_1,\emptyset\}$ différentiel, hétérogène
Perversion	Signifiant	Usage exclusif	$S_2=\{S_1,\pm\}$ différentiel, homogène
Psychose	Holophrase	Usage sui-référentiel	$H_n=\{H_n,\$ \}$ différentiel ou $H_n=\{H_n\}$ non-différentiel
...			

Tableau 1 : Trois usages du signe linguistique et leurs articulations respectives

Légende :  $S_n$  : signe linguistique articulé différemment aux autres éléments de la langue.  $H_n$  : signe linguistique sui-référentiel, éléments d'une langue.  $\emptyset$  : ensemble vide, ne contenant aucun élément.  $\pm$  : valeur qualitative établissant l'homogénéité des signes linguistiques.  $\$$  : locuteur.

## 2 Névrose

Lacan a proposé une théorie des discours de névrose à partir d'une conception assez originale du signifiant (détournement du signifiant saussurien largement critiquable par la linguistique). Rappelons la définition qu'il en donne : « un signifiant est ce qui représente un sujet [un locuteur] auprès d'un autre signifiant » (1968-1969). Il propose que le signifiant représentant le locuteur soit désigné par  $S_1$  et que l'autre signifiant auprès duquel le locuteur est représenté soit noté  $S_2$  ; le sujet locuteur sera noté  $\$$ . Notons d'emblée la dimension sémantique qui se dégage de l'articulation des signifiants lacaniens entre eux : rétroactivement, le second signifiant éclaire le premier ; mais le second est toujours susceptible d'être suivi d'un autre, puis d'un autre.

La clinique atteste au cas par cas de la générativité sémantique infinie du discours de névrose depuis que Freud invente un dispositif sensible aux associations morphématiques et phonétiques qui font sentir combien les discours de névrose sont gros de doubles sens et d'interprétations infinies (Freud, 1898 par exemple). Exemple : Un locuteur peut signifier sa douleur avec un morceau de pain ( $S_1$ ) dont la signification sera éclairée par l'association graphique avec le mot anglais *pain* ( $S_2$ ). Les discours de névrose sont, au moins tacitement, construits sur la mobilisation de l'Autre : le signifiant, tel que Lacan en a fourni une théorisation à l'usage du psychanalyste, est moins lié au signifié qu'aux autres signifiants

avec qui il s'articule dans un système de renvois infinis. Voilà l'usage discursif que fait le névrosé du signe linguistique. Il s'agit d'un usage représentationnel du signifiant, c'est-à-dire que le discours de névrose est l'expérience que fait le sujet de ne pas exister ailleurs que dans la représentation que le langage peut lui fournir de lui. Pourtant ces mots, et l'ordre du langage tout entier, sont insuffisants pour résoudre entièrement l'énigme ontologique : qu'est-il, ce sujet particulier ? La logique sous-jacente à son activité de symbolisation, fût-elle une stochastique, est-elle formalisable, peut-elle s'écrire ?

Les effets de sens se déduisent de la nature différentielle des signifiants entre eux. Cependant, une autre propriété importante distingue le « signifiant maître » ( $S_1$ ) du signifiant second ( $S_2$ ), c'est qu'ils ne relèvent pas du même niveau, ils ne sont pas de même nature. Pierre Bruno le précise ainsi : « L'implication signifiante  $S_1 \rightarrow S_2$  est inintelligible si elle n'est pas entendue comme la conséquence de la contradiction (il n'y en a qu'une, celle que Russell a découverte et formulée en 1901<sup>1</sup>) qui hétérogénéise les deux termes.  $S_1$  doit être entendu comme le signifiant qui fait intrusion, par forçage logiquement indu, dans le langage. Il se veut représentatif, mais ne peut l'être seul.  $S_2$ , renonçant à l'intrusion, n'est pas représentatif. » (2003, 7) Dès lors, si les propriétés linguistiques de  $S_1$  et  $S_2$  sont fondamentalement les mêmes, leurs participations respectives au discours sont rigoureusement hétérogènes, asymétriques. Bruno en fait une démonstration aux accents russelliens comme suit : « Disons  $S_1$  le signifiant qui s'inscrit dans l'Autre alors qu'il ne le devrait pas. C'est le signifiant d'intrusion [...]. Disons  $S_2$  le signifiant second qui ne s'inscrit pas dans l'Autre, car, sinon, il se confondrait avec  $S_1$ . Ce  $S_2$ , ne s'inscrivant pas dans l'Autre, pourrait dès lors s'y inscrire. [...] C'est en ce sens que le signifiant  $S_2$  est tout à fait asymétrique par rapport au  $S_1$ . Il est un représentant à jamais non représentatif, et ce qu'il dénote est un objet qui ne pourra jamais être objectivé, c'est-à-dire appartenir à la collection des objets qui peuvent entrer dans une proposition ayant du sens. » (2003, 91)

L'Autre, « trésor des signifiants » (Lacan, 1999, 206), se définit comme l'ensemble des signes linguistiques s'articulant différentiellement.  $\text{Autre} = \{S_1, S_2, \dots, S_n\}$ . Or, le sujet emploie  $S_1$  pour être représenté dans l'ordre du langage (« l'intrusion, par forçage logiquement indu ») :  $A = \{\$, S_2, \dots, S_n\}$ . Répétons que le sujet n'est pas un signifiant, qu'il n'est que représenté et ce faisant il joue sur la propriété qu'a le signifiant de représenter. Le  $S_2$  en s'articulant différentiellement<sup>2</sup> à  $S_1$  délimite nécessairement un écart (noté  $\emptyset$  et correspondant à l'« objet  $a$  » de Lacan) :  $S_2 = \{\emptyset, S_1\} = \{\emptyset, \{\$\}\}$ , c'est-à-dire cerne l'impossibilité logique à s'articuler sans reste à  $S_1$  du fait de l'articulation *par définition* différentielle des signifiants entre eux. L'intrusion du signifiant, c'est l'évènement langagier nouveau par lequel le locuteur vient à s'inscrire dans le langage, à s'y faire représenter. Manifestations de l'inconscient qu'on a pu lister sans doute trop restrictivement en symptôme, lapsus, acte manqué, rêve, trait d'esprit, oubli... et que l'on généralisera en tout ce qui se tient à fleur de dire (constante équivocité, infinité des interprétations...). Le signifiant non-intrusif s'articulant à  $S_1$  est comme condamné à ne pouvoir représenter (ou entourer) que l'irreprésentable, le non symbolisable, soit à se tenir sur l'accorde de la contradiction. Autrement dit, il indique négativement ce qui du sujet ne peut s'inscrire dans le langage (ce qui demeure donc asémantique).

### 3 Psychose

Le régime psychotique de la parole mobilise le signe linguistique dans un usage

<sup>1</sup> Cf. Russell, 1903, 101.

<sup>2</sup> La nature différentielle du signifiant pourrait être écrite ainsi :  $\forall n, \neg \exists x, S_n = S_x$ . Nous lisons : quel que soit  $n$ , il n'existe pas de  $x$  tel qu'un signifiant  $n$  serait identique à un signifiant  $x$ . Nous écrivons aussi  $S_n = \{S_x, \emptyset\}$ .

autoréférentiel, pour lequel le signe se désigne lui-même (Lacan qualifie cet usage d'holophrastique, bien qu'à proprement parler il ne s'agisse pas exactement du trope du même nom ; à défaut d'un meilleur terme, nous désignerons cet usage du signe comme « holophrase »). Exemple : un locuteur me rapporte son étonnement que le signifié « pain » soit articulé au signifiant « pain », qu'il est étonnant que le signifiant « pain » correspondent justement à la sonorité /pɛ̃/. (On sera sensible à l'effet de concentration particulière sur le signe conduisant à l'expression d'une tautologie et d'une perplexité du locuteur (et de l'allocutaire) ; tout se passe comme si le signifiant « pain » était pris pour signifié ; pour le dire avec Freud : une représentation de mot traitée comme représentation de chose.) On en déduit deux conséquences simultanées alternatives : 1°) le signe holophrasé ne permet pas au sujet d'être représenté dans le langage dans la mesure où l'holophrase ne fait pas référence au sujet mais seulement à elle-même ; et/ou 2°) le signe holophrasé représente effectivement le sujet dans le langage, mais en l'identifiant à son propre signe, c'est-à-dire en dégradant le « locuteur » en pur signe linguistique parlé par l'Autre. Ainsi, si l'Autre de la névrose pouvait être défini comme « trésor des signifiants », ou encore comme l'ensemble des différents signes linguistiques qui ne se représentent pas eux-mêmes (et une telle définition implique une tension paradoxale quant à l'auto-inclusion ou non de l'Autre), l'Autre de la psychose serait un « trésor d'holophrases », dont l'extension serait réduite *in fine* à  $\text{Autre}=\{\text{Autre}\}$ . Penchons-nous sur chacune des deux alternatives. Elles se déduisent en somme de la définition que nous retiendrons de l'holophrase. Ou bien l'holophrase est strictement autoréférentielle, ou bien elle est autoréférentielle mais le sujet peut néanmoins l'employer pour se représenter.

Premier cas, celui de l'holophrase strictement autoréférentielle, « sans perte » :  $H_n=\{H_n\}$ . L'Autre se définit alors comme classe d'éléments ne se référant strictement qu'à eux-mêmes :  $\text{Autre}=\{H_1, H_2, \dots, H_n\}$  ; ensemble duquel l'Autre est remarquablement exclu dans la mesure où il ne se mentionne pas uniquement lui-même et duquel le Sujet est exclu dans la mesure où nul élément de cette classe ne représente autre chose que lui-même. Nous sommes ici face à un système langagier apparemment autonome où nul sujet ne fera jamais intrusion tant la définition strictement autoréférentielle constitue une barrière infranchissable. Cela tient donc moins aux propriétés du langage qu'à la prémisse particulière qui conçoit le signe linguistique comme holophrase ; il s'agit d'une « attitude » particulière à l'égard des éléments langagiers, d'une façon de se tenir face à l'ordre du langage. Or, cette attitude établit la langue comme système de premier ordre et complet. La catatonie peut être une façon muette et immobile de faire correspondre tautologiquement le corps avec lui-même, d'épouser exactement l'espace occupé par un corps exclu de l'ordre langagier.

Second cas, l'holophrase, autoréférentielle, représente aussi le sujet locuteur :  $H_n=\{H_n, \$\}$ . Alors, l'Autre se définit comme classe des éléments se référant à la fois à eux-mêmes et au sujet :  $\text{Autre}=\{\{H_1, \$\}, \{H_2, \$\}, \{H_3, \$\} \dots \{H_n, \$\}, \{A, \$\}\}$  ; ensemble dans lequel chaque élément représente aussi le sujet et dans lequel la classe entière peut être incluse car elle représente le sujet et se représente elle-même toute entière (on reconnaît dans cette notation la signification personnelle que chaque élément du langage est susceptible de prendre dans certains vécus psychotiques). On déduit de cette formalisation de l'Autre, la notation suivante :  $\text{Autre}=\{\text{Autre}, \$\}$  (à la condition que \$ satisfasse aux conditions d'inclusion dans l'ensemble : soit \$ un élément se référant à lui-même et au sujet) et nous ne nous étonnons pas d'y reconnaître la structure même de l'élément holophrasé venue récursivement se refléter dans la structure de l'Autre. Cependant, ce deuxième cas soulève une difficulté : chacun des éléments langagiers y assume *réflexivement* la propriété différentielle du signifiant :  $H_n=\{H_n, \emptyset\}$  et cela tient à notre prémisse qui définissait l'holophrase comme ce qui se désigne *et* représente le

sujet. Autrement dit, chaque holophrase, à représenter *aussi* le sujet, est toujours *un peu* différente d'elle-même, toujours seconde, et l'holophrase ne peut donc pas se désigner exactement elle-même comme le voudrait sa définition. D'où,  $H_n = \{H_n, \$\} = \{\{H_n, \emptyset\}, \$\} = \{S_1, \$\} = S_2 = \{S_2, \$\}$  et  $S_n \neq S_n$  chaque holophrase représente le sujet au lieu logique de la pure différence, au lieu du Réel. C'est-à-dire que chaque élément de ce langage holophrasé s'articule différentiellement, où c'est le sujet représenté dans le langage qui assume la fonction de pure différence entre les signes. Le sujet s'éprouve étrangement comme manquant toujours dans l'Autre, comme objet cause du désir de l'Autre. Il est circonscrit par chacun des éléments du langage. *Il est représenté par une holophrase auprès d'une autre holophrase qui ne diffère que du sujet.* Cette formulation rendrait compte tant de la paranoïa, de l'érotomanie, que dans une certaine mesure de la mélancolie, si par mélancolie, on entend l'identification à l'objet archaïquement (éternellement) perdu.

## 4 Discussion

En mobilisant le vocabulaire théorique de sa discipline (la psychopathologie clinique), l'auteur regrette les confusions possibles avec les concepts de la linguistique (en particulier, la notion lacanienne d'holophrase, réductible à l'autonymie). Il est ouvert aux suggestions pour désigner plus correctement ce dont il parle. Par ailleurs, cette réflexion, largement exploratoire, schématique et limitée par la forme de sa communication, met de côté certaines positions éthiques intéressantes. Notamment, comment le sujet autiste mobilise-t-il le signe linguistique et dans quel type de discours s'engage-t-il ? De plus, précisons bien que cette analyse au niveau sémiotique ne constitue évidemment pas une étiologie des positions éthiques ; elle ne renseigne pas sur les causes de ces usages. On pourra s'interroger sur l'intérêt diagnostique de la démarche. En portant son attention sur le maniement du signe linguistique et de l'Autre (classe d'éléments sui-référentiels ou bien classe d'éléments signifiants), le clinicien peut dès lors ajuster éthiquement sa façon de se rapporter aux sujets-locuteurs auprès de qui il travaille. Enfin, cette étude, très focalisée sur le signe, sert de repérage préalable à une étude sur les implications discursives des grandes modalités d'usage du signe.

## Références

- BRUNO P. (2003). *La passe*, Toulouse : PUM.
- DERRIDA J. (1998). *Psyché, Invention de l'autre* [1984,1986], t. 1, Paris : Galilée.
- FREUD S. (1898). Sur le mécanisme psychique de l'oubli, *Résultats, idées, problèmes*, Paris : Puf.
- KRISTEVA J. (1974). *La révolution du langage poétique*, Paris : Seuil.
- LACAN J. (1999). *Les Ecrits* [1966], t. 2, Paris : Seuil.
- LACAN J. (2006). *Le séminaire XVI, d'un autre à l'Autre* [1968-1969], Paris : Seuil.
- RUSSELL B. (1903). *The Principles of Mathematics*, vol. 1, Cambridge : Cambridge University Press.
- SAUSSURE F. (1972). *Cours de linguistique générale* [1916], Paris : Payot.

## « L'économiquement correct » : une source d'ambigüités

Loredana Ruccella<sup>1</sup>

(1) MoDyCo & Université de Paris X  
loredana.ruccella@univ-nantes.fr

**Résumé** Ce travail vise à montrer de quelle manière l'euphémisation du discours économique, *vulgarisé* par la presse générale, favorise l'élaboration de variabilités interprétatives erronées se déclenchant, chez le lecteur, suite à la *vulgarisation* d'une information, volontairement ou involontairement, ambiguë.

**Abstract** The aim of this work is to study the way in which the euphemization of economic speech, *popularized* by the press, may trigger - following the *popularization* of news that is, intentionally or unintentionally, ambiguous - the reader's incorrect variability in the interpretation.

**Mots-clés :** *vulgarisation*, variabilités, interprétation, ambigüité sémantique, euphémisme, faux-euphémisme, discours économique.

**Keywords:** *vulgarization*, variability, interpretation, semantic ambiguity, euphemism, false-euphemism, economic speech.

## 1. Introduction

Notre quotidien est rempli d'incertitudes, d'ambiguïtés, d'informations voilées qui se répandent dans les médias à travers une rhétorique de plus en plus déstabilisante faite de mots, de termes ou d'expressions qui peuvent décrire la réalité ou la dissimuler, dire le vrai ou induire le faux, informer ou tenter de faire régner l'ignorance. Ces ambiguïtés<sup>1</sup> nous paraissent liées au recours, de plus en plus fréquent dans l'activité de *vulgarisation*<sup>2</sup>, au langage « économiquement correct »<sup>3</sup>, qui à travers l'emploi de *substituts potentiellement euphémiques* (SPE)<sup>4</sup> – répertoriés parmi les principaux procédés métalinguistiques employés par la presse – semble pouvoir mettre en péril la compréhension des événements présentés. Dans ce contexte notre questionnement initial peut s'exprimer ainsi : le langage « économiquement correct » encourage-t-il l'élaboration d'interprétations erronées du discours économique ?

## 2. « Corpus »

A la suite de notre questionnement, nous avons essayé de trouver une réponse à travers l'analyse d'un « corpus »<sup>5</sup> de presse écrite. L'analyse de notre « corpus » – analyse que nous ne pouvons que présenter partiellement au fil de cet article – nous a permis de remarquer que l'« économiquement correct » se présente comme une tendance de la presse générale à diffuser, volontairement ou involontairement, des informations homogènes, indépendantes de l'orientation politique du support dans lequel elles apparaissent et traduisant, à des degrés différents, la doxa libérale. Cela se vérifie via le recours à certaines stratégies discursives – qui ne seront pas traitées ici – et à des SPE pouvant compromettre la compréhension de l'information présentée. C'est à ces derniers que nous nous intéresserons afin de donner une réponse à notre questionnement initial.

---

<sup>1</sup> L'ambiguïté est la propriété d'un énoncé qui présente simultanément une pluralité de lectures et elle se réalise lorsqu'un signifiant unique suscite une pluralité d'interprétations (Villar, 2005).

<sup>2</sup> La *vulgarisation* pourrait, à notre avis, se définir comme une recontextualisation, c'est-à-dire comme l'adaptation - présupposant l'évaluation des connaissances hypothétiques du récepteur - d'un savoir issu d'un univers culturel d'experts à un univers culturel de non-experts.

<sup>3</sup> Expression créée à partir de « politiquement correct ». Cette dernière expression « se dit d'un discours, d'un comportement visant à bannir tout ce qui pourrait blesser les membres de catégories ou de groupes minoritaires en leur faisant sentir leur différence comme une infériorité ou un motif d'exclusion. (Calque de l'Amér. *politically correct.*) ». Définition du *Petit Larousse illustré* (Garnier & Vinciguerra, 2006).

<sup>4</sup> Tout signe qui en remplace un autre, désignant une réalité taboue.

<sup>5</sup> Notre « corpus » se compose d'article de presse écrite – en format numérique et accessible en ligne sur le site Internet du journal de référence – collectés à partir des quotidiens français suivants : *Le Figaro*, *Libération* et *Le Monde*. La période de collecte a été fixée à quatre ans, de juillet 2007 à décembre 2010 et l'intégralité des articles a été collectée à partir de la rubrique économique de ces quotidiens. À partir de cette série d'articles, nous avons, ensuite, procédé à la réalisation d'une base de données collectant les SPE : 1. Se référant à la crise des *subprimes* et aux manifestations économique-financières reliées à celle-ci, à savoir le processus de titrisation, la déréglementation et les effets directs et indirects de la crise (critère thématique) 2. Se présentant sous forme de nom ou de syntagme (critère formel) 3. Susceptibles de modifier la façon dont le lecteur perçoit la réalité (critère pragmatique).

### 3. Analyse des données

Le recours aux SPE répondant à ce que nous avons défini comme le *critère d'ambiguïté sémantique*<sup>6</sup> encourage, à notre avis, l'élaboration d'interprétations erronées du discours économique. Cela s'explique par le fait que plus le degré d'ambiguïté est élevé, majeure est la possibilité que les SPE réalisent leur PE en discours et engendrent, chez certains lecteurs – ceux pour lesquels le SPE devient euphémisme<sup>7</sup> – une variabilité interprétative incorrecte de l'information donnée. En revanche, l'interprétation sera correcte lorsque le lecteur sera en mesure de reconnaître un SPE en tant que faux-euphémisme<sup>8</sup>. En d'autres termes, l'ambiguïté sémantique caractérisant un SPE lui confère une valeur très connotée ; par conséquent, celui-ci sera susceptible de recevoir une interprétation erronée entraînant une variabilité interprétative incorrecte. Afin de mieux comprendre cette idée, nous pouvons observer l'exemple ci-dessous :

« Pour parvenir à ce miracle, le ministère a **optimisé la gestion du personnel**, jusqu'ici trop dispendieuse » (Libération, 16/04/2008)

A partir du sens dénoté, le récepteur pourra conférer au syntagme *optimisé la gestion du personnel*, un sens connoté résultant de l'identification correcte ou incorrecte de certains indices contextuels et pragmatiques. L'identification correcte des indices contextuels et pragmatiques caractérisant la situation énonciative permettra donc de connoter le SPE *optimisé la gestion du personnel* en tant que syntagme appartenant au langage « économiquement correct » et de l'opposer à des termes dénotativement équivalents tels que *licencié*, terme « neutre » utilisé dans la langue standard ou *dégraissé* et *rationalisé*, termes appartenant, eux aussi, au langage « économiquement correct ». Dans ce cas, le SPE ne pourra pas être considéré comme un euphémisme<sup>9</sup>; il s'agit pour nous ici d'un registre de langue différent du registre standard, d'un choix lexical préférant pour maintes raisons le syntagme *optimisé la gestion du personnel* au signe *licencié*, il s'agit tout simplement d'adapter des

<sup>6</sup> La particularité d'un signe/syntagme de posséder des traits sémantiques lui conférant le plus élevé degré d'ambiguïté possible.

<sup>7</sup> Un euphémisme est, pour nous, la réalisation en discours d'un *potentiel d'euphémisation* (PE) qui, à son tour, peut être considéré comme la possibilité intrinsèque à chaque signe de devenir euphémisme. En d'autres termes, on peut considérer l'euphémisme comme un SPE dont le *potentiel d'euphémisation* se réalise effectivement en discours (dans la *parole*).

<sup>8</sup> Le faux-euphémisme est, pour nous, un SPE dont le *potentiel d'euphémisation* ne se réalise pas en discours.

<sup>9</sup> De notre point de vue, l'euphémisme existe uniquement si sa force d'atténuation ou d'amélioration n'est pas perçue par le récepteur. La vraie nature de l'euphémisme ne résiderait, donc pas, comme conventionnellement affirmé, dans sa force méliorative ou d'atténuation, mais plutôt dans l'affectation de la fonction référentielle du langage, c'est-à-dire dans le remplacement d'un faux référent - auquel le signe substitut renvoie - à un vrai. Saisir l'allusion entre un SPE et le signe tabou qu'il remplace comporte une prise de conscience, de la part du récepteur, de la réalité taboue ; cette réalité est bien présente dans la conscience du locuteur et c'est pour cette raison que nous ne pouvons pas considérer qu'elle soit atténuée. L'euphémisme est donc pour nous une hypothèse interprétative erronée, suite à laquelle un récepteur, faute à une analyse non pertinente des éléments linguistiques et extralinguistiques qui caractérisent l'emploi d'un SPE dans une situation communicative donnée, ne saisit pas la relation entre un SPE et le signe tabou qu'il remplace, compromettant ainsi la fonction référentielle du langage et, par conséquent, sa perception de la réalité. C'est pour cela que lorsque le récepteur attribue au syntagme *optimisé la gestion du personnel* et au signe *licencié* la même signification (rupture du contrat de travail de la part de l'employeur) qui renvoie au même référent (perte de l'emploi), nous préférons parler de faux-euphémisme. Dans ce cas, le récepteur reconnaît le lien entre SPE et le signe tabou et il se trouve face à des signes/syntagmes de signification quasi identique pouvant coréférer en discours (synonymie approachante).

mots à la situation d'énonciation et à ses fins. Vraisemblablement, cela impliquera l'élaboration d'une variabilité interprétative correcte. Au contraire, l'identification incorrecte des indices contextuels et pragmatiques caractérisant la situation énonciative donnera lieu à une connotation inexacte du syntagme *optimisé la gestion du personnel* qui sera probablement considéré comme un technicisme appartenant au langage de l'entreprise. Ce syntagme trouvera donc son sens par opposition à des signes dénotativement synonymes tels qu'*amélioré*, *maximisé* ou *perfectionné*, par exemple ; ce qui compromettra la référence – le référent du syntagme *optimisé la gestion du personnel* ne correspond pas au référent des signes considérés erronément comme signes équivalents – et provoquera, par conséquent, une

connotation incorrecte<sup>10</sup>. Seulement dans ce deuxième cas, le *substitut potentiellement euphémique* acquerra la valeur d'euphémisme. Le SPE *optimisé la gestion du personnel*, qui remplace le signe tabou *licencié*, ne joue donc réellement le rôle d'euphémisme que si le récepteur ignore certaines informations supplémentaires qui vont s'ajouter à l'ensemble des traits sémantiques du syntagme *optimisé la gestion du personnel*<sup>11</sup>. Cela impliquera l'élaboration d'une variabilité interprétative incorrecte de l'information présentée.

Suite à ces considérations, nous nous proposerons d'analyser quelques exemples porteurs d'ambiguïté et donc susceptibles d'amener le lecteur à élaborer des variabilités interprétatives inexactes. Il s'agit de SPE tirés de notre corpus de référence et issus du recours, lors de l'activité de *vulgarisation* effectuée par le journaliste, à certains procédés de reformulation<sup>12</sup> marquant une feinte équivalence sémantique entre deux termes ou expressions.

- [1] « Il s'agit ni plus ni moins de scinder les fonds en deux nouveaux produits : l'un provenant des actifs toujours liquides et qui continue à fonctionner normalement, l'autre constitué d'actifs **illiquides** dont le portefeuille sera vendu au fur et à mesure des opportunités. » (Le Figaro, 20/10/2008)

En [1], le morphème /il/ se présente comme un simple procédé de préfixation négative qui oppose *liquides* à *non liquides*, sur le modèle de *légal/ illégal*, par exemple. Cependant, le SPE *illiquides*, contrairement au signe *illégal*, se caractérise par un changement de sens de la base à laquelle le préfixe est adjoint. Si donc, *illégal* signifie *non-légal* (*légal* et *illégal* sont antonymes), *illiquides* ne signifie pas, ici, *non-liquides*, mais *toxiques*. Ici, le syntagme *actifs non liquides* a été créé à partir de l'expression technique *actifs liquides*, relevant du domaine de la finance et désignant des actifs ayant « la faculté d'être mobilisés ou réalisés à vue » (Silem, 2012:509). En d'autres termes il s'agit de stocks ou de créances à court terme qui se

<sup>10</sup> Cela car le sens connoté change sans compromettre la référence.

<sup>11</sup> Dans une telle situation, le SPE est réellement euphémique car le syntagme *optimisé la gestion du personnel* et le signe *licencié* renvoient à un référent propre et, par conséquent, la fonction référentielle du langage est compromise. Telle situation se vérifie lorsque le récepteur ne perçoit pas la relation allusive entre le signe tabou et son *substitut potentiellement euphémique*. En d'autres mots, le SPE est un signe caractérisé par un signifiant dont le signifié est mis en attente ; si le récepteur, ne reconnaissant pas l'allusion, assigne au signifiant, un signifié qui renvoie à un référent différent du signe tabou, on parlera d'euphémisme. Dans ce cas, il aura une réception faussée de l'énoncé, compromise par son incapacité à établir une relation conforme entre signifiant, signifié et référent. Si, au contraire, le récepteur, reconnaissant la relation allusive entre le signe tabou et le SPE, assigne au signifiant, un signifié qui renvoie au même référent du signe tabou, on parlera de faux-euphémisme.

<sup>12</sup> 1. Procédés morpho-lexicaux tels que la préfixation qui permet essentiellement de masquer les changements de sens de la base d'un signe et la nominalisation qui se caractérise par une complexité sémantique par rapport à la construction verbale 2. Procédés sémantiques tels que l'hyponymie, l'homonymie et la polysémie 3. Procédés rhétoriques tels que la métaphore, qui permet de voiler la négativité du sujet tabou et la périphrase qui permet de contourner le tabou.

caractérisent par la capacité d'être échangés rapidement en monnaie. Logiquement, donc, parler d'*actifs illiquides* ou *non-liquide* pourrait amener le lecteur à croire que ce « technicisme » se réfère aux actifs financiers dont le degré de *liquidité* n'est pas à court terme. Cependant, considérant que les actifs en question sont définis en finance par l'expression *actifs immobilisés*, nous pouvons considérer ce syntagme comme un faux-technicisme<sup>13</sup>. Le recours à l'expression *actifs illiquides* ou *non-liquides* est donc susceptible d'engendrer deux variabilités interprétatives ; la première, incorrecte, correspondant à l'idée que l'inutilisation de ces biens s'explique par la logique de les amortir dans le long terme. La deuxième, correcte, considérant ces produits financiers comme des produits difficilement négociables, voire non-négociables issus de la *titrisation*<sup>14</sup> des crédits *subprimes*<sup>15</sup> ; en d'autres termes, comme des créances irrécouvrables ou bien invendables.

[2] « L'annonce de BNP Paribas montre que « **le problème des subprimes** » va prendre plus de temps qu'attendu avant de se dissiper. Il ne faudra pas quelques jours ou quelques semaines, mais des mois avant de connaître son impact total », a expliqué Owen Fitzpatrick, analyste de la Deutsche Bank. » (Le Figaro, 10/08/2007)

En [2], le recours à un hyperonyme, renvoyant à un référent générique (*problème*) à la place d'un hyponyme, renvoyant à un référent spécifique (*crise*) peut causer une certaine ambiguïté sémantique, due à une extension<sup>16</sup> élevée et une intension<sup>17</sup> réduite de l'hyperonyme par rapport à l'hyponyme. Ici, la variabilité interprétative dépend du fait que le SPE permet de jouer sur le degré de négativité que le locuteur peut inférer d'un énoncé. En effet, le signe *problème* n'a pas le même degré de négativité que le signe tabou *crise*, qu'il remplace.

[3] « Nous devons nous interroger sur le rôle exact que doivent jouer ces agences dans la cartographie des risques », a précisé le président français Nicolas Sarkozy. Dans la crise actuelle, leur position est plus fragile car elles ont aussi un rôle de conseil auprès des banques dans la création des **véhicules hybrides**. » (Le Figaro, 17/08/2007)

En [3], l'ambiguïté du faux-technicisme *véhicules hybrides* est telle que nous nous limiterons à proposer deux possibles interprétations sans, pour autant, trancher pour l'une des deux car, à vrai dire, les contextes linguistique et extralinguistique ne nous permettent pas de le faire. D'après nous, donc, le syntagme *véhicules hybrides* pourrait, d'une part, désigner un ensemble de produits financiers très risqués, du type des CDO<sup>18</sup>. Dans ce cas, la création lexicale de ce SPE serait, d'après nous, à rechercher dans l'association d'un synonyme du terme technique *instrument* (*véhicules*) et de l'adjectif *hybrides*, remplaçant

<sup>13</sup> Termes économiques adoptés avec une signification différente par rapport à leur signification dans la langue spécialisée ou bien termes désignant des notions étrangères à la langue économique, des notions qui n'existent pas en économie ou en finance.

<sup>14</sup> Transformation des crédits en titres financiers (Silem, 2012).

<sup>15</sup> Type de crédits hypothécaires (crédits immobiliers) très risqués, *i.e.* proposés à des ménages ne présentant pas les garanties financières normalement exigées pour un emprunt immobilier. (Silem, 2012).

<sup>16</sup> « Concept d'origine logique utilisé dans la description sémantique pour désigner l'ensemble des entités qui constituent la *dénotation* d'un terme (nom ou expression nominale) au niveau de la langue, ou bien l'ensemble auquel *réfère* un terme dans un univers de discours donné au niveau de l'énoncé. » (Arrivé, Gadet, & Galmiche, 2010).

<sup>17</sup> « Concept emprunté à la logique équivalent à celui de *compréhension*. L'intensité désigne l'ensemble des propriétés sémantiquement pertinentes (ou traits distinctifs) d'une unité lexicale, qui permettent de la distinguer de l'ensemble des autres unités appartenant à la même catégorie syntaxique (Arrivé, Gadet, & Galmiche, 2010).

vraisemblablement *dérivés*, terme associé à *instrument* dans le syntagme technique *instruments dérivés*. D'autre part, *véhicules hybrides* pourrait également remplacer le terme technique *véhicules d'investissement structuré* (SIV) qui désigne les « fonds d'investissement, inventé en 1988 par la banque Citigroup, placés hors bilan pour échapper aux normes prudentielles. » (Silem, 2012:838).

[4] « Le directeur financier, Fritz Henderson, a précisé que la **réduction** envisagée **des coûts de personnel** se ferait via des primes au départ et des plans de retraite anticipée. » (Le Monde, 15/07/2008)

En [4], malgré la substitution du signe tabou *chômage*, par le SPE périphrastique *réduction des coûts de personnel*, le locuteur reconnaît facilement l'idée taboue implicite. Ici, on parle de *réduction des coûts de personnel*, mais il serait plus correct de parler de réduction de personnel tout court ou mieux encore de licenciement. Dans ce cas, la variabilité, n'est donc pas issue de la valeur référentielle du SPE – cette dernière étant, à peu près, la même exprimée par l'idée tabou qu'elle remplace – mais du changement de perspective du problème présenté et des effets que ce changement implique sur le locuteur. En effet, si le terme licenciement sous-tend un problème pour les travailleurs, son SPE sous-tend un problème pour l'employeur.

#### 4. Conclusion

Suite à cette analyse, nous avons, d'une part, essayé de montrer de quelle manière le recours aux SPE pourrait déterminer des variabilités interprétatives, engendrées par la présence ou l'absence de certaines variables pragmatiques -dépendantes de l'environnement physique et de l'*environnement cognitif*<sup>19</sup> du récepteur - permettant au récepteur de reconnaître, dans une situation communicative donnée, un SPE en tant qu'euphémisme ou que faux-euphémisme. D'autre part, nous avons voulu souligner le danger du langage « économiquement correct » qui, par le biais de procédés euphémiques, prive le lecteur des éléments nécessaires pour interpréter l'information et l'empêche d'avoir une vision claire des événements qui se produisent dans notre quotidien. L'« économiquement correct » est donc, à notre avis, une source de variabilités interprétatives qui, paradoxalement, naissent de l'homogénéité du langage euphémique. En effet, à travers le recours aux SPE, le journaliste établit ce qui est acceptable pour le lecteur et ce qui ne l'est pas, ce qu'il est préférable de dire et ce qu'il vaut mieux omettre. De son côté, le lecteur concerné est sous-informé, il est privé des éléments nécessaires pour interpréter la réalité économique et il a uniquement accès à une information économique formatée, unique et standardisée, expression de l'absence de points de vue différents et pouvant entretenir une vision irréaliste de la réalité économique.

<sup>18</sup> Sigle pour *Credit Debt Obligation*. Titres de taux émis en contrepartie d'un portefeuille d'ABCP (Sigle pour *Assets Backed Commercial Papers*). Produit d'une opération de *titrisation* correspondant à des créances sécurisées par un portefeuille d'actifs (des crédits, des titres) émises par un véhicule *ad hoc* aussi appelé conduit si les titres proposés sont à court terme) ou d'ABS (Sigle pour *Asset Backed Securities*. Instruments financiers, produits de crédit purs peu risqués émis en contrepartie d'un portefeuille de créances, utilisant la technique de la *titrisation* ; traduction : « titres garantis par des actifs »). On peut aussi constituer des CDO de CDO (CDO *square*). (Silem, 2012).

<sup>19</sup> « Un environnement cognitif d'un individu est un ensemble de fait qui lui sont manifestes », « L'environnement cognitif d'un individu est un ensemble d'hypothèses à sa disposition » (Sperber & Wilson, 1989).

## Références

- ARRIVÉ M., GADET F., GALMICHE M. (2010). *La grammaire d'aujourd'hui. Guide alphabétique de linguistique française*. Paris : Flammarion.
- GARNIER Y., VINCIGUERRA M. (2006). *Le Petit Larousse illustré 2007*. Paris : Larousse.
- GREIMAS A.J. (2002, rééd.). *Sémantique structurale*. Paris : Presses universitaires de France.
- RASTIER F. (1994). *Sémantique pour l'analyse. De la linguistique à l'informatique*. Paris : Masson.
- SILEM A. (2012). *Lexique d'économie*. Paris : Editions Dalloz.
- SPERBER D., WILSON D. (1989). *La Pertinence. Communication et cognition*. Paris : Les Editions de Minuit.
- VILLAR C. (2005), *Pour une théorie du discours diplomatique*. « AFRI » 6. Bruxelles : Editions Bruylant.

## ***Ethos* des locuteurs dans des émissions économiques radiophoniques : variables et variations**

Nathalie Gerber<sup>1</sup>

(1) CREM, Université de Lorraine  
nathalie.gerber@gmail.com

**Résumé** Cet article a pour objet l'étude de l'*ethos* des locuteurs par l'analyse de marques énonciatives dans un corpus d'émissions économiques radiophoniques. Il explore les variables que sont le genre d'émission et le statut du locuteur dans les variations de l'*ethos*. Les marques énonciatives mesurées et analysées pour cette exploration sont les pronoms personnels, les questions et les exclamations. Les mesures et les analyses en cotexte posent le problème de la taille du corpus pour l'analyse de marques énonciatives. Au cours des mesures et des analyses, le genre d'émission et le statut du locuteur s'avèrent comprendre plusieurs variables.

**Abstract** This article is a study of the speakers' *ethos* through the analysis of enunciation markers in a corpus made of economic radio programs. Its aim is to explore the program's type and the speaker's status as variables in the variation of the speaker's *ethos*. The enunciation markers which are measured and analysed are personal pronouns, questions and exclamations. Measures and analyses in the text raise the issue of the length of the corpus for enunciation markers analysis. Through the measures and the analyses, the program's type and the speaker's status turn out to be several variables.

**Mots-clés :** analyse du discours, médias et langage, *ethos*, marques énonciatives, variations.

**Keywords:** Discourse analysis, media and language, *ethos*, enunciation markers, variations.

## 1 Introduction

Les analyses présentées dans cet article s'inscrivent dans un projet d'étude de l'*ethos* des locuteurs dans un corpus d'émissions économiques radiophoniques par l'analyse des marques énonciatives. Nous proposons une réflexion sur les variations de l'*ethos* en fonction du genre des émissions et du statut des locuteurs. Après un exposé des statuts et des genres représentés dans le corpus, l'analyse de quelques marques énonciatives mettra en lumière différentes variables ayant une incidence sur l'*ethos*.

## 2 Le corpus et ses variables

Le corpus étudié est composé de 30 émissions économiques<sup>1</sup> de France Inter correspondant à deux semaines de diffusion, du 30 mars au 5 avril 2009 (S1) et du 13 au 19 avril 2009 (S2), totalisant environ 2 heures et 15 minutes d'émission. Le tableau 1 présente succinctement les six émissions différentes que contient le corpus.<sup>2</sup>

Titre	Périodicité	Animateurs	Chroniqueurs	Invités	Genre
<i>Dimanche soir économique</i>	1 jour sur 7	BJ	-	S1 : ML et CB S2 : FS et PF	débat
<i>Le débat éco</i>	1 jour sur 7	ND ED	BM et DS	-	débat
<i>L'économie autrement</i>	1 jour sur 7	SLR	AB	-	chronique
<i>L'édito éco</i>	5 jours sur 7	ND ED	DS JMV	-	chronique
<i>L'autre économie</i>	4 jours sur 7	PM	MB	-	chronique
<i>L'éco du jour</i>	5 jours sur 7	ND ED	PL BG	-	brèves et analyse

Tableau 1 : Emissions du corpus

### 2.1 Le statut des locuteurs

Les locuteurs se distinguent par leur statut dans le dispositif médiatique. On trouve :

- des animateurs, qui sont des journalistes employés par l'institution médiatique de France Inter et ont la responsabilité du déroulement de l'émission ;
- des chroniqueurs, experts qui collaborent de façon régulière avec France Inter et ont la responsabilité de leurs propos ;
- des invités, experts ou acteurs publics qui sont extérieurs à l'institution médiatique et ont la responsabilité de leurs propos.

<sup>1</sup> Pour définir les émissions « économiques », nous avons retenu celles qui contiennent le mot « éco » ou « économique » et se désignent ainsi par le titre comme s'inscrivant dans le champ économique.

<sup>2</sup> Les intervenants apparaissent sous leurs initiales, en fonction de leur statut médiatique ; certains intervenants titulaires ont été remplacés pendant une semaine ou pour une émission ; les remplaçants apparaissent en italique dans le tableau.

Dans les émissions, certains chroniqueurs et les invités sont présentés succinctement par un ou deux éléments de leur statut social. Les chroniqueurs sont des économistes ; parmi eux, BM est présenté comme chroniqueur économique dans *Charlie Hebdo*, et DS et JMV comme chroniqueurs dans *Les Echos*. Les invités de *Dimanche soir économique* sont présentés ainsi :

- ML : directeur des Etudes économiques d'HSBC France ;
- CB : président du Conseil d'analyse économique et co-auteur avec Le Cercle des économistes de *Fin du monde ou sortie de crise* (éditions Perrin) ;
- FS : président de l'UIMM (Union des industries et métiers de la métallurgie), fédération patronale de la métallurgie ;
- PF : PDG du groupe ALPHA qui conseille les comités d'entreprise et travaille sur la revitalisation des bassins d'emploi.

## 2.2 Les genres médiatiques

Dans le corpus, quatre genres sont représentés : la brève, l'analyse, la chronique, le débat. P. Charaudeau (2005) propose de définir les genres informatifs médiatiques en croisant quatre paramètres :

- le type d'instance énonciative (origine du sujet parlant et degré d'implication) ;
- le type de mode discursif (événement rapporté, commenté ou provoqué) ;
- le type de contenu (macro-domaine sur lequel porte l'événement) ;
- le type de dispositif (matériaux organisés en système sémiologique, canal de transmission, possibles connexions entre les participants et leurs relations).

Les premier et deuxième paramètres sont regroupés dans un schéma (Charaudeau, 2005, p. 174) où sont placés différents genres médiatiques. Dans ce schéma, les quatre genres de notre corpus se lisent ainsi :

- brève : événement rapporté, issu d'une instance externe, peu engagée ;
- analyse : événement commenté, issu d'une instance interne, moyennement engagée ;
- chronique : événement commenté, issu d'une instance interne, très engagée ;
- débat : événement provoqué, organisé par l'instance interne, moyennement engagée.

Dans le cadre de l'analyse de l'*ethos*, nous traitons l'engagement à un niveau énonciatif, sous l'angle de la prise en charge du point de vue par le locuteur.

Le troisième paramètre proposé par P. Charaudeau est une constante pour notre corpus dans la mesure où toutes les émissions concernent le même macro-domaine : l'économie.

Concernant le quatrième paramètre, P. Charaudeau (2005) propose un descriptif général du dispositif de la radio mais ne décrit pas le dispositif d'une émission. Toutes les émissions de radio ont en commun le canal de transmission qu'est le son, qui met les voix au centre du dispositif. Mais les rapports entre les participants varient selon les émissions, et les discours n'ont pas toujours le même degré d'oralité. Dans le corpus, nous distinguons trois dispositifs :

- les émissions comprenant un animateur et un chroniqueur, dans lesquelles le chroniqueur oralise un texte préparé ; le discours est alors un monologue préparé (les brèves, certaines analyses et quelques chroniques) ;

- les émissions comprenant un animateur et un chroniqueur, dans lesquelles le chroniqueur oralise un texte préparé et fait intervenir l'animateur qui lit des répliques écrites par le chroniqueur ; le discours est alors un dialogue préparé (la plupart des chroniques et certaines analyses) ;
- les émissions comprenant trois locuteurs, dans lesquelles l'animateur invite deux intervenants (invités ou chroniqueurs) à confronter leurs opinions ; le discours prend alors la forme d'un dialogue spontané mais encadré (les débats).

Dans l'analyse des variations de l'*ethos*, nous verrons dans quelle mesure les genres médiatiques et le statut des locuteurs peuvent être considérés comme des variables.

### 3 L'analyse de l'*ethos* par les marques énonciatives

#### 3.1 L'*ethos*, un phénomène énonciatif dans une approche sociodiscursive

Dans la conception aristotélicienne, l'*ethos* de l'orateur est une preuve dans l'art de la persuasion. Il est présenté tantôt comme relevant de l'honnêteté de l'orateur et tantôt comme relevant de la capacité de l'orateur à être perçu comme compétent par des techniques oratoires. En analyse du discours, R. Amossy (2010) et D. Maingueneau (1999) exposent une conception énonciative de l'*ethos* et une approche sociodiscursive de la variation de l'*ethos*. L'*ethos* comme phénomène énonciatif et sociodiscursif a donné des études intéressantes sur des *ethè* particuliers ou des stratégies énonciatives liés à des genres discursifs<sup>3</sup>. Ces travaux s'attachent à découvrir des fonctionnements argumentatifs ou discursifs de marques énonciatives. Si ces études permettent d'avancer dans l'analyse de l'*ethos* par les marques énonciatives, il est important de replacer ces stratégies dans des scènes d'énonciation, dans une scénographie, pour percevoir l'*ethos* dans sa fonction de preuve dans le discours.

Si l'*ethos* comme preuve peut s'inscrire dans le non verbal et le paraverbal, s'il peut être inféré du choix des mots et des arguments, s'il repose sur le statut institutionnel du locuteur, son approche en sémantique de l'énonciation se fait par l'étude des marques énonciatives qui indiquent la prise en charge, le point de vue et la voix du locuteur. On peut rassembler sous l'étiquette « marques énonciatives » les déictiques, les axiologiques, les verbes parenthétiques, les verbes médiatifs, les adverbes d'énonciation, les interjections modales, les modalités verbales et les modalités d'énoncés, les adresses. Cet ensemble rend compte d'une acception étroite des marques énonciatives. Dans les études sur les marqueurs d'attitudes énonciatives ou dans la recherche des traces énonciatives<sup>4</sup> sont explorés aussi bien des connecteurs que la focalisation et l'ordre des mots. On revient très vite au constat que fait C. Kerbrat-Orecchioni (2002) : la subjectivité est omniprésente.

Dans une conception de la gestion collective des *ethè* (Amossy 2010, ch.5), nous considérons que l'*ethos* se construit dans le rapport entre le locuteur et ses interlocuteurs et co-énonciateurs, par collaboration, négociation ou antagonisme, aussi bien dans les interactions en face à face que dans la circulation des discours. Ce rapport s'inscrit dans les marques énonciatives dialogiques et dans l'actualisation et le point de vue du locuteur.

---

<sup>3</sup> Voir les articles de F. Grossmann et F. Rinck (p.34-50) et de A. Rabatel (p.18-33) dans *Langages* n°156, 2004, et l'article de L. Perrin dans les Actes du colloque international *Le français parlé et les médias*, Lausanne 2010.

<sup>4</sup> Voir *Langue française*, 2009, n°161 pour les marqueurs d'attitude énonciative, et *Le regard du locuteur* (Paris : Kimé, 1993) et *Le regard du locuteur 2* (Paris : Kimé, 2001) de H. Nølke pour les traces énonciatives.

### 3.2 Quelles marques pour quelles dimensions de l'*ethos* ?

Pour passer de la conception théorique de l'*ethos* à son analyse, nous proposons de considérer l'*ethos* comme la synthèse de traits énonciatifs, ou traits éthiques, issus de l'interprétation des marques énonciatives (au sens large) présentes dans un discours ou un ensemble de discours. Etant donnée l'hétérogénéité catégorielle et sémantique des marques énonciatives, nous proposons de les rassembler en trois catégories correspondant à trois dimensions de l'*ethos* :

- les marques d'actualisation du locuteur, qui peuvent indiquer l'identité du locuteur et convoquer son *ethos* préalable ; elles contribuent à la dimension identitaire de l'*ethos* ;
- les marques monologiques, qui indiquent la prise en charge des énoncés par le locuteur et les modes d'articulation de ces énoncés ; elles contribuent à la dimension subjective de l'*ethos* ;
- les marques interlocutives et dialogiques, qui indiquent le rapport entre le locuteur et ses interlocuteurs et co-énonciateurs ; elles contribuent à la dimension intersubjective de l'*ethos*.

Dans la présente analyse, seules quelques marques sont étudiées : les déictiques de première et de deuxième personnes, les questions et les exclamations marquées. Cette sélection répond à plusieurs critères : la quantité des occurrences permettant des mesures comparatives entre les locuteurs, les écarts saillants d'usage de certaines marques entre les locuteurs, le coût de présentation théorique, l'exploration des trois catégories de marques et enfin l'apport de leurs analyses à la problématique des variables dans les variations de l'*ethos*.

## 4 Variables dans les variations de l'*ethos*

### 4.1 Mesures et analyses de marques d'actualisation du locuteur

Les déictiques de première personne sont des marques explicites d'inscription du locuteur dans son discours. L'absence de ces marques indique une forme d'effacement énonciatif. La présence de déictiques de première personne du singulier marque une inscription individualisée du locuteur comme en 1., tandis que les *nous*, *nos*, *notre* ou suffixe *-ons* de l'impératif marquent une inscription du locuteur dans un collectif : les déictiques de première personne du pluriel réfèrent aux Français dans 2(a), à l'UIMM dans 2(b) et aux interlocuteurs dans 2(c). On distingue deux types de *nous*, le *nous*-inclusif qui inclut dans le collectif le ou les allocutaire(s), et le *nous*-exclusif qui les en exclut : *nous* est exclusif dans 2(b), *regardons* est inclusif dans 2(c), *nous* est inclusif dans 2(a) pour les allocutaires français de l'émission.

1. « **moi-** le chiffre/ qui **m'**a frappé/ c'est celui\ de- la prévision\ de déficit public\ en ↑ Grande-Bretagne/ » (28/48/DS)<sup>5</sup>
2. (a) « c'est-à-dire/ que- **nous** avons beaucoup d'argent en France\ » (28/101/BM)  
(b) « parce que\ **nous** avons la chance d'avoir un réseau qui connaît les entreprises/ et les besoins des entreprises/ » (30/92/FS)

<sup>5</sup> Convention de transcription :

/ - \ : contours mélodiques continuatifs ascendant, plat ou descendant ; // \ : contours mélodiques conclusifs ascendant ou descendant ; ↓ ↑ : descente ou montée de registre ; () : troncation ; **souligné** : chevauchement ; (01/02/AB) : 01 est le numéro du document, 02 est le numéro de la séquence, AB sont les initiales du locuteur.

(c) « alors\ regardons l'avenir// » (02/44/ND)

A ces marques personnelles, nous associons l'étude du pronom indéfini *on* qui a valeur parfois de *je* ou de *nous*, 3(a)-(b), mais aussi de *ils*, 3(c), de *tout le monde*, 3(d), ou de présentatif, 3(e). A la suite de K. Fløttum et al. (2007, p.31), nous distinguons un *on*-inclusif, qui inclut le locuteur dans l'ensemble visé, d'un *on*-exclusif<sup>6</sup> qui l'en exclut. *On* est inclusif dans 3(a), (b) et (d), exclusif dans 3(c) ; *on* étant impersonnel dans 3(e) selon la classification de Fløttum et al. (2007 p.29-30), il n'est considéré ni comme inclusif, ni comme exclusif. Dans l'analyse de l'actualisation du locuteur, seul le *on*-inclusif nous intéresse. Même inclusif, le *on* reste une marque indéfinie qui permet de focaliser l'information sur le prédicat et non sur le sujet ; il est une marque d'effacement énonciatif.

3. (a) « **on** commence\ franchement\ à en douter// » (03/19/PL)
- (b) « et là/ **on** peut parler du rôle du FMI\ » (15/103/ML)
- (c) « qu'**on** aurait dû/ créer/ une sorte de *bad* banque/ mondiale// » (15/31/BJ)
- (d) « **on** sait même/ que Toyota va mal/ » (02/23/DS)
- (e) « là- **on** est dans le la famille/ des documentaires/ plaidoyers/ comme *le cauchemar de Darwin*/ comme *we feed the world*\ » (12/66/AB)

Pour mesurer et analyser ces marques d'actualisation du locuteur, les occurrences sont comptabilisées par locuteur et par type d'émission<sup>7</sup>. Les nombres absolus d'occurrences sont convertis en un taux représentant la fréquence d'occurrences à la minute, dans la durée de prise de parole de chaque locuteur, afin de pouvoir comparer les emplois des marques d'un locuteur à l'autre en gommant les différences de durée de prise de parole<sup>8</sup>. Les fréquences sont ensuite regroupées par genre d'émission et type de locuteur. Le tableau 2 rend compte de ces regroupements : les couples de chiffres inscrits dans les cases sont la fréquence la plus basse et la fréquence la plus haute du regroupement de locuteurs, ce qui laisse voir l'amplitude de la variation de la fréquence d'une marque dans les regroupements susmentionnés.

Genre/statut	Locuteurs	1 <sup>ère</sup> pers sing	1 <sup>ère</sup> pers pl in	1 <sup>ère</sup> pers pl ex	on-inclusif
Brèves	PL - BG	0 – 0,22	0 – 0,22	0	0 – 0,27
Analyses	PL - BG	0,28 – 0,31	0,10 – 0,14	0 – 0,20	0 – 1,65
Chroniques/ Chroniqueurs	BM-DS- JMV	0,40 – 0,61	0,17- 1,20	0,11 – 0,40	0,40 – 1,09
Débats/ Chroniqueurs -Invités	BM-DS- ML-CB-PF	2,35 – 4,98	0,18 – 0,62	0 – 0,55	0,92 – 4,37
	FS	0,88	0	2,81	1,58
Chroniques/ Animateurs	ND-ED	0 – 0,82	0,37 – 0,41	0	0,41 – 1,47
	PM	2,75	0	0	0
Débats/ Animateurs	ND-ED-BJ	1,13 – 2,28	0 – 0,48	0 – 0,48	0,81 – 4,56

Tableau 2 : Marques d'actualisation du locuteur

<sup>6</sup> Les termes inclusif/exclusif relatifs à *on* concernent l'inclusion et l'exclusion du locuteur, à la différence de ceux relatifs à *nous* qui concernent l'inclusion et l'exclusion de l'allocataire dans l'ensemble visé.

<sup>7</sup> Pour DS, BM, ND et ED, les occurrences sont comptées séparément selon qu'elles apparaissent en chronique ou en débat ; pour PL et BG, la distinction est faite entre leurs interventions en brève et en analyse.

<sup>8</sup> Les données de SLR ne sont pas intégrées dans les tableaux 2 et 3 car sa durée de prise de parole n'est que de 18 secondes. Celle des autres locuteurs varie entre 1 minute et 19 secondes et 22 minutes et 55 secondes.

On peut alors procéder à l'analyse comparative des fréquences. La troisième colonne montre une régularité dans la gradation de la présence de marques de première personne du singulier entre les brèves, les analyses, les chroniques et les débats avec un phénomène de seuil. Parmi les débatteurs, le locuteur FS est une exception puisqu'en débat il n'emploie ces marques qu'à 0,88 par minute, dans des formules dialogiques d'assentiment et dans un cas de formule performative. Du côté des animateurs, PM a en chronique une fréquence qui dépasse bien celle des autres animateurs en débat. Nous remarquons déjà que PM n'emploie ni *nous-inclusif*, ni *nous-exclusif*, ni *on-inclusif* ; elle ne s'inscrit que dans des déictiques individuels. Cette particularité ne semble s'expliquer ni par son statut, ni par le genre de l'émission.

Dans les quatrième et cinquième colonnes concernant respectivement les marques de *nous-inclusif* et de *nous-exclusif*, il est difficile de discerner des constantes ou des variations. On note que FS n'utilise pas de *nous-inclusif* ; il n'utilise que des *nous-exclusifs*, à une fréquence bien plus élevée que les autres locuteurs ; le *nous-exclusif* est le déictique le plus utilisé par FS et il renvoie, comme en 2(b), systématiquement à l'UIMM dont FS est le président.

Dans la sixième colonne concernant le *on-inclusif*, on distingue une gradation entre les brèves, les analyses et chroniques et les débats sans phénomène de seuil. On note une grande différence de fréquences entre les locuteurs d'un même genre et même statut. L'écart entre les fréquences les plus élevées en chronique d'une part et en débat d'autre part va de 1,09 à 4,37 pour les chroniqueurs et invités, et de 1,47 à 4,56 pour les animateurs. Cet écart peut s'expliquer par la différence de degré d'oralité entre les chroniques, qui sont de l'écrit oralisé, et les débats, qui sont des échanges spontanés cadrés.

Dans l'emploi du *on-inclusif*, FS se distingue encore : mis à part deux *on-inclusifs* inscrits dans une formule médiative comme en 4(b), toutes les occurrences renvoient à l'UIMM associée aux partenaires sociaux comme en 4(a).

4. (a) « c'que **nous** sommes en train de faire/ avec\ l'ensemble des partenaires sociaux/ c'est effectivement/ essayer/ d'imaginer\ dans notre branche/ comment **on** peut\ utiliser tous les moyens\ qui vont être prévus par euh par la loi/ ou par les conventions diverses/ comment utiliser ces moyens/ au niveau du territoire\ » (30/90/FS)
- (b) « si **on** regarde\ le pacte automobile/ » (30/106/FS)

La spécificité de l'actualisation de FS peut s'expliquer par son statut social. Il est le seul dans le corpus à être un acteur économique représentant les intérêts d'un groupe. Les autres invités sont des économistes (ML et CB) et un conseiller des comités d'entreprise (PF). L'actualisation du conseiller se rapproche de celle des économistes par la position extérieure d'analyste, commune aux conseillers et aux économistes. On voit ainsi apparaître une variation d'*ethos* entre les experts (conseillers et économistes) et les acteurs économiques.

## 4.2 Mesures et analyses de marques interactionnelles

Pour aborder la dimension intersubjective de l'*ethos*, nous avons mesuré les occurrences des marques de deuxième personne du pluriel et des structures interrogatives. Dans le tableau 3, on observe dans la troisième colonne une gradation de la fréquence d'emploi des marques de la deuxième personne du pluriel entre brèves, analyses, chroniques et débats. Il y a une différence notable de fréquence en fonction du statut médiatique du locuteur : les invités n'en emploient aucune, les animateurs les utilisent à une fréquence bien supérieure à celle des chroniqueurs, notamment en débat. Sur 98 occurrences, 51 réfèrent à un interlocuteur spécifique comme en 5(a), 28 peuvent être interprétées comme référant à l'ensemble des

interlocuteurs comme en 5(b), 19 sont des *vous*-génériques comme en 5(c). Les animateurs ne font usage, à l'exception de deux *vous*-génériques, que de *vous* référant à un interlocuteur spécifique. Chez les chroniqueurs, un quart des *vous* réfère à un interlocuteur spécifique, la moitié réfère à l'ensemble des interlocuteurs, le dernier quart étant des *vous*-génériques.

5. (a) « à cours terme/ Dominique Seux/ votre analyse// » (13/291/ND)  
 (b) « ↓je **vous** rappelle qu'entre 2000 et 2007/ le prix/ d'un même logement/ a quand même\ doublé// » (29/25/BG)  
 (c) « en fait/ en France/ **vous** faites partie des ↑10% les moins riches/ si- célibataire/ **vous** touchez/ moins de ↑748 euros après impôts et prestations sociales/ ou si/ en couple/ avec deux enfants/ **vous** ne gagnez pas plus/ de- ↑1983/ euros\\ » (24/33/BG)

Dans le dispositif radiophonique, le *vous* qui ne réfère pas à un interlocuteur spécifique peut produire un effet sur l'auditeur. Bien que ne faisant pas partie des participants ratifiés dans l'interlocution, l'auditeur est le destinataire du message. Dans cette double adresse, le *vous* non spécifiquement adressé tend à brouiller la frontière entre le cadre interlocutif du face à face et le cadre médiatique incluant les auditeurs. L'auditeur peut se sentir interpellé dans ce *vous* non spécifiquement adressé. Les animateurs semblent alors se limiter au cadre interlocutif tandis que les chroniqueurs semblent laisser porter leurs adresses au-delà de ce cadre pour atteindre les auditeurs. Les invités, faisant pourtant partie d'un dispositif fortement interlocutif qu'est le débat, n'actualisent pas leurs interlocuteurs dans leur discours.

Genre/statut	Locuteurs	2 <sup>e</sup> pers pl alloc	interrogatives
Brèves	PL - BG	0 – 0,22	0 – 0,27
Analyses	PL - BG	0 – 0,31	0,51 – 0,55
Chroniques/ Chroniqueurs	BM-DS- JMV	0,35 – 0,80	0,39 – 0,44
Débats/ Chroniqueurs	ML-CB- PF-FS	0	0,18 – 0,98
-Invités	BM-DS	0,73 – 0,94	0,62 – 1,61
Chroniques/ Animateurs	ND-ED- PM	0,41 – 1,65	3,31 – 6,06
Débats/ Animateurs	ND-ED- BJ	2,28 – 4,72	4,25 – 12,15

Tableau 3 : Marques interactionnelles

Dans la quatrième colonne du tableau 3 ont été mesurées les structures interrogatives. Nous reprenons à C. Kerbrat-Orecchioni (1991, p.12) la distinction faite entre la *question*, qui est un acte de langage, et la *structure interrogative*, qui est une structure formelle ; une question peut être posée soit directement par un énoncé de structure interrogative, soit indirectement, avec une assertion à valeur illocutoire de question, par exemple (Kerbrat-Orecchioni, 1991, p.20). Pour la présente analyse, seules les questions directes sont étudiées. On observe une légère gradation de la fréquence d'occurrences de structures interrogatives en fonction du genre, mais l'ordre de la gradation est quelque peu différent : la fréquence est un peu plus élevée en analyse qu'en chronique. Les structures interrogatives sont bien plus fréquentes chez les animateurs que chez les chroniqueurs et invités. En analysant ces occurrences dans leur dimension d'acte de langage, on s'aperçoit que certaines sont adressées à un interlocuteur, nominalement comme en 6(b), ou sans mention du nom comme en 6(a), provoquant la réponse du locuteur adressé ; d'autres questions, après avoir été posées, sont suivies directement d'une réponse formulée par le locuteur qui a posé la question, comme en 6(c) ;

d'autres encore sont posées sans que le locuteur semble n'attendre de réponse comme en 6(d). Nous proposons de faire une distinction entre les *questions adressées* de celles qui ne le sont pas, que nous appelons *questionnements* ; la visée des *questionnements* n'est pas tant d'obtenir une réponse d'un tiers que de poser la bonne question, ce qui apparaît explicitement en 6(e) qui montre le souci de poser « la vraie question ».

6. (a) « est-ce que le décor/ du sommet se précise// » (05/06/ND)
- (b) « et du côté de la Chine/ Bernard // » (01/23/PM)
- (c) « l'A400M/ l'avion militaire transporteur de troupes/ d'Airbus/ verra-t-il vraiment le jour // on commence\ franchement\ à en douter// » (03/18-19/PL)
- (d) « n'aura-t-il pas été le dindon du CAC 40\\ » (19/29/BM)
- (e) « **la vraie question**/ c'est- est-ce qu'on a du ↑travail\ qui fait de la production/ et est-ce que cette production/ vient du travail\\ » (28/121-123/BM)

Dans le corpus, on constate que toutes les structures interrogatives se trouvant dans les brèves, les analyses et les discours des invités sont des questionnements ; trois quarts des structures interrogatives de BM et DS en chronique et en débat sont des questionnements, les autres étant des questions adressées ; toutes les structures interrogatives énoncées par les animateurs sont des questions adressées, à l'exception de deux questionnements énoncés dans une introduction où BJ fait une analyse de la situation économique servant de point de départ au débat. Parmi les invités, seuls CB et ML, économistes, énoncent des questionnements ; FS et PF ne posent ni question adressée ni questionnement. Ces constats permettent d'entrevoir des traits éthiques des locuteurs : les animateurs sont des *questionneurs*, face aux chroniqueurs et invités qui sont les *questionnés* ; les chroniqueurs et invités économistes sont aussi, si j'ose le néologisme, des *questionnants* (locuteurs énonçant les questionnements). Seuls DS et BM adoptent les trois postures : ils sont questionnés, adressent des questions et énoncent des questionnements. Les invités FS et PF n'endossent que le rôle de questionnés. Les questionnements, s'ils ne sont pas adressés spécifiquement à un interlocuteur, ne sont pas énoncés sans destinataire. Si le locuteur énonçant un questionnement n'attend pas de réponse immédiate, le questionnement appelle des réponses dans la circulation des discours. A travers l'analyse des questions directes, les animateurs apparaissent encore dans une posture privilégiant le cadre interlocutif ; les économistes semblent viser un cadre plus large, une dimension dialogique. Les invités PF et FS, en ne posant aucune question, semblent en permanence être en position réactive dans les interactions.

### 4.3 Mesures et analyses d'une marque monologique

Comme marque monologique, nous proposons d'analyser les exclamations comprenant un marqueur exclamatif de nature morphologique ou syntaxique ; ces exclamations marquées, selon J. Bacha (2000, p. 254), comportent « une double information » : « le haut degré » et « le sentiment ressenti par le locuteur à propos du contenu propositionnel (joie, colère, étonnement, admiration...) ». <sup>9</sup> En analysant les exclamations marquées, on étudie l'usage d'un moyen d'expressivité du haut degré et du sentiment du locuteur.

On ne dénombre que onze exclamations marquées dans le corpus. Ce nombre est trop faible pour effectuer des mesures de fréquence. Mais la répartition de ces exclamations entre les locuteurs est notable : huit sont énoncées par BM en chronique ou en débat, comme en 7(a), une est énoncée par un animateur au moment des salutations en début d'émission en 7(b), et

<sup>9</sup> Les exclamations sans marqueur, accompagnées d'une intonation exclamative qui exprime le sentiment du locuteur, n'expriment pas le haut degré : « Il est venu ! » ; exemple repris de la note 181 (Bacha, 2000).

enfin deux sont énoncées par PF et PL ; mais ces deux dernières exclamations enchaînent sur une concession en 7(c) ou une restriction qui minorent la portée du haut degré. L'usage d'exclamations marquées, sans concession ni restriction, semble une particularité de BM dans ce corpus. Elle peut être considérée comme un stylème de BM, à mettre en rapport avec son statut social de chroniqueur à *Charlie Hebdo*, hebdomadaire satirique. A travers ces exclamations, l'*ethos* de BM se démarque des autres locuteurs dans son expressivité.

7. (a) « le petit porteur\ comme il a été choyé\ sollicité\ par tous les gouvernements\ » (19/05/BM)
- (b) « quelle collégialité// » (28/08/ED)
- (c) « excellente initiative/ **mais** insuffisamment de moyens/ aussi/ dans l'accompagnement des personnes/ qui sont concernées par le RSA\ » (30/82/PF)

## 5 Conclusion : analyser, mesurer et contextualiser

L'analyse de l'*ethos* à partir des déictiques personnels, des questions et des exclamations, dans un corpus économique radiophonique, a révélé que la variable genre renferme une sous variable : celle du degré d'oralité des genres. De même, l'analyse a révélé que la variable statut du locuteur comprend plusieurs variables : celle du statut de dispositif (animateur-chroniqueur-invité), celle du statut dans le champ économique (expert-acteur). S'ajoute enfin comme variable le style propre à chaque locuteur ; ce style est toutefois marqué par une gestion de l'*ethos* en cohérence avec les institutions que le locuteur représente. Les mesures et les analyses, pour être pertinentes, demandent d'un côté un corpus large pour obtenir des fréquences valides, et de l'autre, un corpus restreint pour rendre possible des retours au cotexte afin de conserver de la finesse dans les analyses. Si certaines marques énonciatives apparaissent en grand nombre dans le corpus, d'autres marques sont plus rares mais ne sont pas moins des indices de l'*ethos* du locuteur.

## Références

- AMOSSY R. (2010). *La présentation de soi. Ethos et identité verbale*. Paris : Presses Universitaires de France.
- BACHA J. (2000). *L'exclamation : approche syntaxique et sémantique d'une modalité énonciative*. Paris : L'Harmattan.
- CHARAUDEAU P. (2005). *Les médias et l'information : l'impossible transparence du discours*. Bruxelles : De Boeck.
- FLØTTUM K., JONASSON K., NOREN C. (2007). *On : pronom à facettes*. Bruxelles : De Boeck.
- KERBRAT-ORECCHIONI C. (2002 [1980]). *L'énonciation : de la subjectivité dans le langage*, 4<sup>e</sup> éd. Paris : Armand Collin.
- KERBRAT-ORECCHIONI C. (1991). *La question*. Lyon : Presses universitaire de Lyon.
- MAINGUENEAU D. (1999). Ethos, scénographie, incorporation. In Amossy R., *Images de soi dans le discours. La construction de l'ethos*, Paris : Delachaux et Niestlé, 75-100.

## **Modélisation du référencement en langue des signes : Analyse des variantes gestuelles**

Monia Ben Mlouka<sup>1</sup>

(1) IRIT - Université de Toulouse

mlouka@irit.fr

**Résumé** L'étude de la langue des signes nécessite la prise en considération des variantes de production et de perception d'un discours. La langue des signes conjuguant plusieurs sources gestuelles à la fois, le résultat d'intégration de gestes manuels et non-manuels produit des structures linguistiques dont la modélisation est complexe. Dans cette étude, nous nous sommes focalisés sur la structure linguistique de référencement. Le référencement est le marquage d'une zone de l'espace rattachée au discours en mettant en jeu des gestes manuels et/ou non manuels. Nous avons analysé les formes de marquage spatial en vue de construire un modèle qui tienne compte de ses variantes gestuelles. Notre méthode se compose d'analyses temporelles et spatiales des formes de pointage interagissant avec le contenu de l'espace de signation.

**Abstract** The analysis of sign language needs a careful consideration of its variations at enunciation and interpretation levels. Sign language merges many linguistic information sources. The fusion of non-manual and hand gestures results in complex linguistic structures. In this study, we focused on the linguistic structure of the act of reference. Reference in sign language means pointing to discourse's elements by involving manual and / or non-manual gestures. We analyzed reference gesture combinations and represented the concept by means of geometric shapes that takes into account the content of the signing space.

**Mots-clés :** Langue des signes, référencement, geste de pointage, variabilité, apprentissage, modélisation.

**Keywords:** Sign languages, Reference, pointing gesture, variability, learning, modeling.

# 1 Enjeux de l'analyse de la langue des signes

L'analyse automatique de discours en langue des signes a pour objectifs de valider, apporter des précisions et enrichir les modèles établis par les linguistes. Elle se compose de représentations informatiques de composants corporels et des techniques qui permettent de combiner ces représentations du fait de l'aspect multi-linéaire de la langue des signes pour pouvoir traiter automatiquement les événements linguistiques tel que le référencement. Outre les variations gestuelles dans la réalisation de signes manuels, l'aspect multi-linéaire vient rajouter de la complexité dans les tâches que doit gérer un programme informatique pour mener à bien l'analyse d'un énoncé en langue des signes. D'un autre côté, les outils informatiques ne sont pas à un niveau de précision qui permet de détecter, dans une vidéo, d'une manière précise certains gestes comme la direction du regard ou les clignements des yeux (Leroy, 2010). Dans ce travail, nous nous proposons de mettre en place un modèle qui permet la détection du concept de référencement, qui gère ses variantes gestuelles, son aspect multi-linéaire ainsi que l'absence d'information du regard dans certains cas.

## 2 La variation en langue des signes

La langue des signes est une langue gestuelle. Le signeur réalise des signes par les mains, des mouvements corporels comme la rotation de la tête et le balancement du buste. Nous considérons les signes réalisés par un seul composant des signes "mono-sources" et ceux effectués par plusieurs composants corporels des signes "multi-sources".

Dans un premier temps, nous exposerons les modèles proposés pour décrire les propriétés d'un signe "mono-source". Dans un deuxième temps, nous décrirons des exemples de modèles de signes "multi-sources" et mettrons en avant la contribution des gestes non-manuels dans la construction de ce type de signes.

### 2.1 Le signe "mono-source"

En langue des signes américaines, (Stokoe, 1960) identifie quatre unités caractéristiques du signe : la configuration, l'emplacement, l'orientation et le mouvement. L'exemple (1 b)) illustre les caractéristiques manuelles du signe [AVION] : 1)une configuration en [Y], 2)une orientation vers le bas, 3)un emplacement de départ devant le buste du signeur et un mouvement curviligne. En langue des signes française, (Cuxac, 2000, ch.4) définit un signe comme étant un ensemble de morphèmes c'est-à-dire des unités porteuses de sens. Dans l'exemple (1 a), on voit que la configuration du signe [AVION] se compose de deux unités qui sont le pouce et l'auriculaire qui représentent les ailes.

### 2.2 Le signe "multi-source"

(Liddell, 1980) souligne qu'un geste non manuel peut contribuer à la construction du sens, parallèlement ou en décalage avec les mains. L'exemple (2) illustre un [AVION] qui tourne auquel s'ajoute, simultanément, une expression faciale particulière et un mouvement de rotation de la

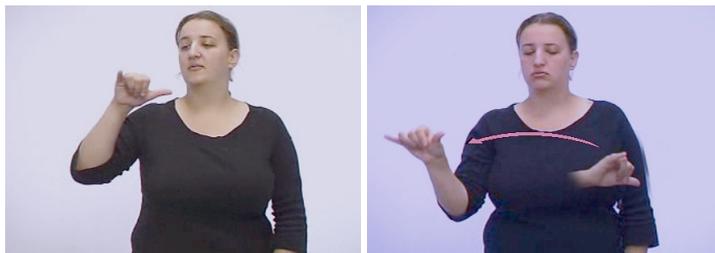


FIGURE 1 – a) Signe [AVION], b) Signe [AVION] en mouvement

tête. D'après le contexte, on comprend que le déploiement simultané d'une mimique faciale particulière et d'une rotation de la tête signifient un détournement d'avion.

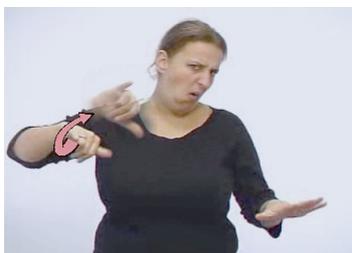


FIGURE 2 – Signe [AVION] en mouvement accompagné d'expression faciale particulière et d'une rotation de la tête

Un signe non-manuel est réalisé par un mouvement de la tête, un changement de la direction du regard, une mimique faciale particulière, des mouvements des épaules ou un balancement du buste. Une fixation par le regard peut être courte ou durer quelques secondes. Un geste de la tête peut être brusque ou lent. Un balancement du buste peut avoir une amplitude faible ou importante. Ainsi, un signe non-manuel peut varier dans la durée, la dynamique et l'amplitude du geste. De plus, dans le cas d'un référencement, la direction d'un mouvement non-manuel dépend du contenu l'espace de signation.

### 2.3 L'espace de signation

Le signeur utilise l'espace pour réaliser des signes par les mains, des mouvements corporels comme la rotation de la tête et le balancement du buste. Certains linguistes se sont penchés sur l'importance de l'espace car il supporte les éléments (*locus*<sup>1</sup>) qui interviennent dans le discours ainsi que leurs relations spatiales (Emmorey & Reilly, 1995). De plus, l'espace de signation représente les relations sémantiques et topologiques qui lient les entités rattachées au discours de manière structurée et pertinente (Emmorey & Reilly, 1995). (Rinfret, 2009) souligne que le déploiement de mouvements non-manuels sont, dans certains cas, motivés par la position de *locus* dans des l'espace de signation. Un mouvement non-manuel permet de réaliser le lien entre les entités du discours telles qu'une zone de l'espace et son référent (le signe qui lui est associé) ou encore mettre le *Focus* sur une zone particulière de l'espace dans le cas de référencement.

1. Zone de l'espace de signation rattachée à l'énoncé

## 2.4 Synthèse

Ainsi, nous pouvons extraire des paramètres mesurables pour caractériser un signe "multi-sources" tels que la position de début et de fin d'un composant réalisant un mouvement non-manuel, le délai entre les mouvements réalisés par des composants différents ainsi que la zone de l'espace vers laquelle se dirige du mouvement. Ceci permettra, dans une première étape de l'analyse, de quantifier un événement de référencement selon que les gestes correspondants soient réalisés par la main, par la tête et/ou par le regard. Une représentation de ces composants corporels est, ainsi, nécessaire pour concrétiser cette première étape. Nous illustrerons par des exemples la manière dont nous avons représenté la tête, les mains et le regard et comment on les a combinés pour représenter un événement de référencement vers une zone de l'espace de signation. L'un des objectifs de l'analyse du concept de référencement est d'apporter des éléments de réponse aux interrogations suivantes : 1) Quelles sont les combinaisons gestuelles associées au concept de référencement ?, 2) comment varient les combinaisons gestuelles dans le temps et dans l'espace ?

Les variations dans les modalités de réalisation de signes (manuels et non-manuels) d'une part et le fait qu'un énoncé est compris d'une manière non ambiguë d'autre part nous amènent à s'interroger sur le processus de compréhension élaboré par l'interlocuteur pour construire le sens du discours.

## 3 Le référencement

Un référencement est la manière dont le signeur met le "focus" sur un élément du discours localisé dans l'espace. (Cuxac, 2000) précise que le référencement se réalise grâce à des directions particulières du regard, de la tête ou de la main. Le geste de pointage manuel avec index tendu représente la variante la plus fréquente en langue des signes. En effet, Dans ses travaux sur les langues des signes émergentes, (Fusellier-Souza, 2004) a mené une étude sur le geste de pointage manuel qui s'est concrétisée par un taux de fréquence moyen de 24.85% et par neuf catégories fonctionnelles dont les pointages désignant une référence spatiale.

D'autre part, (Cuxac, 2000) attribue au regard la fonction de marqueur de zones occupées dans l'espace de signation. De leur part, (Parisot & Rinfret, 2006) associent à la direction du regard le rôle de pointeur spatial et émettent l'hypothèse que l'orientation du buste (tronc) ainsi que celle de la tête sont impliquées dans le marquage de locus. Il existe des signes "multi-sources" incluant la main dominante, le regard, la tête, les épaules et le buste. L'exemple (3 a) illustre un référencement "multi-sources" où le regard et la tête sont orientés vers une zone localisée par la main droite dans sa position d'arrivée. L'exemple (3 b) illustre une posture particulière du corps orientée vers la zone pointée par la main, le regard et la tête.

Bien qu'on puisse donner une définition unique et représentative au concept de référencement, les corpus en langue des signes prouvent qu'on peut associer à ce concept plusieurs combinaisons gestuelles qui varient sur les plans temporel et spatial. À titre d'exemple, pour une même configuration manuelle, "un pointage" varie en directions<sup>2</sup> et en profils de vitesses (Ben Mlouka & Dalle, 2011).

Nous nous sommes intéressés au référencement réalisé par la tête, la main, le buste et/ou le regard. Nous l'avons étiqueté pointage manuel (Man) et pointage non manuel (NM). Dans la

---

2. les directions d'un pointage sont liées à l'emplacement de locus



FIGURE 3 – Un pointage manuel avec : a) main plate, b) index tendu

section suivante, nous décrirons les corpus sur lesquels nous avons effectué notre analyse.

## 4 Corpus et annotations

Le corpus faisant l'objet de représentations informatiques a été enregistré dans le cadre du partenariat franco-québécois (Marqspat)<sup>3</sup>. Les sessions de capture ont été réalisées avec un système de capture de mouvement (VICON)<sup>4</sup>, une caméra vidéo pour filmer le cadre complet de la scène et un système de capture du regard (FaceLab)<sup>5</sup>. Le déroulement des sessions d'enregistrement a été détaillé dans (Ben Mlouka, 2012). Les données de capture de mouvement et celles de suivi du regard ont fait l'objet d'une analyse des mouvements de référencement manuels dans (Ben Mlouka, 2012). Nous décrivons dans le paragraphe suivant la représentation informatique des mains, de l'orientation de la tête et celle du regard.

### 4.1 Représentations informatiques

Nous avons exploité les données de capture de mouvement pour estimer : une enveloppe de la main sous la forme d'une sphère (4 a), la droite qui porte l'orientation de la tête (4 b). Nous avons récupéré les coordonnées spatiales de la cible du regard, à chaque instant, fournies par "FaceLab" (4 c). Le résultat de la combinaison Regard-Main et Tête-Main sont les intersections des modèles géométriques respectifs.

Ainsi, nous avons appliqué les modèles géométriques résultants sur les séquences de sessions d'enregistrement correspondants à des références spatiales manuelles et non-manuelles en gardant en mémoire les coordonnées spatiales de locus dans le but de caractériser l'événement de référencement en mesurant les distances des positions composants corporels par rapport aux positions spatiales du locus. Dans cet article, nous exposerons les résultats relatifs aux composants corporels non-manuels.

Le corpus en langue des signes française est un extrait d'une série de d'enregistrements réalisés

3. Lien vers le site web du projet : <http://www.irit.fr/marqspat/index.html>

4. Il s'agit d'un système de capture composé de 8 caméras infrarouges qui enregistrent les positions 3D de marqueurs réfléchissants posés sur les membres du signeur

5. Un système non invasif composé de deux caméras et d'un émetteur infrarouge. Il fournit sous forme de données 3D, l'orientation du regard. Une caméra de scène qui permet de synchroniser les données vidéos et données 3D

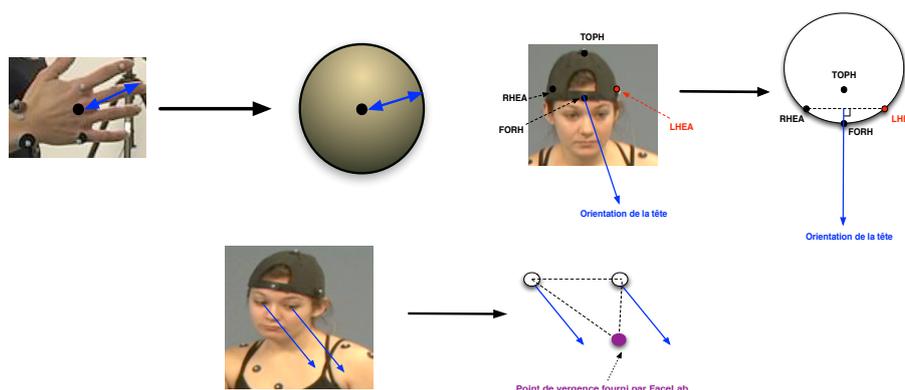


FIGURE 4 – Représentations informatiques : a) de la main, b) de l’orientation de la tête, c) de la cible du regard

à Websourd<sup>6</sup> dans l’objectif de diffuser des brèves d’actualités (La santé, la politique, l’économie, etc.). Les figures (1 a), 1 b) et 1 c)) sont des exemples visuels d’enregistrement. Nous avons choisi ce corpus car les conditions de son enregistrement se rapprochent le mieux de celles de discours naturels qui se compose de deux séries de vidéos. La première, qui contient sept vidéos, a été analysée en vue d’exploiter les résultats et de les appliquer sur la deuxième série qui se compose de deux vidéos. La durée totale des vidéos est de 337,68 secondes.

## 4.2 Annotations manuelles

Le logiciel Elan<sup>7</sup> a été utilisé à la fois pour annoter les vidéos LSF et LSQ en format texte et pour effectuer des recherches de valeurs d’annotation satisfaisant certaines contraintes.

Une annotation manuelle a pour objectif de transcrire les événements de référencement observés. L’annotateur prend connaissance du contenu de l’énoncé, puis repère les moments où le signeur associe un signe à une zone de l’espace. Par la suite, il marque les instants où le signeur fait référence à ces zones. Les indices visuels de marquage spatial sont les configurations manuelles de pointage, une brève fixation du regard et/ ou une rotation des épaules.

Les marqueurs de début et de fin de référencement se manifestent par le changement qui se produit dans l’un des indices listés précédemment :

- Le début d’un pointage manuel est marqué par la stabilisation de la configuration en question. La fin correspond au premier changement de configuration perçu.
- Le début et la fin d’une fixation correspondent au déplacement de l’iris qui précède et suit la fixation.
- Le début de rotation du buste correspond à la rupture avec une posture fixe. La fin du balancement se traduit par une posture fixe.

6. WebSourd est une entreprise qui a pour vocation la mise en œuvre de services d’accessibilité pour les personnes sourdes et le développement d’outils et métiers qui favorisent leur indépendance et leur citoyenneté.

7. (EUDICO Linguistic Annotator) est un outil de création, d’édition, de visualisation et d’annotation de données vidéos et paroles. Il a été développé à l’institut Max Planck en Irlande.

En parallèle, une annotation qualitative : 1) des rotations de la tête : selon les axes ; vertical, horizontal et combinés (les inclinaisons), 2) des rotations des épaules : selon les trois axes, 3) des fixations du regard vers : l'interlocuteur, l'espace ou une partie du corps du signeur et des signes effectués par les mains sous forme de gloses. La figure (5) illustre un exemple de transcription de la rotation des épaules et de la tête de 20 selon l'axe vertical vers la gauche [(+)rot(20/y)], le regard qui cible une zone en haut à gauche du signeur [l-up] et le signe manuel [ENSEMBLE].



FIGURE 5 – Annotations de pointages

Les annotations manuelles d'événements de référencement ainsi que des gloses, dans les vidéos de test, ont été réalisées par un entendant qui a un niveau B2 en LSF et vérifiées par un expert en LSF sourd de naissance. Les annotations qualitatives de mouvements corporels et du regard ont été réalisées séparément par le même annotateur sans vérification.

L'opération de vérification se résume en ces étapes. L'expert : prend connaissance de la méthode d'annotation ainsi que du contenu du discours, élimine les segments qui ne correspondent pas à un référencement, réajuste les limites des segments de référencement et rectifie les motifs de gestes correspondant au référencement.

## **5 Analyses et résultats**

Nous avons pris le parti de quantifier les variations de la position de la tête, de la cible du regard et de la main par rapport à celles du locus dans la réalisation de pointages "mono-sources". Pour cela nous avons appliqué les modèles géométriques (4.1) sur le corpus 3D et sélectionné les valeurs qui correspondent à des segments de référencement. Nous avons mesuré les distances tri-dimensionnelles qui séparent : 1) le centre de la main dominante (locus) de la droite qui porte l'orientation de la tête, 2) le centre de la main dominante (locus) de la cible du regard. Les distances mesurées représentent l'évolution du comportement du regard et de la tête au cours d'une séquence de référencement. Nous avons effectué ces mesures dans le but de caractériser

	Tête	Cible du regard
Main droite	100,4 ± 97,4 (mm)	545,3 ± 9,1 (mm)

TABLE 1 – Valeurs moyennes et écart-types des distances séparant la main droite en phase de création d'un signe spatialisé de 1) la droite qui porte l'orientation de la tête, 2) la cible du regard

Pointage	Zone		Motif	Occurrence	Mesures	(ETR)	(ET)
	(E1)	(E2)					
(Man)	6	3	(RMTE)	2	Precision	0.94	0.66
(NM)	17	13	(ETR)	12	Performance	0.52	0.8
(Man) & (NM)	9	3	(RMT)	3	F-measure	<b>0.67</b>	<b>0.73</b>

TABLE 2 – a) Zones clés de l'espace de signation et comparaison de nombre de pointages réalisés dans les vidéos de modélisation b) Occurrences de combinaisons gestuelles de référencement réalisé par la main droite (M), la tête (T), l'épaule (E) et/ou le regard (R). (vidéos de modélisation) c) Indice d'exactitude de détection de gestes de pointage dans les vidéos tests.

ce comportement en déterminant une loi de distribution de valeurs des distances. Nous avons constaté les mesures du tableau(1).

Le tableau (1) montre que les distances (Tête-Main droite) varient considérablement (20cm environ) et que le locus est référencé d'une manière très précise dans certains cas (distance nulle). Par contre, le regard n'est pas autant précis d'après la mesure de la distance qui sépare la cible du regard de la main droite (de 54cm environ) mais varie faiblement (de 1cm environ).

D'autre part, nous avons mesuré les délais temporels entre les mouvements de pointages "multi-sources". Pour cela, nous avons utilisé les annotations manuelles effectuées sur le corpus de Websourd, extrait les débuts et fins des segments appartenant aux pistes : Main droite, Tête, Les épaules et Cible du regard ( figure 5) qui correspondent à des séquences de référencement. Nous en avons construit des motifs de gestes (exemples : MTE, RET, etc.). Par la suite, nous avons choisi le motif le plus fréquent pour en construire un filtre de détection. Ensuite, nous avons appliqué le filtre sur les annotations de gestes de la seconde série d'enregistrement, puis, recensé le nombre de vrais et de faux positifs.

A partir des résultats du tableau (3), nous avons mesuré la précision et la performance des résultats de détection par filtrage, puis combiné les deux indices à l'aide de la méthode *F-MEASURE*<sup>8</sup> pour obtenir l'indice d'exactitude qui figure dans le tableau 2 c) qui montre que les deux filtrages aboutissent à des résultats complémentaires. En effet, les lignes *Precision* et *Performance* du tableau (2 c) montrent que le filtre (ETR) donne des résultats précis mais pas assez performant alors que le filtre (ET) donne des résultats performants mais pas assez précis.

Ensuite, nous avons pris en considération les zones de l'espace associées à un ou plusieurs éléments du discours pour chaque session d'enregistrement. Nous avons découpé l'espace en zones clés. L'exemple (6) illustre les zones clés associées à une partie de l'énoncé dont le résumé du

8.

$$F - Measure = \frac{2 \cdot Precision \cdot Performance}{Precision + Performance} \quad (1)$$

Motif	Video	$\delta$ (secondes)	VP	VN	FP	FN
(ETR)	884	0.36	15	105	0	10
(ET)	884	0.36	19	89	4	6
(ETR)	887	0.38	1	101	1	5
(ET)	887	0.38	6	108	9	0

TABLE 3 – Résultats de détections de gestes de pointage. **Motif** Référencement effectué par (ET) : les épaules et la tête (ETR) : les épaules, la tête et le regard. **Vidéo** : libellé de la vidéo. " $\delta$ " : Durée moyenne d'une séquence de référencement. **VP** : le nombre de séquences de référencement détectées. **VN** : le nombre de séquences qui ne correspondent pas à un référencement et qui n'ont pas été détectées. VN est calculé selon cette formule " $VN = \frac{D - \sum VP_{duration}}{\delta}$ " où D représente la durée de la vidéo. **FP** : le nombre de faux positifs. **FN** : le nombre de séquences de référencement non détectées.

contenu est : "un groupe de personnes attaque une ambassade en Syrie". Nous avons mesuré le nombre de pointages manuels et non manuels (NM) effectués vers chaque zone que nous étiquetons (E1) et (E2) qui représentent, respectivement, les zones associées à l'énoncé localisées à gauche et à droite du signeur. Nous remarquons que le signeur pointe plus fréquemment vers la zone (E1) et ceci selon les deux modalités de référencement (Man, NM) .

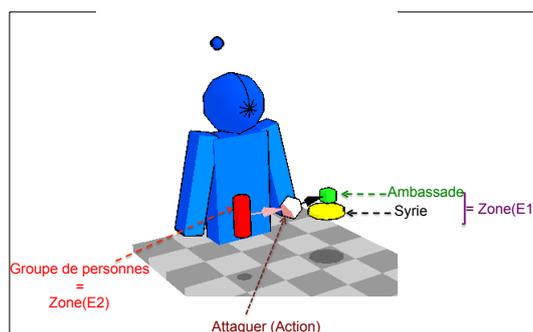


FIGURE 6 – Exemple de zones clés de l'espace associées à l'énoncé

## Interprétations

Les valeurs de distances cible du regard – locus et tête – locus obtenues varient différemment. Ceci montre que nous ne pouvons les caractériser par une même loi de distribution. Ainsi, nous remettons en question le fait que les distances mesurées soient considérées comme des indices pour caractériser l'événement de référencement.

Les composants corporels de référencement qui ont fait l'objet de filtres (2 b) appliqués au corpus de brèves d'actualités sont : les épaules, la tête et le regard. Nous notons que nous avons établi, implicitement, un rapprochement entre les mouvements des épaules et ceux du buste car, d'une part, nous avons constaté que pour détecter une rotation du buste l'annotateur se sert des déplacements des épaules, et d'autre part, les outils d'analyse d'image dont nous disposons qui servent à la détection de mouvement prennent en compte le facteur de contraste dans l'image ce qui signifie les éléments de bord et donc les épaules. (Cuxac, 2000) et (Daigle & Parisot, 2006)

précisent que les gestes de référencement sont réalisés par : le regard, les mains, la tête et le buste. Ainsi, les résultats de détection en apportent une confirmation car les segments correspondants incluent des gestes de combinaisons diverses.

D'un autre côté, l'analyse spatiale (2 a) montre que dans les séquences de référencement, le signeur privilégie une zone plutôt qu'une autre. Ceci peut s'expliquer par les contraintes physiques et/ou l'importance d'un élément discursif. Cet indice servira sans doute à nourrir les outils de traitement informatique injecté sous forme de connaissance *a priori* sur la fréquence de référencement vers une zone donnée.

## 6 Conclusion

Nous avons analysé le processus de référencement, établi des classes de combinaisons de gestes puis proposé deux méthodes : la première qui permet de représenter les variations spatiales d'un référencement mono-modal, la deuxième modélise les variantes temporelles d'un référencement multi-modal.

Nous avons, en partie, apporté des éléments de réponses au processus de compréhension du contenu de l'espace de signation malgré les variations de gestes de référencement. En effet, nous avons introduit une représentation informatique de l'emplacement d'un locus dans l'espace et établi l'association espace-référencement qui se traduit par une modélisation géométrique de la relation entre l'orientation de la tête et celle du locus d'une et entre la cible du regard et celle du locus. La mise en place d'un mécanisme de fusion de ces représentations géométriques mono-modales ainsi que la modélisation temporelle de combinaisons de gestes serait une piste intéressante pour concrétiser la modélisation d'un référencement multi-modal.

## Remerciements

Je remercie M. Patrice Dalle, mon directeur de recherche qui a supervisé ce travail. Je remercie les linguistes pour leurs collaborations dans la mise en oeuvre des corpus. Le projet (Marqspat) est soutenu par le CRSH, dans le cadre d'un partenariat stratégique soutenu par le CFQCU. Le corpus d'illustration (3a et b) a été réalisé en interne par Juliette Dalle, assistante ingénieur de l'équipe Traitement et compréhension de d'Image de l'Institut de Recherche en Informatique de Toulouse (IRIT). Je la remercie d'avoir revu et corrigé les annotations manuelles en tant qu'experte en LSF.

## Références

- BEN MLOUKA M. (2012). Analyse automatique de discours en langue des signes : Représentation informatique et traitement de l'espace de signation. In TALN-RECITAL, Ed., *Chercheurs en Informatique pour le Traitement Automatique des Langues*, p. 182, Grenoble.
- BEN MLOUKA M. & DALLE P. (2011). Analyses d'annotations pour la reconnaissance. In TALN-DEGELS, Ed., *Traitement Automatique de la Langue des Signes*, Montpellier.

CUXAC C. (2000). *La langue des signes française (LSF) - Les voies de l'iconicité*. Paris : Faits de Langues : Ophrys.

DAIGLE D. & PARISOT A. M. (2006). *Surdit  et soci t  : Perspectives psychosociale, didactique et linguistique*. Collection Sant  et soci t . Presses de l'Universit  du Qu bec.

EMMOREY K. & REILLY J. (1995). *Language, gesture and space*. Hillsdale, New Jersey Hove, UK : Lawrence Erlbaum Associates publishers.

FUSELLIER-SOUZA I. (2004). *S miog n se des langues des signes :  tude de langues des signes primaires (LSP) pratiqu es par des sourds br siliens*. PhD thesis, Universit  Paris 8.

LEROY E. (2010). *Didactique de la Langue des Signes Fran aise : Attitudes et strat gies p dagogiques de l'enseignant sourd*. PhD thesis, Universit  Paris 8.

LIDDELL S. (1980). *American sign language syntax*. Approaches to semiotics. Mouton.

PARISOT A.-M. & RINFRET J. (2006). La variation dans le marquage de la sp cificit  en langue des signes qu b coise. In LILLE3, Ed., *Colloque International Syntaxe, interpr tation, lexique des langues sign es*, Lille.

RINFRET J. (2009). *L'association du nom en langue des signes qu b coises : Formes, fonctions et sens*. PhD thesis, UQAM.

STOKOE W. (1960). Sign language structure : An outline of the visual communication systems of the American deaf. *Journal of deaf studies and deaf education*, **10**(1), 3-37.



# Variation et variabilité dans les Sciences du Langage : analyser, mesurer, contextualiser

## Variation and variability in the Language Sciences: analysing, measuring, contextualising

Les notions de variation et de variabilité sont au cœur de nombreux domaines liés aux Sciences du Langage. Le linguiste peut aussi bien être amené à les interroger en tant qu'objet de recherche qu'au travers des problèmes méthodologiques qu'elles posent. La variation comme objet d'étude offre ainsi un terrain de réflexion particulièrement riche et ouvert. Elle permet par exemple d'apporter un éclairage spécifique, et souvent fondamental, à nombre de problématiques à travers les travaux de contextualisation historique, théorique ou sociologique qu'elle peut impliquer. Par ailleurs, le positionnement face à une norme ou un standard, ou l'étude des systèmes et des structures qu'ils imposent représentent des pistes de recherche qui émergent autour du concept de variabilité. Avec le récent développement de la linguistique de corpus qui permet la description, l'analyse et la comparaison des systèmes linguistiques à grande échelle, les questions de changement, d'évolution, d'instabilité ou d'incertitude apparaissent également comme des enjeux scientifiques majeurs. À un niveau plus méthodologique, c'est la manière de traiter variation et variabilité qui soulève des questionnements essentiels. On peut par exemple s'interroger sur la sélection, voire la construction d'outils de mesure ou de cadres d'analyse pertinents, qu'ils soient qualitatifs ou quantitatifs. De nombreuses démarches entrent donc en résonance avec des problématiques liées à la variabilité : codages et transcriptions, traitement de corpus, entretiens et enquêtes de terrain, outils statistiques, logiciels et programmes, sélection et exploitation de variables...

The notions of variation and variability are central to various studies in the Language Sciences. The linguist can investigate these notions as a research object or focus on the methodological issues these same notions raise. Variation as a research object offers a wide range of possible reflexions. It can bring specific and fundamental perspective to the studies in the Language Sciences, through theoretical, historical or sociological contextualisation for instance. Besides, finding one's positioning in relation to a norm, a standard or studying the constraints and structures that a norm, a standard can impose, are as many research directions linked to the concept of variability. With the recent development of corpus linguistics, which allows for the description, analysis and comparison of the linguistic systems of the different languages on a large scale, the issues of change, evolution, instability or uncertainty seem to be of crucial scientific importance. On a more methodological level, taking variation and variability into account seems to be one of the essential issues for the Language Sciences. The linguist can ask himself whether he is selecting and using the best measuring tools or the accurate analytical frameworks (whether they are qualitative or quantitative). Many approaches and many protocols try to take the notion of variability into account, for example through codings and transcriptions, fieldwork investigations and interviews, statistical tools, softwares and computer programming, selection of variables...