

COLDOC 2014

Colloque des doctorants et jeunes chercheurs du laboratoire MoDyCo
<http://coldoc2014.free.fr>

Diversité des langues

Les universaux linguistiques à l'épreuve des faits de langue



Laboratoire MoDyCo – CNRS UMR 7114

Université Paris Ouest, France

Université Paris Descartes, France

Sommaire

<i>Remerciements</i>	3
<i>Introduction : diversité des langues : les universaux linguistiques à l'épreuve des faits de langue</i>	4
<i>Composition des comités scientifique et d'organisation</i>	6
Jovan KOSTOV : Le marquage différentiel de l'objet en tant qu'universel linguistique : le cas du macédonien	7
Lin XIAO : À propos du marquage différentiel de l'objet et de 把 <i>bǎ</i> en chinois	18
Ophélie GANDON : Stratégies de relativisation atypiques dans l'aire Caucase-Iran-Anatolie	27
Guillaume JAUDHUIN : Analyse multimodale et contrastive des interjections émergentes : le cas de <i>genre</i> en français et <i>as if</i> en anglais britannique	42
Audrey NOËL : Les connaissances lexicales : universalité ou variété de la catégorisation sémantique ? Étude du lexique de locuteurs de l'île de La Réunion	54
Sarah HARCHAOUI : La notion de multiethnolecte dans le contexte scandinave	62
Yufei Guo : Politique linguistique de la Chine envers les dialectes : le débat autour du cantonais au début du 21 ^e siècle	78
Rosa RUSSO : Les verbes de mouvement italiens : stratégies de lexicalisation des natifs	95
Simona ANASTASIO : Diversité dans l'expression de l'espace dynamique en L1 : étude contrastive entre italien, français et anglais	106
Emilie RIGUEL : Les <i>phrasal verbs</i> dans les productions écrites des apprenants de l'anglais langue seconde.....	117
Emmanuelle EGGINTON : Universaux et acquisition du lexique chez les jeunes enfants de migrants allophones en France	131
Najima KERROU : Apprentissage du français L3 : spécificités temporelles de la langue française du point de vue d'apprenants japonais en milieu hétéroglotte	141

Remerciements

Le comité d'organisation est reconnaissant aux instances du laboratoire pour leur aide logistique et leur soutien financier, et remercie particulièrement Jean-Luc Minel et Caroline Bogliotti, directeurs du laboratoire MoDyCo, pour leurs conseils avisés et Myriam Djedi et Alexandre Sourie, gestionnaires du laboratoire MoDyCo, pour leur disponibilité et leur efficacité.

Nous remercions également les membres du comité scientifique pour leur rigueur dans l'examen des contributions reçues, en particulier Anne Lacheret, enseignant-chercheur référente, pour son expertise et sa bienveillance ; les conférenciers Françoise Gadet, Marzena Watorek et Alexandru Mardale ; chacun des intervenants dont la plupart des travaux présentés ont été réunis dans ces présents Actes ; ainsi que tous les participants pour l'intérêt qu'ils ont bien voulu porter aux thèmes du colloque.

Nous remercions également l'Université Paris Ouest Nanterre La Défense notamment pour la mise à disposition des locaux et matériels de projection ainsi que des services de reprographie, l'école doctorale 139, le département SDL, l'UFR PHILIA et l'association ATALA pour leur soutien. Nos remerciements vont également à Pablo Rodrigo pour la réalisation de la communication visuelle du colloque (affiches, programme et logo).

Enfin, nous remercions chaleureusement nos collègues et amies Julie Beliaõ, Marine Damiani et Kaja Dolar, membres des comités d'organisation des éditions antérieures de ColDoc, pour leurs conseils, leurs indéfectibles encouragements et leur précieuse aide enthousiaste.

Diversité des langues : les universaux linguistiques à l'épreuve des faits de langue

Pour sa onzième édition, COLDOC s'intéresse à la diversité des langues et de leurs usages, en tant que source de nombreuses interrogations dépassant la sphère de la linguistique. Cet appel à contributions se veut ouvert et cherche à confronter différents points de vue autour de problématiques fortement pluridisciplinaires.

La variation interlangue nous oblige à constater à la fois les spécificités des langues et de leurs usages (pragmatique, sémantique morpho-syntaxe, phonologie), faisant ainsi émerger la question des universaux qui vise à décrire la faculté universelle du langage. Face à ces observables aussi divers que nombreux, comment articuler diversité des représentations linguistiques et universalité de la faculté de langage ?

La présence de plus en plus massive de grands corpus permet aujourd'hui de mener une nouvelle réflexion sur la diversité des langues et de leurs pratiques. L'ensemble de données langagières collectées issues de l'étude de différentes langues constitue une importante source pour une validation empirique des modèles décrivant tant des traits universels que des schémas représentationnels particuliers (par exemple la métaphore et autres contenus implicites).

Ainsi, la question qui est au cœur de ce colloque est celle de la confrontation de la notion d'universaux linguistiques et de la réalité des faits langagiers. Que devient la théorie de ces universaux face à la pratique des langues ? Dans quelle mesure la linguistique moderne, grâce à de nouveaux outils de fouille et de nouvelles méthodes empiriques, peut-elle apporter des réponses à la recherche de régularités ou de diversités linguistiques ?

Linguistes, traducteurs, typologues, psycholinguistes, anthropologues, sociolinguistes, enseignants de langues étrangères, talistes, scientifiques de terrain contraints de rapprocher des systèmes qui découpent le monde différemment, vous êtes confrontés à une diversité bien concrète. Ce colloque souhaite vous apporter une place d'honneur afin que vos outils et vos observations viennent éclairer les théories linguistiques qui cherchent à les prendre en compte.

Nous invitons masterants, doctorants et jeunes chercheurs à venir exposer leurs réflexions sur ce thème à partir de leur propre pratique et ce quel que soit le degré

d'avancement de leur recherche.

Plusieurs questions, mentionnées ici sans exhaustivité, peuvent être envisagées comme source d'inspiration :

- Les langues recèlent-elles des universaux ? Comment les décrire ?
- Dans quelle mesure une langue est-elle un système spécifique ?
- Négocier le sens d'une langue à une autre signifie-t-il le trahir ?
- Enseigne-t-on seulement un lexique et une morpho-syntaxe dans l'enseignement d'une langue étrangère ?
- Comment faire face aux problèmes soulevés par la traduction des contenus posés et implicites ? Comment traduire une métaphore ?
- Existe-t-il des tendances universelles des évolutions des langues (figement, grammaticalisation, perte de désinence...) ?
- Que révèle la constitution de grands corpus multilingues alignés sur l'existence d'universaux linguistiques ?
- Comment éviter les biais de l'ethnocentrisme dans ses pratiques de recherche ?
- Existe-t-il des catégories syntaxiques/sémantiques/phonologiques universelles ? Lesquelles ?
- De quels outils dispose-t-on pour analyser les similitudes et les différences phonologiques, morphosyntaxiques, sémantiques et discursives entre les langues ?
- Quels modèles adopter pour regrouper les langues ? Des modèles typologiques ou des classements par famille linguistique ?
- Quel est le poids des structures universelles ou particulières de la langue qu'apprend l'enfant dans le processus d'acquisition ?
- Au-delà de leur diversité à quels universaux de l'interaction langagière les rituels communicationnels répondent-ils ?

Comité scientifique

- Delphine BATTISTELLI (MoDyCo, Université Paris Ouest)
- Annie BERTIN (MoDyCo, Université Paris Ouest)
- Caroline BOGLIOTTI (MoDyCo, Université Paris Ouest)
- Sarah DE VOGUE (MoDyCo, Université Paris Ouest)
- Guillaume DESAGULIER (MoDyCo, Université Paris Ouest, Université Paris 8)
- Markus EGETMEYER (Université Paris 4 – Sorbonne)
- Benjamin FAGARD (ENS, LaTTiCE, Université Paris 3)
- Françoise GADET (MoDyCo, Université Paris Ouest)
- Sylvain KAHANE (MoDyCo, Université Paris Ouest)
- Maria KIHSTEDT (MoDyCo, Université Paris Ouest)
- Anne LACHERET (MoDyCo, Université Paris Ouest)
- Bernard LAKS (MoDyCo, Université Paris Ouest)
- Denis LE PESANT (MoDyCo, Université Paris Ouest)
- Caroline MELLET (MoDyCo, Université Paris Ouest)
- Jean-Luc MINEL (MoDyCo, Université Paris Ouest)
- Valelia MUNI TOKE (DeDyL – CELIA, IRD)
- Colette NOYAU (MoDyCo, Université Paris Ouest)
- Christophe PARISSÉ (MoDyCo, Université Paris Ouest)
- Sabine PÉTILLON (MoDyCo, Université Paris Ouest)
- Benoît SAGOT (INRIA, Université Paris Diderot – Paris 7)

Comité d'organisation

- Giulia Barreca
- Trang Luong
- Laure Martin
- Ilaine Wang

Le marquage différentiel de l'objet en tant qu'universel linguistique : le cas du macédonien

Jovan KOSTOV

EA 4509 (Sens, Texte, Informatique, Histoire) – 1, rue Victor Cousin, 75 005 Paris
jovan.kostov@paris-sorbonne.fr

RÉSUMÉ

Le marquage différentiel de l'objet (MDO) est un procédé syntaxique très répandu dans les langues indo-européennes, et notamment dans les langues slaves du sud. En macédonien et en bulgare, ce phénomène est traditionnellement appelé « redoublement de l'objet ». C'est le cas le plus typique que l'on retrouve aussi dans les langues romanes (espagnol, roumain, etc.). Différentes études (Bossong, 1985, 1998 ; Lazard, 2001) considèrent ce procédé comme un universel linguistique. Dans cet article, nous étudierons les cas les plus saillants de MDO en macédonien, dans le but de poursuivre nos recherches sur la syntaxe du groupe verbal.

ABSTRACT

Differential object-marking as a linguistic universal: The Macedonian language question

Differential object-marking (D.M.O.) is a syntactic process often found in indo-European languages, and even more so in Southern Slavic languages. The marking of the definite direct object in Macedonian, also known as "object doubling", is a typical feature of this language but also that of some romance languages such as Spanish and Romanian. Several linguists (Bossong, 1985, 1998 ; Lazard, 2001) consider this process as a linguistic universal. The aim of this paper is to consider DMO under its different forms in Macedonian, within the context of linguistic research applied to the verbal group.

MOTS-CLEFS : MDO, verbe, syntaxe, macédonien, objet direct, slave, groupe verbal, marquage différentiel

KEYWORDS: DOM, verb, syntax, macedonian, direct object, slavic, verbal group, differential marking

1 Marquage différentiel de l'objet : définitions préliminaires et cadre de la présente étude

Le marquage différentiel de l'objet (MDO dans la suite de cet article) suscite l'intérêt de nombreux syntacticiens et linguistes-typologues qui travaillent sur le groupe verbal dans différentes langues. Parmi les travaux les plus significatifs sur le MDO dans les langues slaves, et notamment sur l'ensemble dialectal bulgaro-macédonien, on peut noter les travaux de Guentcheva (1990, 1994) qui concernent le bulgare mais qui contiennent des informations précieuses sur le groupe verbal en macédonien, et ceux d'Adamou (2009) qui traitent le MDO dans les parlers slaves du nord de la Grèce et du sud de la Macédoine. Dans cet article, après un bref rappel des définitions communément acceptées de ce phénomène syntaxique et une présentation du fonctionnement du groupe verbal en macédonien, nous allons effectuer une étude des cas les plus courants de marquage différentiel de l'objet direct à travers des exemples concrets. Notre but est de montrer que la saillance de ce phénomène va au-delà des cas typiques que l'on rencontre dans les définitions existantes, notamment par la présence d'autres marqueurs comme les pronoms clitiques et les prépositions caractéristiques.

1.1 Le MDO en tant qu'universel linguistique

Dans son étude sur le MDO, Adamou (2009) définit le marquage de l'objet direct – terme qu'elle emprunte à Bossong (1985, 1998) – comme un phénomène d'apparition de l'objet avec ou sans marque spécifique qui se manifeste soit comme un mot « autonome » préposé ou postposé à un objet, soit comme une marque d'unité casuelle (Adamou, 2009). Par ailleurs, on « marque » un objet en fonction de son actualisation en discours et suivant trois paramètres majeurs :

1. le caractère animé de l'objet direct ;
2. le caractère spécifique de l'objet direct ;
3. le déplacement de l'objet direct.

Cela représente une première approximation du MDO que l'on peut affiner en valeurs graduelles selon l'exemple de Mardale (2012). Le cas typique et, de surcroît, le plus exploité du MDO est celui de l'objet direct en espagnol contemporain précédé par la préposition *a* en fonction du caractère humain ou non de l'objet direct, comme dans les exemples suivants :

- (1) He traicionado tu confianza
 AUX comprar-PTCP POSS confianza-F
 « J'ai trahi ta confiance. »
- (2) He traicionado a María
 AUX regañar-PTCP PREP María
 « J'ai trahi Marie. »

Bossong examine ce phénomène à travers plusieurs exemples (Bossong, 1985) issus de différentes langues européennes. Les mécanismes de manifestation du MDO dans les cas examinés par Bossong dénotent l'existence d'une diversité typologique sur le plan syntaxique et c'est dans ce sens qu'il convient de caractériser le MDO comme un universel linguistique qui pourrait être utilisé dans la classification typologique des langues. En macédonien, ce phénomène semble se manifester de manière tout à fait unique, notamment par l'adjonction des pronoms clitiques et par des prépositions spécifiques. Cela nous a motivé à étudier de manière plus profonde ce phénomène à aller au-delà de ce qui a été observé auparavant.

1.2 Le redoublement de l'objet en macédonien

Le macédonien, le monténégrin et le croate sont considérés comme les langues slaves du sud les plus récemment standardisées (Petruševska, 2014). La codification du macédonien intervient en 1945 avec l'acte de l'Assemblée nationale antifasciste (ASNOM). Après la parution des premières grammaires (Kepeski, 1945 ; Lunt, 1952 et Koneski, 1953) qui étaient principalement orientées vers un recensement et une description très sommaire des catégories des mots, il a fallu plus d'un-demi siècle pour voir paraître la première syntaxe macédonienne (Minova-Gjurkova, 1994) et de décrire le fonctionnement de la phrase macédonienne avec une terminologie inspirée principalement de l'école russe.

Dans cette syntaxe, le marquage de l'objet direct n'est pas traité comme un phénomène linguistique spécifique, mais il est sommairement abordé dans le cadre du chapitre traitant de l'objet direct et de l'emploi des pronoms clitiques. Cette description syntaxique, même incomplète, représente une base solide pour peaufiner la description de la structure et du fonctionnement des unités du discours sur le plan syntaxique et sémantique en macédonien contemporain. Dans la continuité de travaux existants, notamment ceux de Koneski (1953) et de Friedman (2001), et plus particulièrement dans le cadre du programme de la fédération TUL intitulé « Unité et diversité dans le

marquage différentiel de l'objet », nous tenterons d'approfondir l'étude du marquage de l'objet direct en macédonien standard et d'en établir une grille des conditions.

1.3 Le groupe verbal en macédonien

D'un point de vue typologique, le macédonien est une langue slave dont le système nominal a perdu la déclinaison. La disparition des cas est un processus que l'on peut observer dans les écrits du 19^e et dans ceux de la première moitié du 20^e siècle. Aujourd'hui, seul le vocatif persiste dans certains parlers de la Macédoine de l'ouest. Sous l'influence des premiers didacticiens (en macédonien *učebnikari*) et après une tentative avortée de création d'une langue commune bulgare-macédonienne, le macédonien se développe en tant que langue autonome. Son système verbal est caractérisé comme l'un des systèmes slaves les plus complexes et les moins étudiés (Friedman, 1977). Cependant, on peut d'ores et déjà retenir que sa structure phrastique se caractérise par un ordre S-V-O. Les mots sont principalement composés de schémas syllabiques CV(C) et la disparition de l'article favorise l'apparition de marqueurs déictiques postposés¹ au nom ou à l'adjectif-épithète (*kuče-to* [le chien], *žena-ta* [la femme], *maž-ot* [l'homme]), ainsi que celle de déterminants (numériques, indéfinis et démonstratifs) que l'on rencontre également en bulgare.

Pratiquement toutes les grammaires existantes se contentent de dire que le redoublement de l'objet direct défini est obligatoire. En outre, l'objet direct est indispensablement marqué lorsqu'il est animé et/ou spécifique ou, plus précisément, « défini » et/ou « humain » (Mardale, 2012). Voici quelques exemples préliminaires à partir desquels nous pouvons constater que l'objet direct en macédonien standard est marqué :

- (3) Jas **go** vidov **kučeto**
 Je le voir-1-SG-PST chien-DEF
 « J'ai vu le chien. »
- (4) Jas vidov kuče
 Je voir-1-SG-PST chien-INDF
 « J'ai vu un chien. » (non-spécifique)

¹ Pour des questions liées à la tradition terminologique, les linguistes continuent à appeler cette marque déictique « article défini » et rajoutent l'adjectif « tripartite » pour faire référence aux trois séries différentes qui marquent l'espace et le temps. Dans cet article, nous allons nous tenir à cette tradition pour assurer une plus grande compréhension au sein de la communauté scientifique.

- (5) Jas vidov edno kuče
 Je voir-1-SG-PST un chien-INDF

« J'ai vu un chien. » (spécifique/non-spécifique : deux interprétations possibles)

Ce qui nous intéresse particulièrement, ce sont les cas équivalents à l'exemple (3). En effet, la marque du MDO – le pronom clitique *go* avant l'objet – est obligatoire à chaque fois que l'objet est défini. De surcroît, l'objet porte l'article *-to* ou désigne un être humain. Dans l'exemple (5) nous pouvons noter la présence du modifieur *edno*, mais son sens peut être à la fois équivalent de celui de l'exemple (4) – un chien « quelconque » – ou un chien « en particulier ». Dans ce cas, seul le contexte peut nous indiquer si l'adjectif *edno* a la valeur d'un modifieur de *kuče* ou, tout simplement, d'un article indéfini.

2 Le marquage différentiel de l'objet en macédonien standard : étude des différents cas observés

Dans cette partie, nous étudierons quelques cas de marquage de l'objet direct en macédonien et nous essaierons de définir les critères d'apparition selon les classifications existantes (Bossong, 1998 ; Adamou, 2009 ; Mardale, 2010). Dans le cadre de notre recherche doctorale, nous avons travaillé sur la morphologie du groupe verbal, et plus particulièrement sur la modélisation des règles morphotaxiques dans le but de les exploiter dans un système de génération automatique des paradigmes. Cependant, le macédonien est une langue récemment codifiée et son processus de standardisation et de stabilisation est toujours actif. De ce fait, une recherche sur le fonctionnement syntaxique du verbe devenait indispensable et c'est pour cela que nous nous sommes intéressé en particulier aux verbes transitifs directs qui présentaient une double conjugaison. Cependant, les recherches (en cours) ne nous permettent pas encore de dire avec assurance qu'il s'agit d'un phénomène du même ordre. Toutefois, le mécanisme est similaire à celui du MDO. Les trois contextes que nous aborderons dans cet article sont les contextes suivants :

1. le MDO dans les constructions S-V-O ;
2. le MDO dans les constructions contenant une forme nominale du verbe ;
3. les cas de déplacement à gauche de l'objet direct.

2.1 Le MDO dans les constructions verbales S-V-O

Dans l'exemple (3), nous avons vu un cas de redoublement de l'objet direct défini, marqué par la présence du pronom clitique *go*. L'article défini *-to* indique également ce marquage de la définitude de l'objet. Le MDO est obligatoire et la présence du pronom clitique *go* est indispensable. En d'autres termes, le pronom clitique ne pourrait pas être employé si le COD n'était pas défini. Les tableaux qui suivent récapitulent les formes des marques de l'article défini en macédonien et les pronoms objet qui sont utilisés lors d'un marquage de l'objet direct :

	Singulier			Pluriel
	Féminin	Masculin	Neutre	Masculin/Féminin/Neutre
Indéfini	-ta	-ot	-to	-te
Proximal	-va	-ov	-vo	-ve
Distal	-na	-on	-no	-ne

TABLE 1 – Désinences déictiques (articles définis) en macédonien

	Singulier			Pluriel
	Féminin	Masculin	Neutre	Féminin/Masculin/Neutre
1ère personne	me			ne
2e personne	te			ve
3e personne	ja	go		gi

Table 2 – Pronoms clitiques objets en macédonien

Le caractère déictique de l'article postposé en macédonien joue, en effet, un rôle prépondérant dans le MDO et peut être considéré comme un facteur déclencheur du marquage. Nous pouvons constater cette faculté à partir des exemples suivants où nous observons l'absence (6) puis la présence (7) d'un pronom clitique, conditionnée par la définitude du nom :

(6) Nie jadevme salata
 Nous manger-1-PST-PL-IPFV salade-INDF

« Nous avons mangé une salade. »

- (7) Nie ja izedovme salatata
 Nous la manger-1-PST-PL-PFV salade-DEF
 « Nous avons mangé la salade. »

Ces deux exemples nous indiquent que dans ce type de constructions « canoniques » avec une structure S-V-O l'objet direct est marqué à chaque fois qu'il est défini. Le marquage est également obligatoire lorsque l'objet désigne un être humain (humanité) :

- (8) Marija vide avtomobil
 Marie voir-1-PST-PL-PFV voiture-INDF
 « Marie a vu une voiture. »
- (9) Nie go intervjuiravme Stefan
 Nous le interviewer-1-PST-PL-PF Stéphane
 « Nous avons interviewé Stéphane. »

Il est à noter que l'article défini ne peut pas être adjoint aux noms propres. Dans ce cas, le marquage s'effectue uniquement par la présence du pronom clitique, ce que Bossong appelle « un marquage verbal » par opposition à un « marquage nominal » (Bossong, 1998). La spécificité de l'objet direct est aussi un critère de MDO, sauf dans certains énoncés contenant une vérité générale – un cas que nous n'allons pas développer dans cet article car il mérite une étude à part entière.

2.2 Le MDO dans les constructions contenant une forme nominale du verbe

Dans les grammaires existantes (Kepeski, 1945 ; Lunt, 1952 ; Koneski, 1053 ; Minova-Gjurkova, 1997 ; Minova-Gjurkova, 2005), le paradigme verbal inclut, entre autres, la forme nominale du verbe, appelée également « le nom/substantif verbal ». Cette forme, construite sur le thème du présent des verbes imperfectifs avec le suffixe *-nje*, a toutes les caractéristiques formelles d'un substantif et son étude au sein de la classe du verbe est une grande source de confusions, aussi bien pour les apprenants étrangers que pour les locuteurs natifs. Cette forme est employée en particulier dans l'anaphore verbale et elle sert à nominaliser le procès d'un verbe. Dans ces constructions, nous avons décelé un procédé très similaire au MDO avec la préposition *na* que l'on retrouve dans certaines variétés du macédonien comme celle d'Ohrid. Cette préposition joue le même rôle que celui de la préposition *pe* dans le cas de l'accusatif prépositionnel en aroumain avec lequel le macédonien se trouve en contact (Adamou, 2009). Nous étudierons ce

procédé avec une suite d'énoncés dont le lien anaphorique est le verbe.

- (10) Marija prekrasno pee pesni
 Marie magnifique-ADV chanter-3-SG-PRES chanson-PL
 « Marie chante des chansons magnifiquement. »

Cet énoncé est constitué d'un sujet (*Marija*), d'un verbe (*pee*) auquel se rattache un complément circonstanciel de manière (*prekrasno*) et d'un objet direct qui est un nom commun (*pesni*). Dans le cas d'une anaphore où le verbe *pee* est repris par un nom verbal (*peenje*), les autres constituants deviennent ses compléments au sein d'un groupe nominal :

- (11) Peenjeto na Marija e prekrasno
 chanter-DEF de Marie être-3-SG-PRES magnifique-ADJ
 « Le chant de Marie est magnifique. »

- (12) Peenjeto pesni e prekrasno
 chanter-DEF chanson-PL-INDF être-3-SG-PRES magnifique-ADJ
 « Chanter des chansons est magnifique. / C'est magnifique de chanter des chansons. »

Le verbe, qui a morphologiquement dérivé vers la classe des noms, est employé ici dans sa forme définie (article *-to*) et représente le noyau du groupe nominal sujet. Nous pouvons constater que la préposition *na* est obligatoire lorsque le nom verbal est complété par l'élément qui avait la fonction sujet dans l'énoncé précédent, mais proscrite dans le cas d'une complémentation avec le COD. Toutefois, dans le cas d'une complémentation du nom verbal avec un nom défini qui se trouverait en position objet, l'emploi de la préposition *na* devient obligatoire :

- (13) Peenjeto na pesnite e prekrasno
 chanter-DEF des chanson-PL-DEF être-3-SG-PRES magnifique-ADJ
 « Chanter des chansons est magnifique. »

Cette caractéristique est source de nombreuses hésitations lors de l'emploi de la préposition *na* après un nom verbal, en particulier chez les locuteurs natifs. L'ambiguïté est très souvent résolue par le remplacement du nom verbal par une construction avec *da*, ce qui permettrait de construire des énoncés qui sont plus facilement identifiables en termes d'acceptabilité grammaticale :

- (14) *Da peesh Marija e prekrasno
 que-CONJ-SUBORD pee-2-SG-PRES Marie être-3-SG-PRES magnifique-ADJ
 « Chanter Marie c'est bien. »

- (15) Da peesh pesni e prekrasno
 que-(CS) pee-2-SG-PRES chanson-PL être-3-SG-PRES magnifique-ADJ
 « Chanter de chansons c'est bien. »

Ainsi, à chaque fois que le nom verbal peut être remplacé par la construction *da*, il se comportera davantage comme un verbe, tandis que dans le cas contraire il aura toutes les propriétés d'un substantif. Cela ne nous permet pas de dire systématiquement qu'il s'agit d'un cas de MDO avec la préposition *na*. Toutefois, ce phénomène s'en rapproche considérablement et laisse entrevoir que l'origine de ce nom est un verbe. Dans ce cas, le MDO serait conditionné par la définitude du complément de nom qui était antérieurement un complément d'objet direct.

2.3 Le déplacement de l'objet direct

Le troisième cas le plus saillant de MDO se rapproche de celui que nous avons pu analyser dans les constructions canoniques S-V-O (exemple (5)). Nous avons vu que l'emploi de *edno* dénote parfois un objet direct spécifique. Cette modification est d'autant plus saillante lorsque nous avons un déplacement de l'objet direct en tête de phrase, comme dans les exemples suivants :

- (16) Nie čekame edna kniga
 Nous attendre-1-PL-PRES une-DET livre-SG-F
 « Nous attendons un livre. »

- (17) Edna kniga ja čekame cel den
 Une-ART-INDF livre-SG la attendre-1-PL-PRES tout jour-SG-M
 « C'est un livre que nous attendons toute la/une journée. »

Dans l'exemple (16), l'objet *kniga* peut être un nom indéfini ou un défini, suivant le contexte qui l'entoure. Le déplacement de l'objet à gauche n'est permis que pour les cas d'un objet spécifique et l'emploi du pronom clitique COD est alors obligatoire. Compte tenu de l'emploi très particulier des pronoms clitiques comme illustré par l'exemple (17), nous pouvons considérer que l'objet direct spécifique est obligatoirement précédé par le pronom clitique. Dans ce cas, nous avons encore une fois un marquage verbal (et non pas nominal), renforcé par le modifieur *edna*.

3 Conclusion : sur les traces d'un MDO atypique ?

Les exemples que nous avons analysés dans cet article mettent en évidence un fonctionnement très particulier du groupe verbal en macédonien contenant des

procédés très similaires à celui du MDO (Bossong, 1985, 1998). L'apparition de certains marqueurs – notamment des pronoms clitiques objet et de la préposition *na* – est conditionnée par les critères de MDO tels qu'ils sont établis dans les études de Bossong (1985, 1998), d'Adamou (2009) et de Mardale (2010). Concernant les trois cas de figure étudiés, nous pouvons rendre les conclusions suivantes :

- dans le cadre des constructions de type S-V-O, l'objet direct est obligatoirement marqué par la présence d'un pronom clitique COD lorsqu'il est humain, défini et/ou spécifique ;
- les constructions de type forme nominale d'un verbe transitif direct + COD dans le cas de l'anaphore du verbe sont sujettes à un marquage qui se manifeste par la préposition *na*, surtout dans le cas de complémentation d'un nom verbal avec un nom ayant précédemment eu la fonction COD ;
- enfin, les constructions dans lesquelles l'objet direct est déplacé en tête de phrase contiennent aussi un marquage que l'on peut observer par la présence d'un pronom clitique qui précède le verbe à chaque fois que l'objet direct est spécifique.

Dans cet article, nous ne prétendons pas à une révision exhaustive du phénomène du MDO en macédonien, ni à une remise en cause des critères préétablis et amplement théorisés par nos prédécesseurs. Notre but est d'appliquer les études sur le MDO à des données nouvelles appartenant aux langues slaves du sud dont le groupe verbal contient des procédés qui fonctionnent selon le même mécanisme syntaxico-sémantique du MDO, mais avec d'autres marqueurs linguistiques. Cela va dans le sens des conclusions de Bossong (Bossong, 1998) qui montrent que le MDO n'est pas exclusivement nominal, mais qu'il peut se rattacher également au verbe.

Remerciements

Nous voudrions remercier nos collègues du programme « Unité et diversité dans le marquage différentiel de l'objet » de la Fédération « Typologie linguistique et universaux » qui nous ont motivé à entreprendre une étude de longue haleine sur le problème du marquage différentiel de l'objet en macédonien standard et dans ses dialectes, notamment Alexandru Mardale, Anton Antonov et Evangelia Adamou. Nous tenons également à remercier Patrice Pognan pour son encouragement à poursuivre les recherches sur ce phénomène que nous avons abordé dans notre travail de thèse de

manière tout à fait fortuite, mais nullement inintéressante.

Références

- ADAMOÛ E. (2009). « Le marquage différentiel de l'objet en *nashta* et en *pomaque* (Grèce). Retour sur l'hypothèse du contact ». In : *Bulletin de la Société de Linguistique de Paris*, n°1, vol.104, p.383-410.
- BOSSONG G. (1985). *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen : Gunter Narr.
- BOSSONG G. (1998). « Le marquage différentiel de l'objet dans les langues d'Europe », In : J. FEUILLET (ed.) (1998), *Actance et Valence dans les Langues de l'Europe*. Berlin/New York : Mouton de Gruyter, p.193-258.
- ČASULE I. (1989). *Sintaksa na makedonskata glagolska imenka*. Skopje : Studentski zbor.
- FRIEDMAN A.V. (1977). *The Grammatical Categories of the Macedonian Indicative*. Michigan : Slavica Publishers.
- FRIEDMAN A.V. (2002). *Macedonian* (coll. Languages of the World, 177). Munich : Lincom Europa.
- GUENTCHEVA Z. (1990). *Temps et aspect. L'exemple du bulgare contemporain*. Berne : Peter Lang.
- GUENTCHEVA Z. (1994). *Thématisation de l'objet en bulgare*. Berne : Peter Lang.
- KONESKI B. (1953). *Gramatika na makedonskiot literaturnen jazik*. Skopje : Prosvetno delo.
- LAZARD G. (2001). « Le marquage différentiel de l'objet ». In : M. HASPELMATH, E. KÖNIG, W. OESTERREICHER et W. RAIBLE (eds) (2001), *Language Typology and Language Universals*, vol.2, Berlin/New York : Mouton de Gruyter, p.873-885.
- MARDALE A. (2010). « Éléments d'analyse du marquage différentiel de l'objet dans les langues romanes ». In : *Faits de langues. Les Cahiers*, p.161-197.
- MINOVA-GJURKOVA L. (1990). *Sintaksa na makedonskiot standarden jazik*. (Štip, 2-ri Avgust) Skopje : Magor.
- PETRUŠEVSKA M. (2014). *Jazičnata politika i jazičното planiranje vo Republika Makedonija proizlezeni od Ustavot vo 1991 i od amandmanite na Ustavot od 2001 godina*. Skopje : Filološki fakultet « Blaže Koneski ».

À propos du marquage différentiel de l'objet et de 把 *bǎ* en chinois

Lin XIAO

Université Paris-Sorbonne – Paris 4 ; 1, rue Victor Cousin 75 005 Paris

Lacito/CNRS (UMR 7107) ; 7, rue Guy Môquet – 94 801 Villejuif

xlamy1118@gmail.com

RÉSUMÉ

On appelle marquage différentiel de l'objet le fait pour certains compléments d'objet directs (COD) du persan ou des langues romanes d'être marqués par une préposition, mais la notion peut être étendue à d'autres types de langues, dont le chinois. Mais en espagnol *a* est directionnel, tandis qu'en chinois *bǎ* (« prendre ») marque la manipulation de l'objet, ce qui correspond à la phase initiale du scénario. Nous proposons de ne pas isoler la construction de l'objet avec *bǎ* de sa construction verbale sérielle (CVS) d'origine en prenant en considération le fonctionnement typique des CVS, qui consiste à décomposer l'événement en sous-événements présentés dans un ordre souvent iconique.

ABSTRACT

We define 'differential object marker' the fact for some object complement to be marked by a preposition in Persian and in certain Romance languages, and this concept can be extended to other types of language, such as Chinese. For example, *a* in Spanish is a directional preposition, while *bǎ* ("take") in Chinese marks the manipulation of object, indicating the initial phase of scenario. I propose to consider this *bǎ* construction in relation to the serial verb construction (SVC) where was born *bǎ*, and I argue these SVC tend to decompose events into sub-events, which are often presented in an iconic order.

MOTS-CLÉS : construction verbale sérielle, marquage différentiel de l'objet, grammaticalisation, événement, iconique.

KEYWORDS: serial verb construction, differential object marker, grammaticalization, events, iconicity.

1 Introduction

C'est un fait bien connu que dans une partie des langues romanes certains objets directs peuvent être marqués par une préposition (Pottier, 1968). Des phénomènes similaires sont observés dans des langues de type très divers, d'où la notion de marquage différentiel de l'objet en typologie (Bossong, 1982 ; Lazard, 1982). Nous nous proposons de décrire ce phénomène dans la langue chinoise où il est bien attesté avec la construction en 把 *bǎ* + objet direct antéposé au verbe, sans perdre de vue la portée universelle du phénomène aussi bien que la diversité de ces manifestations (Wang, 1943 ; Lazard, 1994).

2 Double statut de 把 *bǎ* : verbe-préposition

Comme les autres prépositions du chinois moderne, issues des verbes, 把 *bǎ* garde son double statut de verbe-préposition (Lemaréchal, 1989), bien qu'il ait subi une évolution d'un état moins grammatical vers un statut plus grammatical, plus précisément, de verbe signifiant « prendre » marque de l'objet direct antéposé.

A. Quand 把 *bǎ* apparaît dans une construction verbale sérielle² (CVS) en y assumant le V1, il est généralement incompatible avec les marques d'aspect. Et, dans un des rares cas où 把 *bǎ* est suffixé par la marque d'aspect inaccompli 着 *zhe*, il subit un changement de sens de « prendre » à « garder » (ne pas laisser) :

(1) 小明把着巧克力不让别人吃

xiǎomíng bǎ -zhe qiǎokèlì bú ràng bié rén chī

NP garder Inacc. chocolat Nég. laisser autres personnes manger

« Xiaoming garde les chocolats et ne laisse pas les autres en manger. »

B. 把 *bǎ* ne s'utilise plus avec une valeur autonome que dans des expressions lexicalisées, comme, par exemple 把门 *bǎmén* « garder la porte », mais aussi dans des expressions nouvelles : 把妹 *bǎmèi* « draguer une fille ».

3 Degré de transitivité et emploi de 把 *bǎ*

Dans une phrase dont l'objet préverbal est marqué par 把 *bǎ*, les rapports sémantiques

² On peut définir d'une façon générale les constructions verbales sérielles (CVS) comme des constructions où deux ou plusieurs verbes se succèdent, sans connecteur ni pause prosodique, qui occupent une place intermédiaire entre coordination et subordination et entre phrase simple et phrase complexe, et qui se comporte comme une seule prédication complexe.

entre le verbe principal et l'objet jouent un certain rôle : le verbe désigne une action concrète, affectant maximalement son objet. Dans l'exemple ci-dessous, si l'on antépose l'objet de « boire-soupe » (ordre normal du chinois, en 2) avec 把 *bǎ*, « soupe » est défini et spécifique, la relation sémantique entre le verbe et l'objet est modifiée : la soupe a été bue en totalité, comme en 3 :

- (2) 他喝了汤了, 可是没喝完
tā hē -le tāng le, kěshì méi hē wán
 3sg. boire Acc. soupe PE mais Nég. boire Compl.rés.finir
 « Il a bu la soupe, mais il n'a pas tout fini. »
- (3) 他把汤喝了, *可是没喝完
*tā bǎ tāng hē -le, *kěshì méi hē wán*
 3sg. BA soupe boire Acc. mais Nég. boire Compl.rés.finir
 « Il a tout entièrement la soupe. »

Sauf erreur de notre part, il semble qu'on n'ait pas relevé un fait que nous croyons important : le prédicat principal de la phrase où figure un syntagme en 把 *bǎ* n'est jamais représenté par un verbe simple monosyllabique ; le verbe doit apparaître au moins sous une forme redoublée, ou, le plus souvent, être accompagné d'autres éléments :

- (4) 把衣服洗洗 (verbe redoublé)
bǎ yī-fu xǐ-xǐ
 BA vêtement laver-Rédup.laver
 « Lavez un peu les vêtements. »
- (5) 把门开着 (verbe + marque d'aspect)
bǎ mén kāi zhe
 BA porte ouverte Inacc.
 « Tenez la porte ouverte. »
- (6) 把门关上 (verbe + complément directionnel/résultatif)
bǎ mén guān shàng
 BA porte fermer Compl.dir.vers le haut
 « Fermez la porte. »
- (7) 你把这本书借给他 (verbe + complément datif)
nǐ bǎ zhè běn shū jiè gěi tā
 2sg. BA Dém. Cl. livre prêter Prép.dat.à 3sg.
 « Tu lui prêtes ce livre. »

- (8) 把他气得话都说不出来了 (verbe + complément d'appréciation)
bǎ tā qì -de = huà dōu shuō bù chū-lái le
 BA 3sg. indigner M. d'appréciation parole même dire Nég. sortir PE
 « (Cela) l'a indigné au point qu'il n'a plus pu rien dire. »

Il nous faudra expliquer cette contrainte.

Avec les verbes de déplacement, l'objet déplacé et les phases initiale/finale du déplacement sont bien définis; l'objet déplacé est toujours antéposé avec 把 *bǎ* : la phase finale est exprimée après le verbe de déplacement, au moyen d'un complément de lieu qui indique la position résultante :

- (9) 你应该把桌子放在这儿
nǐ yīnggāi bǎ zhuō-zǐ fàng zài zhè-r
 2sg. devoir BA table mettre Prép.à ici
 « Tu dois mettre la table ici. »

把 *bǎ* a d'abord été un verbe signifiant « prendre ». Et 把 *bǎ* N1 et *zài zhèr* « être ici » restent dans un ordre iconique mimant celui du déroulement des opérations dans le monde réel; ils deviennent des indicateurs du rôle que jouent chacun des participants de l'action, l'un une préposition marquant l'objet déplacé, et l'autre une préposition marquant la position résultante.

Dans les cas de figure passés en revue ci-dessus (1° V+MAsp, V+Compl. rés/dir ; 2° V+Compl.dat, V déplacement+lieu d'arrivé ; 3° V+ Compl.dit d'appréciation), l'élément postverbal oriente vers la phase finale du scénario. 把 *bǎ* présente l'objet à la phase initiale, au moment où il est introduit dans la situation.

Il faut ajouter que la construction 把 *bǎ* suffit à transitiver certains verbes intransitifs. (exemple emprunté à Iljic, 1987) :

- (10) 犯人跑了
fàn-rén pǎo le
 criminel fuir Asp.
 « Le criminel s'est enfui. »

- (11) 把犯人跑了
bǎ fàn-rén pǎo le
 BA criminel fuir Asp.
 « On a laissé s'enfuir le criminel. »

L'ajout de 把 *bǎ* change l'orientation de verbe dans les exemples présentés ci-dessus.

L'ancien sujet « criminel » de (10) devient l'objet emphatique de 把 *bǎ* en (11) dont la structure syntaxique change complètement : criminel devient le causataire avec le sens « le criminel a été laissé s'enfuir ». Le changement entre emplois transitif et intransitif n'est jamais aussi facile que dans une langue isolante où l'orientation de verbe n'est pas morphologiquement marquée et l'identification de sujet et d'objet sont ainsi litigieuses.

4 Paramètres sémantiques de l'objet marqué par 把 *bǎ*

Dans l'échelle d'individuation, le régime introduit par 把 *bǎ* est obligatoirement défini ou au moins spécifique, et le sujet agent [+contrôle]. Dans la perspective de l'énonciation, le régime de 把 *bǎ* est souvent constitué d'informations connues et mentionnées dans le contexte précédent.

En chinois, l'objet non marqué suit le verbe ; marqué par 把 *bǎ*, il le précède, et il est en général défini. Citons Iljic : « 把 *bǎ* présente toujours le nom qu'il introduit comme quelque chose de posé dans le contexte » (Iljic, 1987 : 250). Le fait que l'objet marqué par 把 *bǎ* soit obligatoirement défini ou spécifique est lié, semble-t-il, à sa topicalité :

(12) 请来了一位大夫
qǐng lái yī wèi dàifū
 inviter Compl.dir.venir Acc. un Cl. médecin
 « (On) a invité un médecin. »

(13) 把大夫请来了
bǎ dàifū qǐng lái le
 BA Médecin Inviter Compl.dir.venir Acc.
 « (On) a invité le médecin. »

La différence entre les deux exemples ci-dessus est une question de définitude du nom médecin. Le complément de l'objet focal en position postverbale est indéfini marqué par « un », *yī* + classificateur ; tandis que l'objet topical antéposé avec 把 *bǎ* est défini et est marqué par zéro.

Quand l'objet antéposé avec 把 *bǎ* est indéfini, il est nécessairement spécifique (exemple emprunté à Teng Shou-Hsin, 1989) :

(14) 我想给她介绍一个朋友
wǒ xiǎng gěi tā jièshào yī gè péngyǒu
 1sg. vouloir Prép.bén. 3sg. présenter un Cl. ami
 « J'aimerais bien lui faire rencontrer quelqu'un qui puisse devenir son ami. »

(15) 我想把一个朋友介绍给她

wǒ xiǎng bǎ yī gè péngyǒu jièshào gěi tā

1sg. vouloir BA un Cl. ami présenter Prép.dat. 3sg.

« J'aimerais bien lui présenter un de mes amis. »

Il est vrai que le syntagme nominal préverbal ou postverbal 一个朋友 *yī gè péng-yǒu* se présente sous une forme identique, mais sémantiquement, le syntagme nominal préverbal marqué par 把 *bǎ* doit être spécifique, tandis que le syntagme nominal postverbal désigne un ami potentiel.

En chinois ancien comme en chinois moderne, on a des cas où un même verbe (même caractère et même prononciation) peut avoir deux orientations contraires ; c'est le cas de 借 *jiè* prêter/emprunter :

(16) 我借他钱了

wǒ jiè tā qián le

1sg. prêter/emprunter 3sg. argent PE

« Je lui ai prêté/emprunté de l'argent. »

(17) 我借给他钱了

wǒ jiè gěi tā qián le

1sg. prêter Prép.dat.à 3sg. argent PE

« Je lui ai prêté de l'argent. »

(18) 我把钱借他了

wǒ bǎ qián jiè tā le

1sg. BA argent prête 3sg. PE

« Je lui ai prêté de l'argent. »

Pour désambiguïser la phrase (16), on peut ajouter la préposition 给 *gěi* marquant le datif, mais 把 *bǎ* peut également désambiguïser en soulignant le contrôle de l'agent sur l'objet : le prêteur, et non l'emprunteur, a le contrôle sur l'objet en phase initial du procès. Ce qui est en cause n'est pas le contrôle de l'agent sur le procès (*emprunter/prêter* est [+contrôle], voir, *emprunte/prête-lui de l'argent, je l'ai persuadé d'emprunter/prêter de l'argent, il a promis d'emprunter/prêter de l'argent*), mais sur l'objet (ce qui correspond tout à fait à la notion de *disposal form* des linguistes chinois). L'ordre des mots reste iconique et la construction syntaxique reste celle de la construction verbale sérielle d'origine.

5 把 *bǎ* + N1 + V2 + N2 : Verbe-préposition ou marquage différentiel de l'objet?

Vu les différents emplois de 把 *bǎ* et la multiplicité de facteurs hétérogènes qui les commandent, comment considérer la construction en 把 *bǎ* ? On la considère comme un cas de marquage différentiel de l'objet (Lazard, 1994), mais la notion de *disposal form* (Wang Li, 1943 ; Chao Yuenren, 1968) est sans doute plus conforme à la réalité des faits.

Nous proposons de replacer 把 *bǎ* dans son environnement d'origine, c'est-à-dire les constructions verbales sérielles, dont relève toujours ce morphème malgré son indéniable grammaticalisation.

La notion de marquage différentiel de l'objet ne cadre pas tout à fait avec les faits du chinois pour les raisons suivantes. Bien que les contraintes sémantiques de l'objet répondent aux critères de sélection de définitude ou detopicalité, l'emploi de 把 *bǎ* n'est pas le seul moyen dont dispose le chinois moderne pour mettre en valeur l'objet. Il faut noter que la position cadrative (19) ou le phénomène de sujet-patient (20) ont aussi pour objectifs de distinguer l'objet non-marqué et que ces différents moyens de distinction de l'objet ne fonctionnent pas de la même manière.

- (19) 桌子, 你应该放在这儿
 zhuō-zǐ, nǐ yīnggāi fàng zài zhè-r
 table 2sg. devoir mettre Prép.à ici
 « La table, tu dois la mettre ici. »

- (20) 桌子应该放在这儿
 zhuō-zǐ yīnggāi fàng zài zhè-r
 table devoir mettre Prép.à ici
 « La table se met ici. »

Le marquage différentiel de l'objet ne fonctionne pas de la même manière dans les différents types de langues. Ainsi, la préposition espagnole *a* est directionnelle et indique la visée sur l'objet : on place l'objet dans la phase finale de l'action, tandis que la préposition chinoise 把 *bǎ* marque la manipulation de l'objet (sens d'origine prendre) qui correspond à la phase initiale du scénario (Lemaréchal, 1998).

Les CVS V1 + N1 + V2 + N2 sont caractéristiques du chinois comme en général des langues isolantes : relevant de ce type, le chinois exploite plus le marquage séquentiel

que le marquage morphologique. Choisir de ne pas isoler la construction de l'objet avec 把 *bǎ* de sa construction verbale sérielle d'origine revient à prendre en considération le fonctionnement typique des constructions verbales sérielles, qui consiste à décomposer l'événement en sous-événements, *chunks of event* (Givon, 1990) présentés dans un ordre qui a tendance à être iconique.

6 Conclusion

D'un point de vue général, l'emploi de marquage différentiel de l'objet est commandé par une multiplicité de facteurs hétérogènes qui jouent ensemble d'une manière subtile: définitude, humanitude, thémativité, transitivité de verbe principal, et consistance des groupes syntaxiques, etc. Ces facteurs ne forment pas toujours le même ensemble et leur poids de considération varie d'une langue à l'autre. Spécifiquement, pour le chinois mandarin, le facteur dominant est le contrôle sur l'objet et l'individuation de l'objet transformé, mais on ne doit pas oublier que ces facteurs sont directement liés au sémantisme de 把 *bǎ* (orientation vers la phase initiale et individuation de son régime), hérité, malgré sa grammaticalisation, du verbe 'prendre' dont il est issu.

Références

- BOSSONG G. (1982). « Der präpositionale Akkusativ im Sardischen ». In : J. HUBSCHMID, O. WINKELMANN et M. BRAISCH (eds) (1982), *Festschrift für Johannes Hubschmid zum 65. Geburtstag: Beiträge zur allgemeinen, indogermanischen, und romanischen Sprachwissenschaft*. Bern : Francke, p.579-599.
- CHAO Y. (1968). *A Grammar of spoken Chinese*. California : University of California Press.
- ILJIC R. (1987). « À propos de *ta ba ge fuqin si-le* ». In : *Cahiers de Linguistique-Asie Orientale*, 16, p.237-257.
- LAZARD G. (1982). « Le morphème *ra* en persan et les relations actancielles ». In : *Bulletin de la Société de linguistique de Paris*, n°1, vol.77, p.177-208.
- LAZARD G. (1994). « Le *ra* persan et le *ba* chinois ». In : *Cahiers de linguistique Asie Orientale*, vol.23, p.169-176.
- LEMARÉCHAL A. (1998). *Études de morphologie en f(x...)*. Louvain-Paris : Peeters.
- LEMARÉCHAL A. (1989). *Les parties du discours sémantique et syntaxe*. Paris : PUF.

- LEMARÉCHAL A. (2014). « Marque d'agent et marque d'objet : mirages et réalités de la grammaticalisation en chinois ». In : *Communication 27e journées du CRLAO* (Paris ; 26 juin 2014).
- POTTIER B. (1968). « L'emploi de la préposition a devant l'objet en espagnol ». In : *Bulletin de la Société de Linguistique de Paris*, n°1, vol.63, p.83-95.
- RYGALOF A. (1973). *Grammaire élémentaire du chinois*. Paris : PUF.
- TENG S.H. (1989). « The semantics of causatives in Chinese ». In : *Journal of the Chinese Language Teachers Association*, p.227-244.
- WANG L. (1943). *中国现代语法 zhōngguó xiàndài yǔfǎ (Grammaire moderne chinoise)*. Beijing : The Commercial Press.

Stratégies de relativisation atypiques dans l'aire Caucase-Iran-Anatolie

Ophélie GANDON^{1,2}

(1) Université Sorbonne Nouvelle, 75 005 Paris

(2) MII, UMR 7528, 94 200 Ivry-sur-Seine

ophelie.gandon@gmail.com

RÉSUMÉ

L'objet de cet article est de présenter des stratégies de relativisation atypiques dans des langues parlées dans la région Caucase-Nord de l'Iran-Est de l'Anatolie. La région est connue pour concentrer une grande diversité génétique et typologique : une cinquantaine de langues réparties en six familles distinctes s'y côtoient (caucasique Nord-Est, caucasique Nord-Ouest, caucasique Sud, turcique, indo-européenne, sémitique). Les langues caucasiennes sont notamment connues pour présenter de nombreux faits linguistiques typologiquement rares tant sur le plan syntaxique que phonologique. L'article vise à montrer comment, en se concentrant sur cette aire, les universaux et tendances couramment observés au niveau de la relativisation sont remis en cause.

ABSTRACT

Untypical relativization strategies in Caucasus-Iran-Anatolia area

This paper aims to show untypical relativization strategies in languages spoken in the Caucasus-North Iran-East Anatolia area. This area is known for its high genetic and typological diversity: more than fifty languages are distributed over six families (North-East Caucasian, North-West Caucasian, South Caucasian, Turkic, Indo-European, Semitic). Caucasian languages are especially known for numerous typologically unusual linguistic facts, both phonological and syntactic. The paper shows how, when focusing on this area, the universals and tendencies commonly noticed about relativization are called into question.

MOTS-CLEFS : relativisation, stratégies atypiques, linguistique typologique, linguistique aréale, aire Caucase-Iran-Anatolie

KEYWORDS: relativization, untypical strategies, typological linguistic, areal linguistic, Caucasus-Iran-Anatolia area

1 Introduction

L'objet de cet article est de montrer combien les langues de la région Caucase-Nord de l'Iran-Est de l'Anatolie peuvent diverger des tendances et universaux généralement observés, au niveau de la relativisation. Cette zone est connue pour concentrer une grande diversité génétique et typologique : une cinquantaine de langues se répartissant sur six familles distinctes y sont parlées. Les plus nombreuses sont les langues caucasiennes, réparties en trois groupes, Nord-Est, Nord-Ouest et Sud, dont les liens de parenté génétique font l'objet de controverses. On y trouve également des langues indo-européennes (notamment des langues iraniennes, mais aussi autres comme l'arménien ou le grec pontique), des langues turciques (turc, azéri, karachay-balkar, nogay, koumyk, etc.), et des langues sémitiques représentées par le néo-araméen et des dialectes arabes, parlés dans le sud-est de la Turquie et l'est de l'Iran.

Il résulte de cette diversité génétique une grande diversité typologique. Les langues caucasiennes sont notamment connues pour présenter des caractéristiques très particulières sur le plan typologique : elles sont ergatives, pour la plupart agglutinantes, et peuvent avoir des systèmes très complexes, tant au niveau phonologique que syntaxique. Certaines propriétés de ces langues, de mieux en mieux documentées par des grammaires récentes (laz, tabassaran, hinuq, etc.³), sont souvent en contradiction avec les tendances jusque-là observées dans les langues plus connues et présentent de ce fait un grand intérêt pour l'étude de la typologie et des universaux.

On se propose ici de montrer quelques phénomènes atypiques au niveau de la relativisation. On entend ici par relative tout constituant phrastique qui modifie un nom en restreignant son champ référentiel. Le constituant phrastique est subordonné à une phrase matrice, et le nom modifié a un rôle syntaxique à la fois dans la matrice et la relative. On reprend ici la terminologie de Keenan (1985) et Creissels (2006), et on appelle le nom ainsi modifié « nom de domaine » (traditionnellement désigné comme « nom tête » ou « antécédent »).

De nombreux paramètres syntaxiques peuvent varier entre les constructions ainsi désignées, ce qui donne lieu à diverses stratégies de relativisation (d'une langue à une autre ou au sein d'une même langue) : la position de la proposition relative par rapport à la phrase matrice et/ou le nom de domaine, son enchâssement ou non dans phrase

³ Respectivement : Lacroix 2009, Babaliyeva 2013, Forker 2013.

matrice⁴, la présence ou non d'un pronom relatif, relativiseur ou subordinateur plus général pour introduire la proposition relative, la finitude de la forme verbale dans la proposition relative, la position canonique du terme relativisé à l'intérieur de la proposition relative (est-elle occupée par le nom de domaine lui-même, un pronom résomptif, ou laissée vide), etc.

La diversité typologique des langues de la région étudiée se retrouve au niveau de la relativisation puisqu'on y trouve quasiment la totalité des stratégies existantes à l'échelle mondiale (Creissels, 2005). On y trouve aussi un certain nombre de stratégies atypiques, objet de cet article, qui sont en contradiction avec des universaux ou tendances couramment proposés.

On reprend ici la terminologie de Hawkins (assez similaire aux autres), qui distingue trois types d'universaux : implicationnels, distributionnels et absolus (Hawkins, 1988 : 4-5). L'universel implicationnel définit une corrélation entre deux paramètres : si une langue a telle propriété, alors elle aura aussi telle propriété. L'universel distributionnel repose sur la fréquence d'un phénomène : les langues de type A sont plus fréquentes que les langues de type B. Enfin, l'universel absolu détermine une propriété que toutes les langues possèdent (ou ne possèdent pas), indépendamment des autres propriétés de cette langue : l'exemple couramment cité est que toutes les langues ont des voyelles.

Bien sûr, avec la description de plus en plus massive de langues, un nombre croissant de contre-exemples aux universaux établis depuis Greenberg émerge et il s'agit aujourd'hui souvent plus de tendances que d'universaux.

Pour chaque type d'universel/tendance, on propose des contre-exemples de langues de l'aire délimitée Caucase-Iran-Anatolie. On montre ainsi qu'en se concentrant sur cette seule zone, les tendances sont très différentes.

2 Universaux implicationnels

2.1 Relatives prénominales à forme verbale finie

Les relatives prénominales font typiquement usage de formes verbales non finies

⁴ La relative enchâssée est un constituant de la phrase matrice qui prend place à l'intérieur de celle-ci (Lehmann 2010 : 3). La subordination n'implique pas l'enchâssement : on peut avoir subordination avec détachement, de même qu'une phrase simple peut avoir des constituants nominaux ou adpositionnels en position détachée (Creissels 2006b : 190).

(Keenan, 1985 : 160 ; Dik, 1997 vol.2 : 55 ; Andrews, 2007 : 208 ; De Vries, 2001 : 235 ; Lehmann, 2010 : 8, *inter alia*). Les formes verbales non finies ont une morphologie réduite avec une flexion moins différenciée, notamment pour ce qui est du marquage temps-aspect-mode et de la personne.

Le laz (caucasique Sud) illustre un contre-exemple de langue faisant usage de relatives prénominales à forme verbale finie (aucune réduction) comme stratégie dominante dans l'aire concernée (Lacroix, 2009a : 749, 2009b) :

(1) Laz (Lacroix, 2009a : 752)

[*ma-na ti me-m-o-k'vat-ase*] *k'oçi ham dunya-s var ren*
 1S-SUB tête PV-III-VAL4-couper-FUT.I3S homme DEM monde-DAT NEG être.I3S
 « L'homme qui me coupera la tête n'est pas de ce monde »

2.2 Relatives prénominales à pronom résomptif

Une autre tendance implicationnelle couramment citée est l'usage rare de pronoms résomptifs avec les relatives prénominales (Keenan, 1985 : 148-149 ; Dik, 1997b : 46 ; Creissels, 2006 II : 239, 242, *inter alia*). On parle de pronom résomptif lorsque la position canonique du nom de domaine dans la relative est occupée par un pronom ou indice pronominal, ce pronom ou indice pronominal pouvant aussi apparaître dans une phrase indépendante comme renvoyant à un référent du contexte (Creissels, 2006 II : 211).

Plusieurs langues à relatives participiales prénominales peuvent faire usage de pronoms résomptifs dans l'aire étudiée : au moins cinq dans la famille caucasique Nord-Est – Lezgi (Haspelmath, 1993 : 342), khwarshi (Khalilova, 2009 : 356), icari dargwa (Mutalov et Sumbatova, 2003 : 185), hinuq (Forker, 2013 : 552, 553) et hunzib (Van den Berg, 1995 : 131-132). L'azéri d'Iran (turcique) présente aussi une telle situation :

(2) Azéri d'Iran (Dehghani, 2000 : 284)

- a. *män or-da gal-yan bag*
 I there-LOC stay-PTCP garden
 « The garden where I stayed »
- b. *män on-učün oku-yan kiši*
 I 3SG-BEN sing-PTCP man
 « The man whom I sang for »

Les pronoms résomptifs sont fréquemment utilisés avec les stratégies de relativisation

postnominales, notamment lorsque la position syntaxique relativisée est basse sur la hiérarchie d'accessibilité (Keenan et Comrie, 1977 : 92 ; Comrie, 1989 : 23 ; Creissels, 2006 : 216, 225, *inter alia*). L'explication est probablement fonctionnelle car ce sont des rôles syntaxiquement moins liés au verbe (non nucléaires, obliques) et dont le rôle sémantique peut être plus difficile à interpréter. Inversement, les pronoms résomptifs sont rares dans le cas de relativisation de l'objet et extrêmement rares avec la relativisation du sujet (« The only languages we know of which regularly present subject NP RELS as pronouns are Urhobo and Yidish », Keenan, 1985 : 147). Dans ce cadre-là, le hunzib, qui peut faire usage de pronoms résomptifs avec ses relatives participiales prénominales, présente un cas encore plus atypique dans la mesure où les pronoms résomptifs sont possibles même avec la relativisation du sujet (« The S, A or Exp of the relative clause can be rendered by a reflexive pronoun⁵ », Van der berg, 1995 : 131).

(3) Hunzib (Van der Berg, 1995 : 132)

(*ʒu*) *ø-ut'-uru suk'u-u maʎu n-ac'ə-r*

(self) I-sleep-PTCP man-DAT dream/V V-see-PAST

« The man who slept saw a dream »

2.3 Position du relativiseur

La marque de subordination, s'il y en a une, est généralement placée à la marge (gauche ou droite⁶) de la relative, ou attachée au verbe (Downing, 1978 : 390, 396 ; Creissels, 2006b : 195-196). Plusieurs langues dans l'aire concernée présentent un fonctionnement différent.

En laz (caucasique Sud), dont un exemple est cité en (1), la marque de subordination *-na* s'attache enclitiquement à un des constituants précédant le verbe (il n'est pas toujours évident de déterminer lequel), et proclitiquement au verbe si aucun constituant ne le précède dans la relative (Lacroix, 2009 : 751-756).

Les langues caucasiennes du groupe Nord-Ouest illustrent un autre cas assez atypique de relativiseur. Un relativiseur *-z(e)- (-d(ə)-/-t-* pour l'oubykh) prend la place de l'indice pronominal du rôle relativisé, qui correspond donc à une position préverbale dans le

⁵ S = subject of intransitive verb, A = agent, Exp = Experiencer.

⁶ Généralement à gauche pour les relatives postnominales, et à droite pour les prénominales.

cas de la relativisation d'un rôle nucléaire (ergatif ou datif⁷, exemple 4); mais qui peut aussi être attaché à une adposition ou un nom dans le cas de la relativisation d'un oblique ou d'un génitif (respectivement exemples 5a et 5b) :

(4) Ubykh (Charachidze, 1989 : 441)

- a. [za-mà ø-sə-də-t'ə-n] tət
ART-apple (3SG-)1SG-REL-give-PRES man
« The man who gives me an apple »
- b. [za-mà ø-də-s-t'ə-n] tət
ART-apple (3SG-)REL-1SG-give-PRES man
« The man to which I give an apple »

(5) Abkhaz (Hewitt, 1987 : 202)

- a. zə-da w-aa-z a-wayoð də-z-dər-wa-yt'
REL-without 2SG-come-VNF DEF-man 3SG-1SG-know-DYN-FIN(PRES)
« I know the man without whom you came »
- b. sarə zə-ko'ət'ə` Merəb (ə)yə-zá-z a-phoə's də-z-dər-wa-yt'
I REL-chicken Merab it-3SG-steal-VNF DEF-femme 3SG-1SG-know-DYN-FIN(PRES)
« I know the woman whose chicken Merab stole »

Le relativiseur s'apparente à un pronom relatif, car de par sa position il indique le rôle relativisé. Lander le qualifie d'ailleurs de pronom relatif (Lander 2010, 2014). Si on le considère ainsi, on a alors avec les langues caucasiques Nord-Ouest violation de l'universel selon lequel les pronoms relatifs ne sont pas compatibles avec les relatives prénominales (Downing, 1978 : 392, 396 ; Keenan, 1985 : 149 ; Dik, 1997b : 46 ; De Vries, 2001 : 235 ; Andrews, 2007 : 208, 218, *inter alia*). Toutefois, si l'on s'en tient à la définition classique du pronom relatif, celui-ci n'apparaît pas dans la position canonique du terme relativisé mais doit être déplacé à la marge de la relative (Downing, 1978 : 385 ; Comrie, 1989 : 150 ; Creissels, 2006 II : 228, *inter alia*), ce qui n'est pas le cas ici.

Le relativiseur s'apparente en fait aussi à un pronom résomptif, puisqu'il apparaît à la place de l'indice pronominal du rôle relativisé. Mais il ne varie pas en personne et n'apparaît pas en phrase indépendante, ce qui empêche de le qualifier de pronom

⁷ Pour la relativisation de l'absolutif, la stratégie diffère un peu : l'indice pronominal reste inchangé, sauf pour l'abkhaz où il est remplacé par le préfixe -y(ə) (identique en forme avec l'affixe personnel absolutif de 3e personne).

résomptif. Wu le caractérise ainsi de mi-pronom relatif mi-pronom résomptif, sans être ni l'un ni l'autre (Wu, 2011 : 52).

2.4 Position de la relative et ordre des constituants

Downing propose une corrélation entre l'ordre des constituants et la position de la relative par rapport au nom de domaine : les langues plaçant le verbe en position finale font généralement usage de relatives prénominales (Downing, 1978 : 375, 389) ; inversement les langues pour lesquelles le verbe précède l'objet font usage de relatives postnominales (ibid : 383) – il s'agit de l'ordre le plus neutre pour les langues à ordre des constituants flexible.

Le persan, langue à verbe final faisant majoritairement usage de relatives postnominales, est un contre-exemple couramment cité. Dans la même famille on peut y ajouter le kurde kurmandji (Rizgar, 2005 : 19, 210), et toujours dans la même aire le kashkay, langue turcique parlée en Iran. Le kashkay est une langue à verbe final (typique des langues turciques) dont la stratégie de relativisation principale est une relative enchâssée postnominale introduite par le complémenteur *ki*, résultat très probable de l'influence du persan (Dolatkhah, 2012 : 189-190). Enfin, l'udi (caucasique Nord-Est) constitue également un autre contre-exemple : comme les autres langues caucasiennes Nord-Est, c'est une langue à verbe final ; mais qui comme stratégie dominante fait usage de relatives postnominales introduites par un pronom relatif (Schulze-Fürhoff, 1994 : 481, 502-503). Il s'agit là encore probablement du résultat d'influence(s) ; selon Schulze-Fürhoff, la syntaxe de l'udi aurait subi de profondes modifications sous l'influence de langues voisines, d'abord l'arménien ancien puis le persan et l'azéri (ibid : 488). L'arménien oriental aussi fait usage, entre autres, de relatives postnominales introduites par un pronom relatif (Dum-Tragut, 2009 : 478). Il peut être intéressant de noter également la similarité dans la construction du pronom relatif avec des langues de la famille Caucasique Sud : un pronom interrogatif fléchi + une particule, *-te* en udi, *-t(i)* et *-c(a)* en mégrélien et en géorgien (Hewitt, 2001 : 107 ; Harris, 1991 : 332-333).

2.5 Autres

Les cas atypiques dans les langues de l'aire étudiée contraires à des tendances ou universaux avancés dans la littérature peuvent encore être multipliés.

Lehmann associe la nominalisation de la forme verbale à l'accessibilité moindre des

différentes positions syntaxiques à la relativisation : plus la forme verbale est nominalisée, moins on aura de positions syntaxiques accessibles à la relativisation (Lehmann, 2010 : 8). Ceci est peut-être valable pour les langues indo-européennes, pour lesquelles les relatives à formes verbales non finies constituent une stratégie secondaire (par exemple les participiales du français). Mais ce n'est pas du tout le cas dans les langues turciques et caucasiques Nord-Est, où les relatives participiales constituent la stratégie dominante⁸ et peuvent relativiser un large éventail de positions syntaxiques⁹.

De Vries avance l'impossibilité pour les relatives prénominales d'avoir une fonction sémantique appositive (non restrictive) (De Vries, 2001 : 238). Les relatives prénominales des familles turciques, caucasiques Nord-Est et Nord-Ouest peuvent avoir une fonction restrictive comme non restrictive, sans marquer de différence morphosyntaxique, intonative ou autre¹⁰. En laz (caucasique Sud), les relatives prénominales qui sont la stratégie dominante peuvent aussi avoir une fonction restrictive comme non restrictive (Lacroix, 2009a : 755-758).

Toujours avec les relatives prénominales, lors de la relativisation d'un complément d'une adposition, Keenan dit ne pas connaître d'exemple où l'adposition ne disparaîtrait pas (« we have no clear cases where a prenominal RC gaps the object of a postposition leaving the postposition itself stranded in SREL. », Keenan, 1985 : 154). C'est effectivement ce qui est généralement observé avec les relatives prénominales. L'exemple 5a en abkhaz ci-dessus en est toutefois un contre-exemple. Il serait intéressant d'élargir à d'autres exemples pour voir s'il s'agit ici juste d'une exception, ou si la rétention des adpositions est la norme en abkhaz.

Plusieurs universaux et tendances implicationnel(le)s sont donc contredits par les langues de l'aire Caucase-Iran-Anatolie. Ces langues sont aussi en contradiction avec les tendances distributionnelles observées à l'échelle mondiale.

⁸ À l'exception du kashkay et de l'udi, comme on l'a vu précédemment.

⁹ La position syntaxique pouvant parfois donner lieu à des restrictions, notamment pour les langues caucasiques Nord-Est, étant le possesseur : en ingush par exemple, pour la possession seuls les termes en relation avec un membre de la famille ou un membre du corps peuvent être relativisés (Nichols 1994 : 131).

¹⁰ Entre autres : Hewitt (1989 : 72) pour l'abkhaz (caucasique Nord-Ouest), corpus dans Ergönenç Akbaba (2009 : 257) pour le turc et le nogay (*babasını öldüren Orkan*), et pour les langues caucasiques Nord-Est : Nichols (1994 : 131) pour l'ingush, Magomedova (2004 : 59) pour le Chamalal, Sumbatova et Mutalov (2003 : 186) pour l'icari dargwa, Haspelmath (1993 : 343) pour le lezgi, etc.

3 Universaux distributionnels

Les grands types de stratégies de relativisation n'ont pas la même fréquence à l'échelle mondiale. Dans les études typologiques, les relatives enchâssées postnominales sont de loin les plus courantes (De Vries, 2001 : 235 ; Creissels, 2006 : 223). Dryer (2013) dans *The World Atlas of Language Structures* (WALS), qui se base sur un échantillon de 824 langues dans le monde, obtient la distribution suivante :

postnominales (579) > prénominales (141) > mixtes (64) > relatives à nom de domaine interne (24) > détachées (15) > double tête (1)

La moins grande fréquence des relatives prénominales s'explique probablement par le coût cognitif : une relative prénominale nécessite la mémorisation de l'ensemble de la proposition relative avant de pouvoir accéder au nom qu'elle modifie.

Pourtant, si l'on se concentre sur la seule zone Caucase-Iran-Anatolie, la distribution devient très différente. Les relatives enchâssées prénominales sont largement représentées et sont très certainement plus courantes que les postnominales : elles sont la stratégie dominante dans les familles turcique, caucasique Nord-Est et Nord-Ouest, la famille caucasique Nord-Est étant la plus nombreuse et comptant à elle seule plus de trente langues (Chirikba, 2008 : 5-6).

4 Universaux absolus

En plus de contredire certaines tendances implicationnelles et distributionnelles, les données de langues de l'aire Caucase-Iran-Anatolie remettent aussi en cause l'universalité même de la catégorie des relatives. En effet, plusieurs langues de la famille caucasique Nord-Est (Daniel et Lander, 2010 ; Comrie, Khalilova et Forker, 2013) ainsi que le karachay-balkar pour la famille turcique (Comrie et Khalilova, 2013) ne disposent pas d'une construction spécifique aux relatives, mais font usage d'une construction plus générale dont la seule fonction est d'indiquer une relation de subordination (appelée « Generalized Noun Modifying Clause Construction », désormais GNMCC).

L'interprétation comme relative, complétive ou autre dépend de facteurs essentiellement sémantico-pragmatiques :

- (6) Hinuq (caucasique Nord-Est, Comrie *et al.*, 2013 ; Forker, 2013 : 555-556)
- a. [*ked-i r-u-ho goła]* *xok'o-be*
 girl-ERG NHPL-make-ICVB be.PTCP khinkal-PL
 « the khinkal (a kind of dumpling) that the girl made »
- b. [*uži: mecxer b-ik'ekko goła]* *xabar*
 boy.ERG money(III) III-steal.ICVB be.PTCP story
 « the news (or story) that the boy stole the money »
- c. [*de mašina toł-o goła]* *mecxer*
 me.ERG car sell-ICVB be.PTCP money
 « the money from my selling the car »
- d. [*() obu Ø-uh-o goła]* *ked y-a:-ho*
 GENI father(I) I-die-ICVB be.PTCP girl(II) II-yell-PRES
 i. 'The girl whose father died is crying'
 ii. 'When the father dies the girl cries'

Ce type de constructions avait déjà été décrit notamment pour le japonais (Matsumoto, 1997), ainsi que certaines langues asiatiques (Comrie, 1996, 1998), où elles sont présentes.

On retrouve des faits très similaires en nogay et en koumyk (turcique) (-*yan* et -*gen* sont les allomorphes d'un même morphème -*Gan*) :

- (7) Nogay (Ergönenç Akbaba, 2009 : 183, 247, 185)
- a. [*qızar-yan]* *yüz-i*
 rougir-PTCP visage-POSS.3SG
 « son visage qui rougit »
- b. [*sıñar bir quday alla bol-yan]* *yañi din-ge de iynan-ıp tur-di*
 seul un dieu Allah être-PTCP nouveau religion-ALL aussi croire-CVB AUX-PF
 « Il a aussi cru à la nouvelle religion selon laquelle il n'y a qu'un seul dieu Allah »
 litt. « la religion que il n'y a qu'un seul dieu »
- c. [*"Al yuldız-lar bel-in-de"* *de-gen]-di qaytiğ anlay-sız?*
 take star-PL taille-POSS2/3SG-LOC say-PTCP-ACC how understand-2PL
 « Comment tu comprends la phrase "Prends les étoiles dans ta/sa taille" ? »
 litt. « Comment tu comprends le fait de dire »

Cependant, en nogay, il est fréquent d'avoir une marque de possession co-référente au sujet de la subordonnée (lorsqu'il y en a un) sur le participe, ou sur le nom de domaine

dans le cas d'une relative avec nom de domaine¹¹. Il semble que cette marque soit plus ou moins facultative (à l'inverse du turc où elle est beaucoup plus systématique) ; des recherches plus poussées seraient nécessaires pour comprendre si l'on a effectivement à faire à des GNMCCs ou non. On observe des faits similaires en koumyk (Pekacar, 2007 : 996-997), pour lequel des exemples n'ont pu être insérés faute de place.

5 Conclusion

Les langues de la région Caucase-Iran-Anatolie fournissent donc des contre-exemples à nombreux universaux et tendances généralement observés au niveau de la relativisation, qu'ils soient implicationnels, distributionnels ou absolus.

Sur le plan des tendances implicationnelles, les relatives prénominales peuvent faire usage de formes verbales finies et de pronoms résomptifs, et peuvent avoir une fonction non restrictive ; le relativiseur, s'il y en a un, peut avoir une position autre qu'à la marge de la relative ou attaché au verbe. Plusieurs langues à verbe final de la région font usage de relatives postnominales. Enfin, la nominalisation de la forme verbale n'implique pas une restriction dans l'accessibilité à la relativisation. Sur le plan distributionnel, les relatives prénominales sont largement majoritaires sur les postnominales. Enfin sur le plan absolu, l'universalité même de la catégorie des relatives est remise en cause.

Les tendances ne sont donc plus du tout les mêmes lorsque l'on se concentre sur la seule zone Caucase-Iran-Anatolie, pourtant très riche génétiquement et typologiquement. Comment peut-on expliquer une telle divergence ? Il est probable que les nombreux phénomènes de contact de la région, où la plupart des gens sont bilingues voir trilingues ou plus, aient un rôle à jouer dans le développement de stratégies atypiques. Le fait que l'on ait des relatives postnominales en kashkay et en udi par exemple, langues à verbe final, alors que les autres langues de ces deux familles (respectivement turcique et caucasique Nord-Est) utilisent majoritairement des prénominales, est probablement le résultat de contact avec des langues faisant usage de postnominales.

Les données de l'aire nous montrent donc l'importance de prendre en compte un très grand nombre de langues aussi variées que possible pour la recherche d'universaux et tendances. Mais elles nous montrent aussi et surtout les limites des universaux : la description de plus en plus massive de nouvelles langues auparavant ignorées ne cesse

¹¹ Dans le cas d'une relative libre, la marque de possession est sur le participe.

de fournir de nouveaux contre-exemples aux universaux et tendances jusque-là observés. On peut ainsi se demander s'il resterait encore quelques universaux ou même tendances viables, si l'on avait la possibilité de prendre en compte les 6000 langues existant sur la planète, dont la majorité reste à décrire.

Abréviations¹²

1/2/3 : 1ère/2e/3e pers.	ERG : ergatif	PRES : présent
I/II/III/IV/V : classe nominale	FIN : fini	PST : passé
ABS : absolutif	FUT : futur	PTCP : participe
ACC : accusatif	ICVB : converbe imperfectif	PV : préverbe
ALL : allatif	LOC : locatif	REL : relativiseur
BEN : bénéfactif	NEG : négation	S : indice sujet
CVB : converbe	NHPL : pluriel non humain	SG : singulier
DAT : datif	OBL : oblique	SUB : subordonateur
DEF : défini	PAST : passé	VAL1/2/3 : opérateur de valence
DEM : démonstratif	PL : pluriel	VNF : forme verbale non finie
DYN : dynamique	POSS : possessif	

Références

ANDREWS A.D. (2007). « Relative clauses ». In : T. SHOPEN (ed.) (2007) *Language typology and syntactic description, vol. II : complex constructions*. Cambridge : Cambridge University Press, p.206-236.

BABALIYEVA A. (2013). *Études sur la morphosyntaxe du tabasaran littéraire*. Thèse de doctorat en Histoire, Textes et Documents, École Pratique des Hautes Études, Paris.

CHARACHIDZÉ G. (1989). « Ubykh ». In : B.G. HEWITT (ed.) (1989), *The indigenous languages of the Caucasus, Vol. 2 : The Nord West Caucasian languages*. Delmar/New York : Caravan Books.

CHIRIKBA V.A. (2008). « The problem of the Caucasian Sprachbund », In : P. MUYSKEN (ed.) (2008) *From Linguistic Areas to Areal Linguistics*. 25-93, Amsterdam/Philadelphia : John Benjamins Publishing Company.

COMRIE B. (1989). *Language universals and linguistic typology*. Oxford : Basil Blackwell.

¹² Parmi les exemples cités, les gloses de certains ont pu être légèrement modifiées pour uniformiser, et les exemples non glosés ont été glosés.

- COMRIE B. et Z. KHALILOVA (2013). « Generalized noun modifying clause constructions in the North Caucasus ». In : *Societas Linguistica Europaea – 46th Annual Meeting* (Split, Croatie ; 18-21 septembre).
- COMRIE B., Z. KHALILOVA et D. FORKER (2013). « GNMCCs in Bezhta and Hinuq (Nakh-Daghestanian) ». In : *10th Biennial Conference of the Association of Linguistic Typology* (Leipzig ; 15-18 août).
- CREISSELS D. (2005). « Typologie de la relativisation et données de l'aire Caucase-Iran-Anatolie ». In : *Atelier de syntaxe arménienne* (Pithiviers ; 23-25 mai).
- CREISSELS D. (2006). *Syntaxe générale : une introduction typologique, 2 volumes*. Paris : Lavoisier.
- DANIEL M. et Y. LANDER (2010). « A girl of word, meat of a ram, and a life of longing. On peculiar cases of relativization in East Caucasian languages ». In : *Conference Syntax of the World's Languages IV* (Lyon ; 23-26 septembre).
- DEGHANI Y. (2000). *A grammar of Iranian Azari (including comparisons with Persian)*. Munich : Lincom Europa.
- DE VRIES M. (2001). « Patterns of relative clauses ». In : *Linguistics in the Netherlands*, vol.18, p.231-243.
- DIK S. (1997) *The theory of Functional Grammar*. Berlin : Mouton De Gruyter.
- DOLATKHAH S. (2012). *Elements for a grammar of Kashkay (a Turkic language of Iran)*. Thèse de doctorat en linguistique, École Pratique des Hautes Études, Paris.
- DOWNING B.T. (1978). « Some universals of relative clause structure », In : J. GREENBERG, C.A. FERGUSON et E.A. MORAVCSIK (eds) (1978), *Universals of human language, vol.4 : Syntax*. Stanford : Stanford University Press, p.375-418.
- DRYER M.S. (2013). « Order of Relative Clause and Noun », In : M.S. DRYER et M. HASPELMATH (eds) (2013) *The World Atlas of Language Structures Online*. Leipzig : Max Planck Institute for Evolutionary Anthropology. [URL : <http://wals.info/chapter/90>, consulté le 19/08/2014].
- DUM-TRAGUT J. (2009). *Armenian : Modern Eastern Armenian*. Amsterdam : Benjamins.
- ERGÖNENÇ AKBABA D. (2009). *Nogay Türkçesi Grameri*. Ankara : Grafiker Yayınları.
- FORKER D. (2013). *A grammar of Hinuq*. Berlin/Boston : De Gruyter Mouton.

- HARRIS A.C. (1991). « Mingrelian ». In : A.C. HARRIS (ed.) (1991), *The indigenous languages of the Caucasus, vol.1 : The Kartvelian languages*. Delmar NY : Caravan books, p.315-389.
- HASPELMATH M. (1993). *A grammar of Lezgian*. Berlin : Mouton de Gruyter.
- HAWKINS J.A. (1988). « Explaining language universals ». In : J.A. HAWKINS (ed.) (1988), *Explaining language universals*. Oxford/Cambridge : Blackwell, p.3-28.
- HEWITT B.G. (1987). *The typology of subordination in Georgian and Abkhaz*. Berlin : Mouton de Gruyter.
- HEWITT B.G. (2001). « Convergence in language change: morpho-syntactic patterns in mingrelian (and Laz) ». In : *Transactions of the Philological Society*, n°1, vol.99, p.99-144.
- KEENAN E.L. (1985). « Relative clauses ». In : T. SHOPEN (ed.) (1985), *Language typology and syntactic description, vol.2 : Complex constructions*. Cambridge : Cambridge University Press, p.141-170.
- KEENAN E.L. et B. COMRIE (1977). « Noun Phrase Accessibility and Universal Grammar ». In : *Linguistic Inquiry*, n°1, vol.8, p.63-99.
- LACROIX R. (2009a). *Description du dialecte laze d'Arhavi (caucasique du sud, Turquie), Grammaire et textes*. Thèse de doctorat en sciences du langage, Université Lumière Lyon 2.
- LACROIX R. (2009b). « Laz relative clauses in a typological and areal perspective ». In : P. AUSTIN, O. BOND, M. CHARETTE, D. NATHAN et P. SELLS (eds) (2009), *Language Documentation and Linguistic Theory 2*. London : School of Oriental and African Studies, University of London.
- LANDER Y. (2010). « Relativization in Shapsug Adyghe ». In : *Rice Working Papers in Linguistics*, vol.2, Spring 2010, p.75-91.
- LANDER Y. (2014). « On Apparent Specifics of Circassian Relativization ». In : *Linguistic variation and contact : Syntax of relativization in the languages of Western Asia (Iran, Caucasus, Anatolia, Near-East)*, (Paris ; le 7 février).
- LEHMANN C. (2010). « On the typology of relative clauses ». [URL : http://www.christianlehmann.eu/ubl/typology_relative_clauses.pdf, consulté le 22/11/2013].
- MUTALOV R.O. et N.R. SUMBATOVA (2003). *A grammar of Icarî Dargwa*. München : Lincom Europa.

- NICHOLS J. (1994). « Ingush ». In : R. SMEETS (ed.) (1994), *The Indigenous languages of the Caucasus, vol.4 : North east caucasian languages*. Delmar/New York : Caravan Books, p.79-145.
- PEKACAR Ç. (2007). « Kumuk Türkçesi ». In : A.B. ERCILASUN (ed.) (2007) *Türk Lehçeleri Grameri*. Ankara : Akçağ Yayınları, p.938-1008.
- RIZGAR B. (2005). *Uygulamalı Kürtçe dersleri*. Dersên Kurdî, Istanbul : Doz Yayınları.
- SCHULZE-FÜRHOFF W. (1994). « Udi ». In : R. SMEETS (ed.) (1994) *The Indigenous languages of the Caucasus, vol.4 : North east caucasian languages*. Delmar/New York : Caravan Books, p.447-514.
- VAN DEN BERG H. (1995). *A grammar of Hunzib*. München. Newcastle : Lincom Europa.
- WU T. (2011). *La relativisation prénominale*. Thèse de doctorat en sciences du langage, Paris 3.

Analyse multimodale et contrastive des interjections émergentes : le cas de *genre* en français et *as if* en anglais britannique

Guillaume JAUDHUIN

Université Sorbonne Nouvelle – Paris 3, PRISMES–SeSyLiA, 75 006 Paris
guillaume.jaudhuin@univ-Paris3.fr

RÉSUMÉ

Cette étude est la première étape d'une recherche doctorale sur l'analyse multimodale et contrastive des interjections émergentes *genre* en français et *as if* en anglais. Bien que ces deux interjections soient particulièrement proches et fréquemment utilisées chez les adolescents et les jeunes adultes, elles ont, à ce jour, suscité peu de recherches. Dans cet article, nous défendons l'idée que les interjections sont multimodales par nature, d'où la nécessité d'adopter une approche multimodale dans laquelle les ressources verbales, vocales, et posturo-mimo-gestuelles sont analysées. Ce travail pionnier permettra à la fois d'identifier la complexité des ressources et des procédés utilisés par les locuteurs, mais également d'établir un parallèle entre ces deux interjections relativement récentes. Notre recherche a également pour but d'étudier l'ensemble des stratégies multimodales utilisées par les locuteurs lors de l'utilisation de ces deux interjections en interaction.

ABSTRACT

A Multimodal and Contrastive Analysis of Emerging Interjections: The case of “genre” in French and “as if” in British English

This study is the first step in a doctoral research project presenting the multimodal and contrastive analysis of the emerging interjections “genre” in French and “as if” in English. Although these interjections share a considerable number of similarities and are regularly used by teenagers and young adults, so far little research has been conducted on them. In this paper, I argue that interjections are multimodal by nature, hence the necessity to adopt a multimodal approach in which speakers' verbal, vocal and visual resources are analysed. This innovative research will allow me to identify the complexity of how and in what way these expressions are used, as well as to draw parallels between these two relatively new interjections. This research also aims to study the range of multimodal strategies with which speakers accompany these two interjections in face-to-face interaction.

Mots-clés : multimodalité, analyse contrastive, interjections, genre, as if, corpus oral vidéo.

Keywords: multimodality, contrastive analysis, interjections, genre, as if, audiovisual corpus.

1 Introduction

Proportionnellement à leur présence dans le discours oral spontané, les interjections n'ont pas reçu, à ce jour, l'attention qui leur était due. Depuis longtemps, les interjections ont constitué un sujet d'étude très controversé car leur statut linguistique a toujours soulevé de nombreuses questions. L'interjection fait partie du discours pour les uns, d'un groupe spécifique pour les autres, ou est complètement passée sous silence. La principale raison est que, pour bien des chercheurs (voir, entre autres, Swiatkowska, 2000), l'indépendance syntaxique caractérisant l'interjection est un frein à la détermination de son sens. Pourtant, le contexte joue, à l'écrit comme à l'oral, un rôle fondamental dans la spécification de son sens.

L'une des originalités de cette étude est alors de proposer d'élargir la notion de « contexte » à d'autres ressources communicationnelles que les seules ressources verbales. En effet, lorsqu'un locuteur emploie une interjection en situation d'interaction, ses ressources suprasegmentales et posturo-mimo-gestuelles jouent elles aussi un rôle crucial dans la détermination du sens. Nous proposons ainsi d'élargir une définition du « langage » longtemps restée trop restreinte et négligeant les nombreuses ressources sémiotiques utilisées en interaction (Liddell, 2003). Partant de ce postulat, le langage est considéré comme un système linguistique intégrant parole et gestualité de manière organisée.

Notre travail de recherche, ancré, d'une part, dans l'étude des interactions orales et, de l'autre, dans l'étude des interjections, se focalise sur l'étude des deux interjections émergentes *genre* en français et *as if* en anglais. Une originalité de ce travail est de traiter ensemble ces deux phénomènes interjectifs relativement proches qui, avant notre mémoire de Master 2, n'avaient jamais été étudiés en regard l'un de l'autre. Ces deux phénomènes interjectifs récents sont présents dans la langue familière, plutôt dans la langue spontanée, telle qu'on la trouve à l'oral ou à l'écrit dans les blogs, sur les réseaux sociaux, les chats et les forums. Bien qu'ils soient très régulièrement utilisés par les adolescents et les jeunes adultes, ces deux marqueurs ont suscité peu de recherches en tant que phénomènes interjectifs à proprement parler.

L'interjection – multimodale par nature – est un élément fortement saillant de notre activité langagière. L'analyse d'un tel phénomène nécessite de l'expliquer en l'intégrant à une réalité plus large, d'où notre recours, en situation d'interaction, à une analyse

multimodale comprenant modalité verbale (syntaxe et organisation du discours), modalité vocale (prosodie et intonation) et modalité posturo-mimo-gestuelle (expressions faciales, mouvements, postures et gestes). Nous proposons alors une analyse intégrée de divers canaux d'expression pour rendre compte de la richesse, mais aussi de la complexité de ces deux phénomènes interjectifs récents, en essayant d'inclure toutes leurs facettes et leurs fonctions en contexte.

Cette étude se caractérise également par sa dimension contrastive, qui permet de faire ressortir les différences et les similarités entre ces deux interjections émergentes. La littérature consacrée aux interjections a très souvent mis de côté ou bien oublié les interjections émergentes. Ces dernières mériteraient pourtant d'être étudiées de façon systématique. Notre approche multimodale et contrastive permettra à la fois d'éclairer le statut de l'interjection grâce à la prise en compte du non-verbal, mais également de mettre en saillance la proximité linguistique des interjections émergentes *genre* et *as if*, et sera donc un travail pionnier dans le domaine. Partant de ce postulat, plusieurs questions dirigeront nos recherches actuelles : 1) les modalités verbales, vocales et posturo-mimo-gestuelles se synchronisent-elles lors de l'utilisation de ces interjections ? Si oui, la synchronisation de ces trois modalités d'expression favorise-elle la co-construction du sens en interaction ? et 2) à quel point les interjections *genre* et *as if* sont-elles proches (proximité linguistique) ?

Ce travail de recherche, mené en France sur un corpus de langues française et anglaise, vise également à approfondir le dialogue entre approches françaises et anglo-saxonnes. Bien que les scientifiques parlent parfois d'enjeux similaires en des termes différents, ils développent également chacun leurs analyses spécifiques. Une telle mise en perspective ne peut qu'être synonyme d'enrichissement pour l'étude multimodale et contrastive des interjections. De manière à rendre compte de la dimension multimodale et notamment visuelle de l'expression des interjections *genre* et *as if*, un corpus oral vidéo de discussions semi-guidées a été recueilli.

2 Précédentes recherches

Cette étude a été amorcée dans le cadre de nos mémoires de Master 1 et de Master 2, sous la direction de Madame le Professeur Aliyah Morgenstern. En nous appuyant sur les travaux de Fleischman (1999), Rosier (2002), Vladimirska (2013), Dufaye (2013, 2014) et Vigneron (à paraître) pour *genre* et de Quirk (1985), Pinson (2009) et

Andersen (à paraître) pour *as if*, nous avons pu retracer leur passage à la forme interjective, tout en faisant ressortir les emplois autour desquels ces deux interjections semblent s'être constituées, comme le montre la figure ci-dessous.

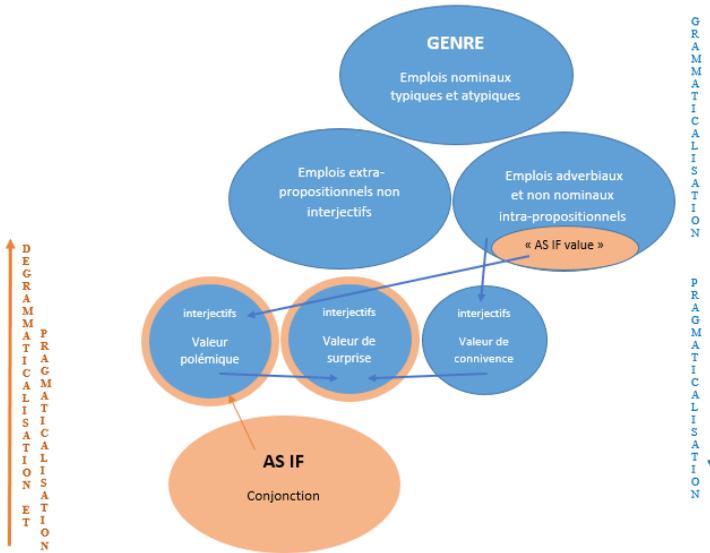


FIGURE 1 – Schéma récapitulatif des processus de (dé)grammaticalisation et de pragmaticalisation subis par les interjections émergentes *genre* et *as if*

2.1 Les valeurs de polémique, de surprise et de connivence

La valeur de polémique de ces deux interjections semble s'être formée autour de la valeur de comparaison de type conjonctif, comme dans *Il fait genre il veut pas venir (il fait comme si...)* et dans *He talks as if he had a potato in his mouth*. Le passage de *genre* + P à *genre* Ø et de *as if* + P à *as if* Ø fournit des informations explicites quant à la façon d'interpréter les attitudes du locuteur envers une proposition P, à savoir que le locuteur rejette fermement P, d'où le phénomène de pragmaticalisation. On retrouve la valeur de polémique dans les cas de postposition à une énonciation préalable qui, bien souvent, peut s'élargir à un concept, une idée, une image, etc. Ces deux interjections à valeur de polémique peuvent être utilisées en contexte dialogique, mais aussi en contexte monologique (avec un dialogisme sous-jacent) où le dialogisme peut être simplement simulé, au sens où l'énonciateur porte un regard critique sur ce qu'il vient de dire, comme dans *J'ai pas le temps en ce moment... Genre !* L'interjection *genre* à valeur de polémique est équivalente à *Style !*, *(Mais) bien-sûr !*, etc., et à l'arabisme

Zarma ! (Gadet et Ludwig, 2014). L'interjection *as if* à valeur de polémique est, quant à elle, équivalente à *As if it were true!* (Dunn, 2005), *Yeah right!*, *In your dreams!* (Andersen, à paraître), *I very much doubt that!*, *You must be joking!* ou *Some hope!* (Pinson, 2009).

La valeur de surprise, plutôt hybride, véhicule à la fois une proximité certaine entre un locuteur 1 et un locuteur 2, tout en apportant un certain ressenti vis-à-vis d'une proposition qui aura préalablement été prononcée. Nos études précédentes nous ont permis d'identifier cette valeur, ce qui, par conséquent, confirme que les interjections *genre* et *as if* sont encore en voie de (dé)grammaticalisation. La valeur de surprise ne se construit, bien entendu, qu'en contexte dialogique. Dans bien des cas, *genre* à valeur de surprise peut être substitué par *ah !* ou *ouah !*, et *as if* par *oh!* ou *whoa!*

La valeur de connivence, quant à elle, véhicule une proximité entre un locuteur 1 et un locuteur 2, à la manière d'un *tu vois !* Ainsi, on retrouve une implication du locuteur 2 qui participe, aux côtés du locuteur 1, à la ré-identification de cette situation type. Ce schéma implique trois sujets : le thème de la narration, un locuteur qui construit un point de vue par rapport à ce sujet, et enfin un locuteur 2 qui est censé adopter la position du locuteur 1, d'où découle la valeur de connivence qui, bien entendu, ne se construit qu'en contexte monologique avec dialogisme sous-jacent. Cependant, seule l'interjection *genre* semble pouvoir véhiculer cette valeur.

3 Corpus, protocole et transcription

Notre corpus de travail s'appuie sur des données extraites du *Paris and Canterbury Student Corpus* (PCSC), un corpus vidéo d'interactions orales dyadiques de même langue. Dans la partie française du corpus, on retrouve des étudiants scolarisés dans différentes universités parisiennes, tous locuteurs natifs de français et, dans la partie anglaise, des étudiants de l'Université du Kent à Canterbury, tous locuteurs natifs d'anglais britannique. Le *Paris and Canterbury Student Corpus* est un corpus original récent, collecté entre 2013 et 2014 par nous-même. Les participants ont entre 18 et 25 ans, et leurs conversations s'organisent autour d'une seule et même tâche, correspondant à notre protocole de recherche (conversation semi-guidée).

La totalité du PCSC sur laquelle se fonde notre travail de recherche est une collection de discussions à deux sur un thème très large, celui de la vie quotidienne. Le PCSC compte 6 conversations entre 12 participants (6 hommes, 6 femmes) qui discutent deux

par deux, dans la même langue, pendant 20 minutes, soit 2 heures de vidéo. Les 6 dyades comptent soit deux hommes, soit deux femmes, soit un homme et une femme. Les participants sont assis, se font face aux trois-quarts, et sont filmés par une seule caméra sur le côté. Au cours de chaque enregistrement, les participants, filmés par deux, piochent à tour de rôle des questions écrites sur des petits morceaux de papiers. Le candidat qui a pioché le morceau de papier doit poser la question à l'autre candidat et y répondre également avant de passer à la question suivante. Ils continuent ainsi pendant 20 minutes environ, jusqu'à ce que le chercheur les arrête.

Dans chaque paire, les deux participants sont amis et se connaissent donc bien. Nous avons fait ce choix pour deux raisons principalement. La première raison est que cela simplifiait le recrutement. La seconde raison est que la familiarité des locuteurs favorise considérablement la spontanéité des échanges, ce qui est un avantage crucial pour la qualité de notre corpus. Les questions que les candidats doivent se poser sont des questions que l'on se pose habituellement entre amis et suscitent des réponses spontanées (« Ta vie, tu la vois comment dans 10 ans ? », « Tu as eu quoi pour Noël ? », etc.), apportant des informations nouvelles qui, par la même occasion, peuvent faire réagir l'interlocuteur par rapport à ce qu'il connaît déjà de l'autre. Le protocole de collecte a été conçu de manière à ce que les participants se sentent tout à fait libres vis-à-vis des questions. Ils peuvent, par exemple, ne pas répondre à une question et passer directement à la suivante.

Nous avons ensuite élaboré un codage adapté aux objectifs de notre recherche, où les interjections *genre* et *as if* ont fait l'objet d'un ensemble de codages : la valeur des interjections et le contexte d'énonciation (dialogique ou monologique). Nous avons également étudié les doubles associations polémiques de type *ouais/hum/bah genre* et *yeah/hum as if*. Les modalités verbales et posturo-mimo-gestuelles ont été annotées sous ELAN. La modalité vocale a, quant à elle, été annotée à l'aide d'une convention de transcription proposée par Psathas (1995).

4 Résultats

Notre corpus comporte 22 occurrences (16 *genre* et 6 *as if* interjectifs, toutes valeurs confondues) au total. Dans le cas de *genre*, la valeur de polémique représente l'écrasante majorité (9 occurrences en contexte dialogique et 4 en contexte monologique = 13 occurrences au total, soit 81,25%) de l'ensemble des interjections

genre codées. La valeur de connivence en représente, quant à elle, 12,5% (2 occurrences, contexte monologique exclusivement) et la valeur de surprise, 6,25% (1 occurrence, contexte dialogique exclusivement). En ce qui concerne *as if*, la valeur de polémique représente l'écrasante majorité (5 occurrences en contexte dialogique et 0 en contexte monologique, soit 83,33%) de l'ensemble des interjections *as if* codées. La valeur de surprise en représente, quant à elle, 16,67% (1 occurrence, contexte dialogique exclusivement).

Pendant, seul un corpus de plus grande envergure nous permettra de réaliser des analyses approfondies et d'ajouter une dimension quantitative à ces recherches principalement qualitatives.

Les 22 occurrences présentes dans le corpus montrent que les ressources verbales, les variations de la voix et les ressources posturo-mimo-gestuelles se synchronisent lorsque les locuteurs utilisent ces deux interjections. Qui plus est, la synchronisation de ces trois modalités favorise la co-construction du sens en interaction, tout en contribuant à véhiculer les valeurs soit de polémique, soit de surprise, ou encore de connivence de l'interjection *genre* et les valeurs soit de polémique, soit de surprise de l'interjection *as if*.

5 Analyses

5.1 La valeur de polémique

En contexte dialogique, on retrouve la valeur de polémique des interjections *genre* et *if* dans des cas de postposition à une énonciation préalable. On remarque ici que les locuteurs utilisent un large éventail de ressources non-verbales lors de l'utilisation de ces interjections.



FIGURE 2 – *genre* et *as if* de polémique en contexte dialogique

Les ressources non-verbales communes aux deux interjections sont : le haussement des sourcils, les yeux fixés sur l'interlocuteur, la tête inclinée en avant ou en arrière, une

posture figée, et une emphase prosodique de durée variable avec intonation ascendante ou descendante. Nous avons par ailleurs remarqué que les énoncés (très) antérieurs que nos paires de locuteurs ont pu échanger semblent également jouer un rôle important quant à l'utilisation et l'intensité des ressources non-verbales. Lorsque *genre* et *as if* sont utilisés par les locuteurs, les modalités verbales, vocales et posturo-mimo-gestuelles se synchronisent et contribuent à véhiculer une valeur de polémique. On remarquera que les trois ressources non-verbales les plus saillantes pour exprimer cette valeur sont : le haussement des sourcils, l'inclinaison de la tête, et l'emphase prosodique. Dans certains cas, *as if* implique un schéma intonatif bien particulier (\uparrow *Yeah* \downarrow *as if*) qui rappelle les doubles affirmations polémiques du type \uparrow *Yeah* \downarrow *right*. À l'instar d'un \uparrow *ouais/hum* \downarrow *genre*, l'association *Yeah as if* semble elle aussi recréer iconiquement la dissociation des sources de prise en charge, primaire et secondaire.

En contexte monologique, seule l'interjection *genre* est présente dans notre corpus. Nous l'avons souligné précédemment, l'interjection *as if* à valeur de polémique peut également être utilisée en contexte monologique. Lorsque l'interjection *genre* à valeur de polémique est utilisée par un locuteur en contexte monologique, les ressources verbales, les variations de la voix et les ressources posturo-mimo-gestuelles se synchronisent et contribuent à véhiculer une valeur de polémique, à l'instar du *genre* dialogique.



FIGURE 3 – *genre* de polémique en contexte monologique

Les ressources non-verbales utilisées par les locuteurs sont les mêmes que pour les interjections *genre* et *as if* utilisées en contexte dialogique, à une exception près : l'intonation semble toujours être ascendante en contexte monologique.

5.2 La valeur de surprise

Les interjections *genre* et *as if* à valeur de surprise ne sont utilisées qu'en contexte dialogique dans des cas de postposition à une énonciation préalable. Lorsque les locuteurs utilisent ces interjections, une emphase prosodique lourde et de durée

variable avec intonation ascendante s'associe à des ressources posturo-mimo-gestuelles qui varient en fonction du degré de surprise exprimé par les locuteurs.



FIGURE 4 – *genre* et *as if* de surprise en contexte dialogique

Une fois encore, les trois modalités d'expression se mobilisent pour la co-construction du sens en interaction. Dans le cas de *as if*, on notera une intonation ascendante répliquée ($\uparrow as \uparrow if$).

5.3 La valeur de connivence

Lorsque les locuteurs utilisent l'interjection *genre* à valeur de connivence, une emphase prosodique légère et de durée très courte avec intonation descendante s'associe à un regard fuyant (rotation de la tête du côté opposé).



FIGURE 5 – *genre* de connivence en contexte monologique

Sans une approche multimodale, il serait bien difficile de faire la différence entre un *genre* à valeur de polémique utilisé en contexte monologique et un *genre* à valeur de connivence. À ce stade, nous souhaitons souligner que l'on ne peut se passer d'une approche multimodale lorsqu'on analyse de telles interjections.

6 Conclusion générale

Dans cette étude, notre objectif était de montrer que les locuteurs ne s'appuient pas seulement sur les interjections en tant que marqueurs verbaux en situation d'interaction. L'étude détaillée de ces deux interjections nous a permis d'identifier une large variété de stratégies multimodales mises en œuvre en concert avec l'interjection pour l'expression des valeurs de polémique, de surprise et de connivence. Pour étudier les interjections *genre* et *as if* dans une approche multimodale et contrastive, nous avons

enregistré un corpus original de discussions semi-guidées entre étudiants de même langue. Nous avons ensuite transcrit et annoté le corpus avec le logiciel ELAN. Rappelons que la modalité vocale a été annotée avec ce même logiciel, en utilisant la convention de transcription proposée par Psathas (1995).

L'analyse des extraits vidéos nous a permis de conclure que les ressources verbales, les variations de la voix et les ressources posturo-mimo-gestuelles se synchronisent lorsque les locuteurs utilisent ces deux interjections. Plus précisément, la synchronisation de ces trois modalités d'expression favorise la co-construction du sens en interaction et contribue donc à véhiculer les valeurs soit de polémique, soit de surprise, ou encore de connivence de l'interjection *genre* et les valeurs soit de polémique, soit de surprise de l'interjection *as if*. Dans notre corpus, les expressions faciales, le regard, la prosodie et l'intonation sont les ressources centrales utilisées par les locuteurs. L'utilisation des gestes manuels reste faible. Tout au long de cette étude, nous avons également souligné que ces deux interjections sont proches à différents niveaux (syntaxe, sémantique, prosodie, expressions faciales, postures, etc.), bien qu'elles aient connu des scénarios différents.

L'apport de cette recherche a été de proposer une analyse intégrée de divers canaux d'expression pour rendre compte de la richesse, mais aussi de la complexité de ces deux phénomènes interjectifs récents, en essayant d'inclure un maximum de facettes. Par ailleurs, l'une des particularités fortes de notre travail a été de nous intéresser à deux phénomènes interjectifs, pourtant si proches, qui, avant cette étude, n'avaient jamais été étudiés en regard l'un de l'autre.

7 Perspectives pour des recherches futures

Dans le cadre de notre thèse de doctorat, nous continuons à étudier les deux interjections émergentes *genre* et *as if* dans une approche multimodale et contrastive. Nous souhaitons également créer un corpus vidéo d'interactions orales dyadiques de plus grande envergure, tout en y incluant des locuteurs natifs d'anglais américain. Nous nous baserons sur des analyses à la fois qualitatives et quantitatives de certains aspects formels. Il s'agira d'allier le qualitatif au quantitatif de manière à aboutir à des résultats recevables et les plus objectifs possibles. Nous calculerons, entre autres, la fréquence d'utilisation de ces deux interjections par nombre d'énoncés et par nombre de mots.

Nous observerons également combien de *genre* et de *as if* les locuteurs produisent en

tout, quelle que soit la catégorie grammaticale. Nous élaborerons ensuite un codage adapté aux objectifs et hypothèses de notre recherche. Les extraits seront également analysés de manière beaucoup plus différenciée par rapport aux réactions des locuteurs et au positionnement de chacun. Une telle entreprise permettra à la fois d'éclairer l'éventuelle complexification des ressources et des procédés utilisés par les locuteurs, mais également d'établir un parallèle entre ces deux interjections. Les outils utilisés seront les logiciels ELAN (transcription et codage des vidéos) et PRAAT (analyse prosodique).

Références

- ANDERSEN G. (à paraître). « Relevance ». In : C. RÜHLEMANN et K. ALJMER (eds), *Corpus Pragmatics*. Cambridge : Cambridge University Press, p.20-21.
- CHUQUET H., R. NITA et F. VALETOPOULOS (eds) (à paraître). *Des sentiments au point de vue : Perspectives contrastives*. Rennes : Presses Universitaires de Rennes.
- DUFAYE L. (2013). « Genre ou le scénario d'une grammaticalisation ». In : *Lynx*. Université Paris Ouest Nanterre.
- DUFAYE L. (2014). « Genre : from Category Identification to Intersubjective Differentiation ». In : *Symposium international Stance and (Inter) subjectivity* (Université Sorbonne Nouvelle – Paris 3 ; 22 mars 2014).
- DUNN M. (2005). *Zounds : a Brower's Dictionary of Interjections*. New York : St. Martin's Griffin.
- FLEISCHMAN S. (1999). « Pragmatic Markers in Comparative Perspective : A contribution to cross-language pragmatics », *PRAGMA 99* (Tel-Aviv ; 13-16 juin 1999).
- GADET F. et R. LUDWIG (2014). *Le français au contact d'autres langues*. Paris : Ophrys.
- GUMPERZ J., J. JOSEPH et S. LEVINSON (eds) (1996). *Rethinking. Linguistic Relativity. Studies in the social and cultural foundations of language*. New-York/Cambridge : Cambridge University Press.
- JAUDHUIN G. (2013). *A Multimodal and Multilevel Analysis of the Interjection as if in Spoken and Written Discourse*. Mémoire de Master 1 Recherche sous la direction de Madame le Professeur Aliyah Morgenstern, Université Sorbonne Nouvelle – Paris 3.
- JAUDHUIN G. (2014a). *The Paris and Canterbury Student Corpus. Corpus d'interactions*

dyadiques entre étudiants de même langue. Université Sorbonne Nouvelle – Paris 3.

JAUDHUIN G. (2014b). *Analyse multimodale et contrastive des interjections émergentes genre en français et as if en anglais : étude d'un corpus oral vidéo de discussions semi-guidées en français et en anglais britannique*. Mémoire de Master 2 Recherche sous la direction de Madame le Professeur Aliyah Morgenstern, Université Sorbonne Nouvelle – Paris 3.

LIDDELL S. (2003). *Grammar, Gesture, and Meaning in American Sign Language*, Cambridge : Cambridge University Press.

OCHS E. (1996). « Linguistic Resources for Socializing Humanity ». In J.J. GUMPERZ et S.L. LEVINSON (eds) (1996), *Rethinking Linguistic Relativity*. Cambridge : Cambridge University Press, p.407-437.

PINSON M. (2009). *Norme et variation en anglais contemporain : étude variationniste de quelques subordonnants complémentation de help et like conjonction*. Thèse de doctorat en Linguistique anglaise. Université Sorbonne Nouvelle – Paris 3.

PSATHAS G. (1995). *Conversation Analysis : The Study of Talk-In-Interaction*. Sage : Thousand Oaks, CA.

QUIRK R. (1985). *A Comprehensive Grammar of the English Language*. London : Longman.

ROSIER L. (2002). « Genre : le nuancier de sa grammaticalisation ». In : *Travaux de linguistique*, n°44. Bruxelles : De Broeck Université.

RUHLEMANN C. et K. ALJMER (eds) (à paraître). *Corpus Pragmatics*. Cambridge : Cambridge University Press.

SWIATKOWSKA, M. (2000). *Entre dire et faire. De l'interjection*. Krakow : Wydawnictwo UJ.

VIGNERON, J. (2013). « Genre en français, like en anglais et so en allemand : mise en scène et mise à distance dans le discours direct ». In H. CHUQUET, R. NITA et VALETOPOULOS F. (eds) (2013), *Des sentiments au point de vue : Perspectives contrastives*. Rennes : Presses Universitaires de Rennes, p.195-210.

VLADIMIRSKA, E. (2013). « Entre le dire et le monde : le cas du marqueur discursif genre ». *Colloque international Approximation et Précision* (Tel-Aviv ; 11-12 juin 2013).

Les connaissances lexicales : universalité ou variété de la catégorisation sémantique ? Étude du lexique de locuteurs de l'île de La Réunion

Audrey NOËL

LCF, 15, avenue René Cassin, BP 7151, 97 715 Saint-Denis Messag Cedex 9

naudrey86@aol.com

RÉSUMÉ

Cet article traite du lien entre organisation lexicale et organisation sémantique. La catégorisation sémantico-lexicale étant un processus universel et nécessaire lors de l'acquisition, on peut supposer que le système sémantique d'individus de langues différentes présente une organisation similaire. Retrouve-t-on également des similarités au niveau lexical ? Pour répondre à cette question, nous avons proposé une épreuve de fluence verbale concernant 21 catégories sémantiques, à 142 locuteurs d'âges variés bilingues français – créole réunionnais. Les données obtenues ont été comparées à celles issues d'études portant sur des sujets métropolitains et allemands (Dubois et Poitou, 2002). Les résultats montrent que certaines catégories sémantiques semblent être universelles, tandis que d'autres sont soumises, tout au long de la vie, à plusieurs facteurs de variabilité, que nous avons tenté de dégager. Cette recherche apporte un nouvel éclairage sur l'interaction entre langage et cognition, et permet d'approfondir les connaissances sur le fonctionnement langagier des locuteurs réunionnais.

ABSTRACT

Lexical knowledge : universality or variety of semantic categorization ? Study of lexicon of the Reunion Island's speakers

This article deals with the link between lexical organization and semantic organization. Semantic-lexical categorization is a universal and necessary process in acquisition ; so it can be assumed that the semantic system of individuals speaking different languages has a similar organization. Can we also find some similarities in the field of lexicon ? To answer this question, we have proposed a test of verbal fluency regarding 21 semantic categories, to 142 bilingual French – Reunion creole speakers of various ages. The data obtained were compared with those from studies on French and German subjects (Dubois and Poitou, 2002). The results show that some semantic categories seem to be universal, while others are subject throughout life to several factors of variability, that we tried to analyze. This research lights up on the interaction between language and cognition, and enables to increase our knowledge about the language functioning of Reunionese speakers.

MOTS-CLÉS : lexique, acquisition, catégorisation sémantique, La Réunion, contact de langues

KEYWORDS: lexicon, acquisition, semantic categorization, Reunion island, language contact

1 Introduction

L'acquisition du lexique est en lien étroit avec certaines opérations cognitives (David, 2000), ce qui amène à s'interroger sur le lien entre l'organisation lexicale et l'organisation du système sémantique. Si l'on part du présupposé que le système sémantique présente une organisation conceptuelle à caractère universel, serait-ce également le cas pour le système lexical ? Retrouverait-on alors, chez des locuteurs de langues différentes, les mêmes listes lexicales ? Une étude de Dubois et Poitou (2002) a montré de grandes similitudes entre le lexique de locuteurs français et de locuteurs allemands, ce qui appuie l'hypothèse d'une organisation sémantico-lexicale universelle. Néanmoins, ces auteurs retrouvent également des différences pour certains champs lexicaux, ce qui les amène à considérer que les listes lexicales sont également soumises à des facteurs de variabilité.

Nous nous sommes intéressée à l'organisation lexicale de locuteurs réunionnais (Lebon-Eyquem et Noël, 2014), partant du constat qu'il existe entre français et créole réunionnais, deux langues génétiquement et structurellement proches, des différences lexicales non négligeables. Nous nous sommes alors posée la question suivante : l'organisation sémantico-lexicale des locuteurs réunionnais diffère-t-elle de celle des locuteurs métropolitains ? Dans ce cas, quels facteurs peuvent expliquer ces différences ?

2 Notions théoriques

2.1 L'acquisition du lexique

On situe généralement l'apparition des premiers mots vers 11-12 mois (Kail et Fayol, 2000). S'ensuit une période qualifiée d'« explosion lexicale » : le vocabulaire augmente rapidement, l'enfant acquiert plusieurs mots par jour. Simultanément, il apprend à organiser son lexique (Rondal, 2003 ; Lemaire et Blaye, 2007). Pour Gopnik et Meltzoff (1987), l'explosion du vocabulaire est en relation directe avec la capacité de catégorisation : l'enfant développe son lexique parce qu'il se crée des représentations mentales et conceptuelles, et il élabore ses connaissances autour de *prototypes lexicaux*, items présentant les caractéristiques définitionnels de la catégorie (Rondal, 2003).

2.2 Le lexique des locuteurs réunionnais

La Réunion est un département français où se côtoient deux langues structurellement et génétiquement proches, le créole, idiome majoritairement usité par les locuteurs natifs, et le français, langue officielle. Le créole et le français ont la particularité de posséder 80 % de lexique commun (Chaudenson, 2003 ; Adelin et Lebon-Eyquem, 2010). Il est cependant encore fréquemment pointé des différences lexicales conséquentes : l'existence de « faux-amis », qui peuvent générer des incompréhensions dans la communication (« amuser » signifie en créole « traîner, flemmarder »), d'un lexique insulaire (« zanbrokal », plat typique de la cuisine créole), ou encore des phénomènes de sur-extension (« linge » peut être utilisé pour désigner tout type de vêtement) et de sous-extension (« boutique » désigne l'épicerie de quartier). Le lexique des locuteurs réunionnais s'avère-t-il alors significativement différent de celui des locuteurs métropolitains, en dépit de la part importante de lexique commun ?

3 Méthodologie

3.1 Protocole utilisé

Pour investiguer l'organisation lexicale de notre population, nous nous sommes inspirée de l'étude de Dubois (1983, 2002). Cette dernière a proposé des tâches de fluence verbale, déclenchées par un terme inducteur (nom de la catégorie sémantique). Nous avons retenu 21 catégories sémantiques, en reprenant certaines catégories adoptées par Dubois, et d'autres préconisées par Boisseau (2005). Nous avons réparti les différentes catégories en deux protocoles distincts (tableau 1) afin de diminuer la longueur de la tâche :

Protocole A	Animaux, fleurs, fruits, insectes, jouets, légumes, meubles, moyens de transport, outils, vêtements, aliments
Protocole B	Arbres, armes, boissons, couleurs, formes, instruments de musique, parties du corps, poissons, récipients, sports, aliments

TABLE 3 – Protocoles de passation

Nous avons opté pour une passation orale : les sujets interrogés devaient citer, à partir du nom de la catégorie, le maximum de mots possibles en une minute.

3.2 Recueil des données

La collecte des données s'est déroulée entre janvier et juillet 2013. Les passations ont eu

lieu dans des contextes variés (au domicile des sujets, en milieu scolaire...). Chacune a duré approximativement 15 minutes (une minute par catégorie plus un temps d'explication), et a fait l'objet d'un enregistrement puis d'une retranscription.

Nous avons précisé ne pas nous intéresser à une langue précise, afin d'éviter que les enquêtés ne fournissent uniquement des items lexicaux français ou créoles. Nous avons cherché à varier la langue de passation (français ou créole) afin de pouvoir par la suite analyser l'influence de la langue usitée par l'examineur sur les réponses des sujets.

3.3 Caractéristiques de la population

Les sujets de l'étude sont tous des locuteurs créolophones (ayant *a minima* des capacités de compréhension du créole), principalement issus du nord et de l'ouest de l'île. Notre échantillon comprend 142 témoins, âgés de 4 à 81 ans. Pour faciliter les analyses, nous avons réparti notre population en six tranches d'âge : les 0-4 ans (8 sujets), les 5-9 ans (69 sujets), les 10-19 ans (26 sujets), les 20-39 ans (18 sujets), les 40-59 ans (18 sujets) et les 60 ans + (7 sujets).

3.4 Analyse des données

Pour nos analyses, nous avons préalablement supprimé les redondances et les intrus sémantiques, puis nous avons mené des analyses d'occurrences de citation, permettant d'appréhender le phénomène de *fréquence lexicale*. Nous avons également calculé le nombre moyen de mots produits par catégories, et nous avons effectué des comparaisons inter-groupes (tests de Mann-Whitney et de Kruskal-Wallis).

4 Résultats

4.1 La fréquence lexicale

Il apparaît de prime abord, pour toutes les classes d'âge, que les items les plus fréquemment cités ne sont pas forcément des items du lexique créole : par exemple, pour la catégorie « fruits », « pomme » arrive en première position, avec 57 occurrences. Si l'on compare à titre d'exemple les premiers items de certaines catégories sémantiques, on constate de grandes similitudes entre les réponses données par les locuteurs réunionnais et celles données par les locuteurs métropolitains :

Catégories sémantiques	Locuteurs métropolitains	Locuteurs réunionnais
Animaux	Chien, chat, lion	Chien, chat, lion
Fleurs	Rose, tulipe, marguerite	Rose, marguerite, tulipe
Poissons	Sole, truite, raie	Requin, baleine, poisson rouge

TABLE 4 – Comparaison des premiers items lexicaux de trois catégories

Des différences sont néanmoins constatées : les premiers items n'apparaissent pas forcément dans un ordre similaire (pour « fleurs » par exemple), et pour certaines catégories (ici, « poissons »), les premiers items cités sont radicalement différents.

4.2 Possibles facteurs de variabilité

L'âge

On constate une augmentation de la moyenne de mots donnés en fonction de l'âge ; les 0-4 ans et les 5-9 ans se distinguent par ailleurs de façon significative des autres tranches d'âge pour une très grande majorité des catégories sémantiques (test de Kruskal-Wallis). Néanmoins certaines catégories sémantiques sont très tôt bien connues des jeunes enfants (exemple : « fruits »), tandis que d'autres sont moins élaborées (« fleurs »). On constate également, en analysant les occurrences de citation, que la structuration de certaines catégories sémantiques reste globalement stable au fil du temps (exemple : « animaux »), tandis qu'elle varie plus fortement pour d'autres catégories (« vêtements »). L'âge semble donc être un premier facteur de variabilité, qui n'affecte pas les catégories de la même manière.

L'input linguistique

Ayant interrogé des familles entières, nous avons constaté que certaines particularités lexicales sont partagées par les membres d'une même famille. Par exemple, tous les membres de la famille A. (les parents et leurs deux filles) produisent des transformations phonologiques par agglutination (exemple : « zoranj (orange) »), que l'on atteste chez certains locuteurs créolophones. Le terme « poupette » (« poupée ») est cité par deux d'entre eux pour la catégorie « jouets », et n'est pas attesté dans le reste du corpus. Ceci appuie l'hypothèse de l'influence de l'input linguistique sur le développement lexical.

Le milieu socio-économique

Les analyses en fonction de la variable socio-économique (uniquement menées pour la

tranche d'âge 5-9 ans) mettent en évidence que les enfants issus d'un milieu plus favorisé ont une richesse lexicale plus importante : ils citent plus de mots pour chaque catégorie sémantique. La différence n'est cependant significative (test de Mann-Whitney) que pour cinq catégories : la différence inter-catégorielle est donc importante. L'influence des milieux socio-économique et socio-culturel sur la construction lexicale est donc à relativiser.

La langue et les référents culturels

Les items les plus fréquemment cités par les locuteurs réunionnais n'ont pas forcément été des items créoles. La langue utilisée par l'examinateur n'a eu qu'un faible impact sur les réponses des sujets : ceux pour qui la passation s'est principalement déroulée en créole n'ont pas significativement donné plus d'items créoles. Par contre, le milieu culturel semble avoir une influence plus importante sur la structuration du lexique : dans la catégorie « arbres » par exemple, « bananier » et « palmier » sont très fréquemment cités. On atteste là encore des différences inter-catégorielles : l'influence culturelle a plus d'impact, par exemple, sur la catégorie « poissons » que sur la catégorie « animaux ».

5 Conclusion

Comment s'organise le lexique des Réunionnais ? Il apparaît déjà que dès le plus jeune âge, le lexique des locuteurs réunionnais témoins, tout comme celui des locuteurs de la France hexagonale, se structure en catégories sémantiques similaires. Toutefois, toutes les catégories ne sont pas acquises et développées au même moment ; la *variabilité inter-catégorielle* s'avère donc importante. Certaines catégories s'élaborent très tôt tandis que d'autres se structurent plus tardivement, et nous pouvons penser que ce phénomène s'attesterait également chez les locuteurs métropolitains.

Notre étude a montré qu'il existe de grandes similitudes entre les listes lexicales des locuteurs réunionnais et des locuteurs métropolitains ; toutefois des différences sont également constatées, ce qui amène donc à émettre l'hypothèse que certaines catégories sémantiques ont un caractère universel (apparaissent tôt, restent stables dans le temps, sont peu sensibles à la variation). D'autres catégories semblent quant à elles plus mouvantes et sont sensibles à différents facteurs de variation (âge, input linguistique, référents culturels, langue et milieu socio-économique *a minima*).

Cette recherche sur l'organisation sémantico-lexicale nous permet de mieux comprendre

comment se structure le lexique, et montre comment les faits universaux du langage s'articulent à la singularité des situations linguistiques.

Références

ADELIN E. et M. LEBON-EYQUEM (2010). *Adaptation de la didactique du français aux situations de créolophonie. Guide du maître : La Réunion. Niveau : Cycle 2 (CP et CE1)*. Paris : OIF / LeWebPédagogique.

BOISSEAU P. (2005). *Enseigner la langue orale en maternelle*. Paris : Retz.

CHAUDENSON R. (1974). *Le lexique du parler créole de La Réunion*. Paris : Honoré Champion.

CHAUDENSON R. (2003). *La créolisation : théorie, applications, implications*. Paris : L'Harmattan.

DAVID J. (2000). « Le lexique et son acquisition : aspects cognitifs et linguistiques ». In : *Le français aujourd'hui*, n°131, p.31-40.

DUBOIS D. (1983). « Analyse de 22 catégories sémantiques du français : organisation catégorielle, lexique et représentation ». In : *L'année psychologique*, n°2, vol.83, p.465-489.

DUBOIS D. et J. POITOU (2002). Les « normes catégorielles » : catégories sémantiques et/ou listes de mots. Structuration des catégories sémantiques : «entre» le cognitif et le linguistique ». In : D. DUBOIS (ed.) (2002), *Cahier du LCPE*, n°5 : « Normes catégorielles » et listes lexicales. Paris, p.7-29.

GOPNIK A. et A.N. MELTZOFF (1987). « The development of categorization in the second year and its relations to other cognitive and linguistic developments ». In : *Child Development*, n°58, p.1523-1531.

KAIL M. et M. FAYOL (2000). *L'acquisition du langage, tome 1 : Le langage en émergence, de la naissance à trois ans*. Paris : PUF.

LEBON-EYQUEM M. et A. NOËL (2014). « La structuration lexicale chez les locuteurs réunionnais : spécificités et normes catégorielles ». In : A. CARPOORAN (ed.) (2014), *Langues créoles, mondialisation et éducation. Actes du XIIIe colloque du Comité International des Études Créoles* (Maurice ; 5-9 décembre 2012). Maurice : CSU-ELP, p.361-387.

LEMAIRE P. et A. BLAYE (2007). *Psychologie du développement cognitif de l'enfant*. Bruxelles : De Boeck.

RONDAL J.A. (2003). *L'évaluation du langage*. Liège : Pierre Mardaga.

La notion de multiethnolecte dans le contexte scandinave

Sarah HARCHAOUI

Laboratoire du CeLiSo, Université Paris-Sorbonne (ED V, Concepts et Langages)

sarah.neslii@gmail.com

RÉSUMÉ

Il s'agira de discuter des différentes perspectives qu'implique la notion de multiethnolecte en premier lieu dans le contexte scandinave, la Suède ayant été un des terrains pionniers de la recherche, et la Norvège abritant la capitale où notre corpus de thèse a été enregistré. En nous référant à la définition exprimée par Clyne dans laquelle il distingue l'ethnolecte du multiethnolecte, nous voudrions montrer de quelle manière l'élargissement induit par le préfixe *multi-* constitue une avancée dans l'appréhension de la réalité des pratiques langagières orales relevées chez les adolescents scandinaves ces dix dernières années, tout en démontrant que le terme ne peut cependant pas être restreint à son acception première. Dans un second temps, nous voudrions revenir sur l'ambivalence du terme dans une perspective linguistique globale. Par définition, le lecte et ses dérivés vont à l'encontre de l'universel, pourtant plusieurs études démontrent que certaines variables prosodiques, syntaxiques, lexicales considérées comme constitutives des multiethnolectes sont communes aux grands espaces urbains européens. Un fil conducteur semble donc exister derrière ces pratiques langagières et cela malgré la diversité des profils des locuteurs. Un dénominateur commun fédérateur, tel qu'incarné dans l'urbain ou l'adolescence pourrait-il supplanter le facteur (multi)ethnique ?

ABSTRACT

The concept of multiethnolect in the Scandinavian context

This paper discusses the different perspectives raised by the notion of *multiethnolect*, first in the Scandinavian context because Sweden has hosted one of the pioneering fieldwork, but also because the corpus mainly used in my dissertation has been recorded in Oslo, Norway. According to Clyne's definition of multiethnolect, this paper points out how the expansion induced by the prefix *multi-* is an important step to understand the oral practices of adolescents in the last 10 years. But the term *multiethnolect* can not be restricted to its first meaning. Its ambivalence has to be raised in a global linguistics perspective. *Lect* and derivatives go, by definition, against universal values. Yet, recent studies in Europe state that some prosodic, syntactic and lexical variables are common to multiethnolects in urban areas. This implies that ethnic factors do not longer constitute a primary value, and moreover that ethnicity has been supplanted by a broader factor such as urbanity.

MOTS-CLEFS : multiethnolecte, style, diversité, adolescents, milieu urbain, Scandinavie

KEYWORDS: multiethnolect, style, diversity, youngster, urban area, Scandinavia

1 Apparition et évolution du terme en Scandinavie

1.1 L'apport de Kotsinas en Suède

En Suède, la fin des années 1980 est marquée par un renouveau dans l'analyse linguistique des pratiques langagières relevées au sein de groupes d'adolescents vivant en milieu urbain. Dans son article de 1988, Kotsinas soulève pour la première fois la question de l'existence et de l'autonomie du *rinkebysvenska* (litt. « le suédois de Rinkeby »), une variété développée lors d'échanges entre pairs de locuteurs adolescents évoluant à Rinkeby, banlieue de Stockholm connue pour sa diversité sociolinguistique et ethnique. Bien qu'elle se compose de variables phonétiques, prosodiques, syntaxiques et lexicales qui étaient jusque-là jugées comme écarts ou lacunes par rapport à la variété standard nationale, Kotsinas démontre que cette variété ne résulte ni d'un apprentissage imparfait du suédois comme seconde langue, ni d'une situation d'interlangue. Elle écarte ainsi toute analyse des variables entrant dans le cadre de l'acquisition d'une seconde langue (SLA), ce qui le cas échéant sous-entend que les jeunes locuteurs ne sont pas locuteurs natifs de par le fait que l'un de et/ou leurs deux parents sont issus de l'immigration. Le *rinkebysvenska* constitue au contraire une variété indépendante se développant en parallèle d'autres variétés stockholmoises, c'est pourquoi son étude doit légitimement s'inscrire dans les domaines de la sociolinguistique et de la dialectologie scandinaves (Kotsinas, 1987, 1988). Pour justifier cela, les deux principaux arguments mis en avant sont d'une part l'emploi de la variété par des locuteurs aux profils très divers mais qui ont tous grandi dans un milieu suédophone (donc considérés comme locuteurs natifs) et d'autre part son usage qui ne relève pas d'une contrainte linguistique mais bien d'une revendication identitaire consciente. C'est notamment sur ce point que nous reviendrons à la fin de notre article.

Cette perspective structuraliste permettant la description formelle des variables relevées dans le discours des jeunes locuteurs par rapport à une variété standard, a dès lors soulevé intérêt et controverse au sein des milieux linguistiques scandinaves et internationaux. Elle reflète également la volonté des linguistes de trouver comment étiqueter ces pratiques langagières en tant que variété (Bodén, 2004 ; Fraurud et Bijvoet, 2004). Le phénomène donne l'impulsion à une série de projets nationaux, dont le premier « Language and language use among young people in multilingual urban settings » est mené en 2002 à Stockholm, Gothenburg et Malmö. À cette occasion, le

terme de SMG (Swedish on multilingual ground) est proposé afin de se référer à l'objet d'étude analysé en évitant les questions de catégorisation du phénomène. Au terme du projet, Bodén (2007) réaffirme le caractère autonome et non lacunaire de la variété observée une décennie plus tôt par Kotsinas. Elle rappelle également que les variables observées se retrouvent aussi bien chez des locuteurs adolescents dont les parents sont issus de l'immigration que ceux dont les deux parents sont locuteurs natifs du suédois. Cette dernière information est particulièrement pertinente puisqu'elle témoigne de la manifestation des linguistes à faire état des pratiques sans pour autant établir une corrélation entre le recours à une des variables linguistiques et l'origine du locuteur qui l'emploie, dès lors que le locuteur est natif de la langue nationale. En effet, comme le rappellent Gadet et Hambye (2014 : 190) :

« shared ethno-cultural roots are neither a necessary nor a sufficient condition for sharing a culture, and it is impossible to affirm that the linguistic practices of social groups with foreign ethno-cultural roots are due to this common ethno-cultural background »¹³.

En d'autres termes, même si un groupe ethnique est lié par le partage de codes culturels, cela n'implique pas une relation de cause à effet entre le facteur ethnique et la culture, et encore moins entre l'ethnicité et les pratiques linguistiques développées.

1.2 Le multiethnolecte danois

À la même époque au Danemark, Quist (2000) mène de son côté des recherches dans la banlieue de Copenhague, à Nørrebro et Vesterbro où elle découvre un phénomène équivalent qu'elle nomme le *ny københavnsk multiethnolect* (litt. « le nouveau multiethnolecte copenhagois »). Avec une perspective variationniste, elle met en évidence une certaine systématisme à différents niveaux linguistiques entre le suédois et le danois. Elle choisit cependant le terme de multiethnolecte formulé par Clyne la même année (2000 : 87) car la variété observée n'est pas liée à un groupe ethnique particulier mais résulte « d'une situation multiethnique dans laquelle les locuteurs ont des origines (y compris danoise) et des langues premières différentes »¹⁴ (Quist, 2008 :

¹³ Traduction de l'auteure : « les origines ethnoculturelles communes ne sont ni une nécessité ni une condition suffisante pour partager une culture, et il est impossible d'affirmer que les pratiques langagières de groupes sociaux avec des origines ethnoculturelles étrangères sont dues à cet héritage ethnoculturel commun ».

¹⁴ Traduction de l'auteure.

49). De plus, elle justifie le suffixe *-lecte* en invoquant le fait qu'il s'agit d'un « *parallel phenomenon to other “lects” like sociolects, dialects, chronolects, and so forth, and it ought to be regarded as such, that is, as something mundane and not “exotic”* »¹⁵ (2008 : 49).

Deux ans plus tard, une autre danoise, Christensen (2002) fait une thèse de doctorat sur l'usage et les choix langagiers des adolescents en milieu multiethnique (traduction de l'auteure) à Århus (seconde ville du Danemark) au cours de laquelle elle découvre une variété similaire qu'elle appelle *århusiansk etnolekt* (litt. « l'ethnolecte d'Århus »). Christensen appuie son utilisation du terme d'ethnolecte sur le fait que la plupart des influences linguistiques relevées chez ses informateurs proviennent de l'arabe. Contrairement à Quist qui opte pour le préfixe *multi-*, Christensen s'en réfère donc à la définition du premier type d'ethnolectes formulée par Clyne (2000 : 86), créant une relation directe entre l'origine arabe des locuteurs et leur emploi de certaines variables. Paradoxalement, elle observe que les locuteurs ayant recours à cet ethnolecte utilisent de façon significative des variables dialectales typiques de la ville d'Århus, variables tombées en désuétude chez les locuteurs natifs originaires de cette ville. Cette dernière observation, sur laquelle nous reviendrons, est particulièrement importante car elle met en exergue un point commun essentiel à l'ensemble des variables linguistiques rencontrées en milieu urbain scandinave : elles sont porteuses d'un ancrage local fort et s'inscrivent au même titre que d'autres variétés dans l'ensemble des pratiques développées à l'échelle nationale, régionale et locale. En effet, il est non seulement indispensable d'inscrire l'analyse de ces variétés multiethniques dans un paysage linguistique local (qui dans les pays scandinaves se traduit avant tout par l'omniprésence des dialectes au quotidien¹⁶), mais surtout on ne saurait les comprendre en les isolant de l'espace local social dans lequel elles sont employées et développées.

1.3 Le discours multiethnolectal norvégien

La dimension sociale est encore plus marquante si l'on reprend les travaux menés récemment en Norvège dans le cadre du projet national UPUS (*Utviklingsprosesser i*

¹⁵ Traduction de l'auteure : « un phénomène parallèle à tous autres “lectes” comme les sociolectes, les dialectes, les chronolectes, etc., et qu'il devrait être regardé en tant que tel, c'est-à-dire comme quelque chose d'ordinaire et non “exotique” ».

¹⁶ À ce propos, mentionnons un proverbe populaire norvégien qui dit : *Man kan si at det finnes like mange dialekter som det finnes steder*, litt. « On peut dire qu'il existe autant de dialectes qu'il existe de lieux ».

urbane språkmiljø) – « Processus de développement linguistique en milieu urbain » lancé entre 2006 et 2008 par le laboratoire linguistique de l'Université d'Oslo afin d'observer et de recenser les pratiques langagières dans deux circonscriptions de la capitale. Les premiers résultats révèlent des pratiques langagières disparates entre les locuteurs des quartiers Ouest et Est dans leur emploi du lexique et notamment concernant l'origine des emprunts utilisés. Cette différence découle vraisemblablement de l'évolution socioéconomique de la capitale et de sa scission progressive d'Est en Ouest et du centre vers les périphéries. Dans les faits, les quartiers Ouest et le centre sont aujourd'hui reconnus pour être les plus huppés de la capitale, où l'immigration plus faible provient majoritairement des pays européens ainsi que d'Amérique du Nord, d'Australie et de Nouvelle-Zélande¹⁷. Opsahl, Røyneland et Svendsen (2008) ont démontré que les adolescents venant de ces quartiers empruntent des mots à l'anglais, l'espagnol ou l'allemand, tandis que dans les quartiers Est fortement marqués par l'immigration non-occidentale, les adolescents empruntent majoritairement du lexique à l'arabe, au berbère, à l'ourdou, au turc ou encore au panjabi, langues correspondant aux plus grands groupes de populations immigrées (Bureau des Statistiques Nationales – SSB, 2008). Ces emprunts extra-européens marquent d'une pierre blanche l'évolution de la linguistique norvégienne, puisque c'était la première fois¹⁸ que l'on recense du lexique extra-européen dans des pratiques orales entre pairs, lexique utilisé y compris par des adolescents locuteurs natifs du norvégien. C'est d'ailleurs autour de cette variété uniquement présente dans les quartiers multiculturels de l'Est que les recherches d'UPUS se concentrent, et que l'on qualifie de *multiethnolektisk norsk* (litt. « norvégien multiethnolectal ») ou *multiethnolectal speech* (litt. « discours multiethnolectal »).

Ce qui est intéressant, c'est l'attribution des causes de ces emprunts formulées par Drange (2002) et Hasund (2006), deux linguistes qui travaillent sur le lexique dans les parlars jeunes norvégiens au sens large. A l'Ouest d'Oslo, on serait en présence

¹⁷ Le Bureau des Statistiques Nationales norvégien, le SSB, prend le pays d'origine comme critère de recensement et de classement des populations ayant un passé à l'immigration sur le territoire norvégien. En effet, une distinction est faite entre a) les pays scandinaves, b) les pays d' UE et de l'EEE, c) les pays européens hors UE et EEE, d) l'Amérique du Nord, l'Australie et la Nouvelle-Zélande, et e) les pays d'Asie (Turquie comprise), d'Afrique, d'Amérique Latine et du Sud.

¹⁸ En 1995, Aasheim mène un travail pionnier à Oslo qui met en évidence le recours à des emprunts non-européens dans la capitale. Elle est également à l'origine de la dénomination *kebabnorsk* (litt. « Norvégien kebab ») actuellement reconnue par l'opinion publique (qui fut la première à l'utiliser) pour être une expression péjorative et stigmatisante.

d'emprunts de « contact culturel » c'est-à-dire qui découlent de la relation qu'exerce indirectement une langue sur l'autre par le biais de son rayonnement culturel et de son influence linguistique en termes de littérature, de musique, d'art ou de mode. À l'inverse, les emprunts des quartiers Est seraient liés à un « contact direct », qui apparaît dans une situation de contact prolongé entre plusieurs langues. Selon nous, la première explication est pertinente puisque l'on retrouve la présence d'un lexique anglais dans les pratiques langagières des jeunes norvégiens en général, en dehors d'Oslo, ce qui rejoint l'idée d'une influence globale particulièrement remarquable dans les domaines des nouvelles technologies, de la musique ou de la mode.

Néanmoins dans le second cas, nous estimons que la cause invoquée n'est pas convaincante, car elle sous-entend que la situation des quartiers Est est le produit d'un « contact de langues », ce qui nous renvoie à la même problématique avec le *rinkebysvenska* en Suède que nous avons précédemment soulevée. Or le norvégien multiethnolectal apparaît dans un contexte sociogéographique relativement similaire à celui des variétés multiethniques suédoise et danoise puisqu'il est parlé dans la plus grande ville démographique de Norvège. L'attachement à la localité se retrouve d'abord sur le plan socio-économique, et reflète la bipartition historique de la ville, où l'Ouest traditionnellement bourgeois fait face à l'Est anciennement ouvrier, devenu aujourd'hui multiculturel. Cette scission géographique se retrouve également au niveau linguistique avec l'émergence de deux variétés orales souvent analysées en tant que sociolectes, et qui se rapportent respectivement aux quartiers Ouest et Est : le *vestkantdialekt* (litt. « le dialecte du côté Ouest ») et l'*østkantdialekt* (litt. « le dialecte du côté Est »). Mais les représentations métalinguistiques des locuteurs qui reposent sur des critères socio-économique, culturel et esthétique exercent une influence considérable sur la façon dont est jugée la langue parlée dans les différents quartiers de la capitale, l'Ouest bénéficiant d'un prestige relativement haut. À l'instar de leurs voisins scandinaves, les locuteurs du multiethnolecte norvégien ont des profils hétérogènes, qu'il s'agisse de la diversité de leurs origines, de leur parcours social, culturel et linguistique. Enfin, les variables prosodiques, syntaxiques et lexicales qui composent la variété multiethnolectale norvégienne sont elles aussi pointées comme divergentes de la norme. Notons d'ailleurs à ce sujet que Svendsen et Røyneland (2008 : 80) défendent l'existence de « trans-Scandinavian multiethnolectal patterns », ce qui renforce l'idée de pratiques expansionnistes et communes aux espaces urbains scandinaves.

Cependant, deux éléments mis en avant par le projet UPUS amènent à appréhender les variétés non plus par la seule description de leurs traits linguistiques (approche structuraliste) mais plutôt par la façon dont elles sont utilisées dans le discours (approche sociopragmatique ou fonctionnelle). Le premier est de constater que les locuteurs des quartiers Est sont capables de passer d'une variété à une autre en fonction de la situation d'énonciation dans laquelle ils se trouvent. En effet, la variété multiethnolectale est majoritairement utilisée lors de conversations entre pairs où les adolescents discutent en groupe, sans la présence d'un adulte (cf. le « paradoxe de l'observateur » : « observer la façon dont les gens se servent du langage quand on ne les observe pas », Labov, 1976 : 116). Certains adolescents sont ainsi capables d'éviter complètement les variables multiethnolectales dès lors qu'ils sont interviewés par un adulte. Cela alimente l'idée que nous défendons et selon laquelle les variétés sont en réalité un ensemble de pratiques stylistiques faisant partie du répertoire linguistique et sociétal du locuteur, ce dernier puisant dans toutes les ressources qui sont à sa disposition, aussi bien par son patrimoine linguistique (langues maternelles, apprises à l'école) que par ses choix personnels (langues qu'il juge prestigieuses). Peu importe que le style multiethnolectal se compose de variables jugées incorrectes par les locuteurs de la variété normée, il n'en demeure pas moins intégré dans le répertoire de l'adolescent et son utilisation le plus souvent contrôlée. Le deuxième élément présenté est que les adolescents sont conscients d'utiliser plusieurs types de variétés orales lors de leurs échanges entre pairs. Lorsqu'on les interroge sur leurs pratiques, ils reconnaissent avoir recours ou non à la variété multiethnolectale (sans la nommer ainsi – soit ils ne lui donnent aucun nom, soit ils lui attribuent le plus souvent le nom de la localité où ils vivent), et cela dans des domaines particuliers car ils savent que cela peut nuire à leur image dans un cadre conventionnel, notamment à l'école, ou lors d'un entretien professionnel.

Après avoir brièvement retracé l'apparition du terme pays par pays, nous avons donc pu dégager deux grandes approches pour appréhender le phénomène des multiethnolectes scandinaves. Nous avons commencé par l'approche structuraliste, démarche pionnière développée par Kotsinas et Quist (2000) mais qui présente néanmoins quelques limites. Bien qu'elle permette de décrire une structure linguistique, et de mettre en évidence une certaine systématité des variables, elle présuppose une vision figée de la variété, où l'on aurait une opposition binaire entre variété standard / variété multiethnique, et donc par extension entre majorité / minorité. Cependant,

Svendsen et Røyneland (2008 : 66) rappellent que cette approche est nécessaire et constitue une pré-condition à une analyse de ses fonctions dans le discours. Enfin, insistant sur l'hétérogénéité de l'espace social et sociolinguistique dans lequel les variables multiethnolectales apparaissent (qui présentent une certaine homogénéité due à la proximité des 3 langues scandinaves – suédois, norvégien, danois), nous avons peu à peu glissé vers l'approche fonctionnelle. La conscience qu'ont les locuteurs de leur usage de la langue selon diverses situations d'énonciation nécessite selon nous de s'intéresser à la manière dont ils utilisent le discours en tant que ressource pour négocier leur identité et se positionner dans l'espace social qui les entoure.

2 De l'approche structuraliste à la pratique stylistique

Dans un article publié en 2008, Quist revient sur ses recherches passées et envisage désormais le multiethnolecte comme une pratique stylistique, une direction qui est à rapprocher de ce que Svendsen et Røyneland nomment fonctionnelle ou sociopragmatique. Quist (2008 : 50-51) définit la pratique stylistique comme « the process through which signs and differences become meaningful resources in daily enterprises and activities »¹⁹, et ajoute que la pratique stylistique couvre « processes that connect different resources (linguistic and non linguistic ones) in meaningful relationships in association with the participants' identity negotiations »²⁰. En d'autres termes, une corrélation existe entre le recours à un ensemble de ressources dont le locuteur dispose et les valeurs sociale et sociétale sous-jacentes à ces ressources qui vont permettre au locuteur de négocier son identité lors d'un échange avec son interlocuteur. Cette perspective est proche de la notion de style d'Eckert (2001 : 123), définie comme « a clustering of linguistic resources, and an association of that clustering with social meaning »²¹. À la différence près que les ressources de Quist ne se limitent pas au champ linguistique, et incluent entre autre des ressources de type corporel et/ou esthétique.

Ce qui suscite notre intérêt dans la notion de pratique stylistique ou plus généralement

¹⁹ Traduction de l'auteure : « le processus à travers lequel signes et différences deviennent des ressources significatives dans les entreprises et les activités quotidiennes ».

²⁰ Traduction de l'auteure : « la pratique stylistique couvre les processus qui connectent différentes ressources (linguistiques et non-linguistiques) dans des relations significatives en association avec les négociations de l'identité des participants ».

²¹ Traduction de l'auteure : « un éventail de ressources linguistiques et l'association de cet éventail avec une signification sociale ».

de style, c'est l'idée du choix, conscient ou non, du locuteur d'avoir recours dans une situation d'énonciation donnée à une collection de ressources, aussi variées soient-elles. Ce qui nous intéresse également, c'est que ces choix sont régis à leur tour par des processus à travers lesquels le langage et l'identité se lient et modèlent la signification sociale (Quist, 2008 : 49). Cette quête du sens peut trouver échos dans ce que Le Page et Tabouret-Keller (1985) ont formulé dans un contexte différent de celui des métropoles scandinaves (puisque'il s'agissait d'une situation de contact créole à Belize) mais qui néanmoins explique un changement dans le choix et l'usage langagier des individus en termes de construction de l'identité sociale. L'acte d'identité permet en effet de « révéler à la fois la propre identité mais aussi la recherche d'un rôle social »²² (1985 : 14). Ainsi, identifier les mécanismes dans des situations où les locuteurs choisissent d'utiliser certains traits langagiers au détriment d'un autre est primordial pour saisir les constructions de l'identité sociale et catégoriser un individu comme membre d'un groupe. Svendsen et Røyneland (2008 : 66) précisent que dans ce cas, « le choix n'est pas seulement dénotatif mais motivé par sa valeur symbolique sur le marché linguistique », faisant ainsi référence aux travaux de Bourdieu (1991) et d'Auer (2007). Cela nous ramène donc à la négociation de l'identité lors des échanges entre pairs, et notamment à la représentation que l'on a de soi-même et à la façon dont cette représentation se construit à travers le discours.

Dans son article « 'Style' as distinctiveness : the culture and ideology of linguistic differentiation », Irvine (2008 : 21-23) considère fondamental le caractère distinctif du style. Elle réaffirme que la distinction opérée par un individu passe avant tout par son cadre social, ce qui repose sur une évaluation sociale voire esthétique. Mais ce qu'elle met en avant, c'est bien la dimension idéalisée des représentations lors de leur construction. La définition du style qu'elle formule qui prend en compte l'autoreprésentation de soi dans le discours et les valeurs que le locuteur juge centrales pour lui, ou pour le groupe auquel il veut être associé ou dissocié, pour reprendre les mots de Le Page et Tabouret-Keller, nous semble être particulièrement pertinente pour appréhender les multiethnolectes scandinaves.

En effet, si l'on s'intéresse de plus près aux locuteurs adolescents du projet UPUS, il paraît indiscutable que les variables identifiées comme multiethnolectales font partie de ces ressources qui deviennent socialement significatives. De plus, le fait qu'une partie

²² Litt. « a series of acts of identity in which people reveal both their personal identity and their search for social roles ».

d'entre eux soient natifs du norvégien (appartenant donc à la majorité) et qu'ils aient recours à ces variables, nous semble significatif d'un renouveau des valeurs indexées jusqu'alors par ce registre. Nous faisons là référence aux propos suivants de Clyne (2000 : 87) :

« In some cases, where members of the dominant (ethnic) group, especially young people, share it with the ethnic minorities in a 'language crossing' situation. It is the expression of a new kind of group identity»²³.

3 Vers un dépassement du critère ethnique

Dans l'optique d'une dichotomie majorité / minorité d'une part, et variété standard / variété multiethnolectale d'autre part, il apparaît que les adolescents des quartiers Est d'Oslo ont recours à l'argot et aux emprunts de langues dites « immigrées » (litt. *innvandringspråk*) en vue de plusieurs motivations. Les stratégies discursives lexicales relevant traditionnellement de l'argot (Calvet, 2007) passent d'abord par une dimension ludique pour crypter le discours, mais sont aussi utilisées à des fins de connivence afin de maintenir la cohésion des membres d'un groupe qui partagent les mêmes codes. Si les membres de ce groupe constituent une communauté linguistique telle que Labov la définit, « moins par un accord explicite quant à l'emploi des éléments du langage que par une participation conjointe à un ensemble de normes » (1976 : 187), c'est-à-dire comme « un groupe de locuteurs qui ont en commun un ensemble d'attitudes sociales envers la langue » (1976 :338), alors il est intéressant de se demander de quelle nature sont les solidarités d'attitudes des locuteurs des multiethnolectes scandinaves. Alors que Clyne faisait référence à une situation de *crossing* (cf Rampton, 1995), Svendsen et Røyneland (2008 : 80) constatent que

« in both of the Oslo corpuses there seems to be a link between the use of the multiethnolectal speech style and engagement within music scenes such as hip hop and rap »²⁴.

²³ Traduction de l'auteure : « Dans certains cas, où les membres d'un groupe (éthnique) dominant, notamment les jeunes, le (multiethnolecte) partagent avec les minorités ethniques dans une situation de "croisement linguistique". C'est l'expression d'une nouvelle forme d'identité de groupe ».

²⁴ Traduction de l'auteure: « dans les corpus d'Oslo, il semble y avoir un lien entre l'utilisation du style langagier multiethnolectal et un engagement à travers la musique telle que le hip hop et le rap ».

En effet, dans le cas d'Oslo, les clivages historiques à l'origine de la scission de la ville d'Est en Ouest et qui reposaient alors principalement sur une stratification sociale et ethnique sont en pleine évolution. En reprenant la notion de champ indexical (*indexical field*) d'Eckert (2008 : 454) qui pense que

« the meanings of variables are not precise or fixed but rather constitute a field of potential meanings – an indexical field, or constellation of ideologically related meanings, any one of which can be activated in the situated use of the variable »²⁵,

il est probable que les locuteurs d'UPUS aient indexé des valeurs originaires de la culture hip-hop américaine dans leur discours. Au-delà de ce que Potter (1995 : 68) considère comme l'expression « d'un profond engagement à pousser, tordre et briser les limites », définissant le hip-hop comme une culture de résistance dont la langue serait un « vernaculaire de résistance » (*resistance vernacular*) dans lequel la variation et l'improvisation sont utilisées afin de déformer et de repositionner les règles de l'intelligibilité mises en place par la variété dominante, deux valeurs semblent émerger particulièrement de la culture hip-hop et se refléter dans le discours des locuteurs de la variété multiethnolectale à Oslo à plusieurs niveaux.

1) L'urbanité : l'appartenance identitaire à un groupe ne semble plus être définie par le partage d'origines communes mais plutôt par le partage d'un territoire local commun, de façon générale, un territoire urbain, à travers lequel se manifeste un attachement à l'endroit où le locuteur a grandi et où il continue à évoluer. Cet ancrage local transparait par les noms attribués à la variété par les locuteurs-mêmes et qui font référence aux noms des lieux au sein desquels ils vivent. Remplacer le facteur ethnique par un facteur local est peut-être influencé par le rayonnement mondial de la culture hip hop qui fédère la jeunesse urbaine dans son ensemble, mettant de côté les différences ethniques et religieuses. Sur le plan morphologique, l'emploi de formes nominales non marquées²⁶, l'assouplissement de la position verbale dans les

²⁵ Traduction de l'auteure : « les significations des variables ne sont ni précises ni fixes mais constituent plutôt un champ de significations possibles – un champ indexical, ou constellation de significations idéologiquement liées, dont chacune peut être activée dans l'usage situé de la variable ».

²⁶ Dans les langues scandinaves, il existe 2 genres : commun et neutre (le norvégien compte également un genre supplémentaire, celui féminin). Une des variables relevées dans le discours des jeunes locuteurs des multiethnolectes scandinaves est l'utilisation par défaut du genre commun, c'est-à-dire non-marqué, aux dépens du genre neutre. Ainsi, on retrouve par exemple l'expression *det kroppssproget* (litt. « le langage corporel ») qui est au genre neutre en danois qui devient *den kroppsspråket*, l'article *den* se

propositions principales²⁷, de même que l'utilisation de marqueurs discursifs transscandinaves voire transeuropéens tels que *wallah*²⁸ pourraient matérialiser une position d'ouverture au monde des locuteurs adolescents. La variété multiethnolectale serait alors utilisée avec une dimension vernaculaire, c'est-à-dire comme une variété de langue expansionniste, qui reflète une volonté de ne pas être limité par des barrières linguistiques ou terrestres.

2) Toujours dans le cas des quartiers multiculturels à l'Est d'Oslo, les locuteurs se définissent eux-mêmes comme « honnêtes et authentiques » (Harchaoui, 2012) une valeur qu'ils considèrent primordiale pour leur identité, et qui entre en opposition avec le caractère « prétentieux, faux et superficiel » des adolescents des quartiers Ouest. **L'authenticité** serait donc la seconde valeur que les locuteurs matérialisent dans leur langue par l'emploi d'expressions ayant à l'origine un caractère religieux mais qui sont aujourd'hui désacralisées. Ainsi, des expressions comme *wollah* (litt. « je jure au nom d'Allah »), immédiatement suivies de *helt ærlig* (litt. « complètement sincère ») relevées dans le corpus UPUS témoignent bien de la volonté des locuteurs de refléter une authenticité à travers leur parole.

4 Conclusion

Le facteur ethnique à l'origine de la dénomination du concept de multiethnolecte semble bel et bien insuffisant pour comprendre l'ensemble des réalités que couvre le

rapportant au genre commun.

²⁷ Dans les langues scandinaves, l'ordre syntaxique des propositions principales comporte la contrainte du verbe en deuxième position selon l'ordre typologique XYV (X = sujet ; Y = verbe ; V = complément, soit SVO). Dans les variétés identifiées comme multiethnolectes scandinaves, la deuxième position obligatoire du verbe n'est pas appliquée, le verbe se trouvant parfois à la 3e position comme dans l'exemple suédois suivant :

Sen hon gick till skolan, au lieu de *Sen gick hon till skolan* (litt. « puis elle alla à l'école »).

CONJ PRON-3PS V-PST PREP ÉCOLE

²⁸ *Wollah* est la transcription phonétique norvégienne du terme *wa-llāh(i)* ; *Wa-* étant la particule qui porte le sens de jurer et *llāh(i)* le mot pour désigner Dieu (avec éventuellement le suffixe génitif *-i*) est en premier lieu une expression d'origine arabo-musulmane qui signifie « jurer au nom d'Allah », « Allah » désignant « Dieu » dans l'islam. Selon Madsen (2008 : 150) : « It is often used in conversations as an interjectional either as an oath, which involves the truth obligation of the speaker, or as an enhancement stating to which degree the content of the utterance is the case ». Traduction de l'auteur : « Il est souvent utilisé soit en tant qu'interjection ou tel un serment dans les conversations, ce qui implique l'obligation morale du locuteur de dire la vérité, ou soit en tant qu'intensificateur indiquant dans quelle mesure le contenu de l'énoncé est vrai ».

phénomène dans son ensemble. En prenant comme terrain de recherche les milieux urbains scandinaves et notamment le norvégien, nous avons pu mettre en évidence deux perspectives d'analyse. La première structuraliste a révélé des correspondances quant à la nature des variables linguistiques présentes dans les variétés des jeunes évoluant en milieu urbain multiethnique en Scandinavie. La seconde, fonctionnelle, a soulevé une réflexion d'ordre général sur les mécanismes qui sous-tendent l'indexicalisation de valeurs (morales, esthétiques, etc.) à des variables linguistiques en vue d'une négociation de l'identité dans et par le discours. En somme, la mise en scène consciente ou non de la parole et de la représentation du soi à l'échelle individuelle mais aussi collective. De plus, le concept de champ indexical proposé par Eckert offre dans le cas des pratiques multiethnolectales un renouvellement quasi illimité de la signification sociale que chacun attribue aux variables dans son discours.

Cependant, une question demeure : le terme multiethnolecte peut-il perdurer pour décrire les phénomènes linguistiques en présence ? Car même si le préfixe *multi-*pourrait à première vue combler la lacune sémantique de *ethnolecte*, il n'en demeure pas moins contre-productif (car élargi mais demeurant restrictif) et incapable de refléter la diversité des espaces sociaux et énonciatifs desquels il émerge.

Un critère hors ethnicité pourrait alors fédérer ces pratiques : celui du caractère adolescent. Cependant, en 2011, Rampton met en évidence dans son article le maintien des pratiques à l'âge adulte en ce qui concerne son terrain de recherche en Angleterre. Il suggère alors une nouvelle dénomination : le vernaculaire urbain contemporain (*Contemporary Urban Vernacular*). Selon lui, la notion de vernaculaire est plus appropriée que celle de *lecte*, car elle n'implique pas un rapport à une norme établie, de même qu'elle ne réduit pas l'usage à une tranche d'âge (idée de durabilité à l'âge adulte). Le facteur ethnique ou multiethnique tel qu'il est envisagé aujourd'hui peut quant à lui être remplacé par la notion d'urbain qui présuppose une diversité sans cantonner aux problèmes déjà soulevés.

En ce qui concerne le terrain scandinave, nous ne disposons pour l'instant pas d'assez de recul ni même de moyens pour savoir si les variables multiethnolectales ont perduré chez les informateurs à l'âge adulte. Bien que la notion de *vernaculaire* soit discutable à son tour (car faisant elle-même écho à une tradition linguistique bien définie), cette nouvelle dénomination n'en reste pas moins ouverte sur son époque et sur le monde, laissant entrevoir une portée plus universelle du phénomène.

Références

- AUER P. (2007). « Social identities and styles. Alternative approaches to linguistic heterogeneity ». In : P. AUER (ed.) (2007), *Social identity and communicative style : Alternative approaches to variability in language*. Berlin : Mouton de Gruyter, p.1-21.
- AASHEIM S.C. (1995) “Kebab-norsk” : framandspråkleg påverknad på ungdomsspråket i Oslo. Hovedoppgave. Université d’Oslo.
- BODÉN P. (2004). « A new variety of Swedish ? » In : S. CASSIDY, F. COX, R. MANNELL et S. PALETHORPE (eds) (2004), *Proceedings of the Tenth Australian International Conference on Speech Science and Technology* (Macquarie University, Sydney ; 8th-10th December, 2004). Sydney : Australian Speech Science and Technology Association (Inc.), p.475-480.
- BODÉN P. (2007). « “Rosengårdssvensk” fonetik och fonologi ». In : L. EKBERG (ed.) (2007), *Språket hos ungdomar i en flerspråkig miljö i Malmö*. Nordlund 27. Småskrifter från Nordiska språk. Lund: Nordiska språk, Språk- och litteraturcentrum, p.1-47.
- BOURDIEU P. (1991). *Language and symbolic power*. Cambridge, MA : Polity Press.
- CALVET L.J. (2007). *L’argot*. (3e éd.) Paris : Presses Universitaires de France (« Que sais-je ? »).
- CHRISTENSEN M.V. (2004). « Arabiske ord i dansk hos unge i multietniske områder i Århus ». In : S. DABELSTEEN et J. ARNFAST (eds) (2004), *Taler de dansk ? Aktuel forskning i dansk som andetsprog*. Københavnerstudier i tosprogethed, bind 37. Københavns Universitet, p.33-49.
- CHRISTENSEN M.V. (2010). *8220, 8210 – Sproglig variation blandt unge i multietniske områder i Århus*. Århus Universitet.
- CLYNE M. (2000). « Lingua franca and ethnolects in Europe and beyond ». In : *Sociolinguistica*, n°14, p.83-89.
- DRANGE E.M. (2002). « Fremmedspråklige slangord i norsk ungdomsspråk ». In : A.B. STRENSTRÖM (ed.) (2002), *Jallaspråk, slanguage og annet ungdomsspråk i Norden*, p.9-17.
- ECKERT P. (2001). « Style and social meaning ». In : P. ECKERT et J.R. RICKFORD (eds) (2001), *Style and sociolinguistic variation*. Cambridge : Cambridge University Press, p.119-128.

- ECKERT P. (2008). « Variation and the indexical field ». In : *Journal of sociolinguistics*, n°12, p.453-476.
- FRAURUD K. et E. BIJVOET (2004). « Multietniskt ungdomsspråk och andra varieteter av svenska i flerspråkiga miljöer ». In : K. HYLSTENSTAM et I. LINDBERG (eds) (2004), *Svenska som andraspråk : I forskning, undervisning och samhälle*. Lund : Studentlitteratur, p.389-417.
- GADET F. et P. HAMBYE (2014) « Contact and Ethnicity in “Youth Language” Description : in Search of Specificity ». In : R. NICOLAÏ (ed.) (2014), *Questioning Language Contact. Limits of Contact, Contact at its Limits*. Leiden : Brill Academic Pub, p.183-216.
- HARCHAOU S. (2012). *Le parler jeune en milieu urbain. Analyse et réception des pratiques langagières adolescentes à Oslo*. Mémoire de Master 2. Université Paris-Sorbonne.
- HASUND I. (2006). *Ungdomsspråk*. Bergen : Fagbokforlaget.
- IRVINE J. (2001). « Style as distinctiveness: The culture and ideology of linguistic differentiation ». In : P. ECKERT et J. RICKFORD (eds) (2001), *Stylistic variation in language*. Cambridge : Cambridge University Press, p.21-43.
- KALLMEYER W. et I. KEIM (2003). « Linguistic variation and the construction of social identity in a German-Turkish setting. A case study of an immigrant youth group in Mannheim, Germany ». In : J. ANDROUTSOPOULOS et A. GEORGAKOPOULOU (eds) (2003), *Discourse constructions of youth identities*. Amsterdam : John Benjamins, p.29-46.
- KOTSINAS U.B. (1987). « Rinkebysvenskan – en dialekt ? ». In : P. LINELL, V. ADELWARD, T. NILSSON et P.A. PETERSSON (eds) (1987), *Svenskans Beskrivning*, vol.16. Linköping : Tema Kommunikation, Universitetet i Linköping, p.264-278.
- KOTSINAS U.B. (1988). « Immigrant children’s Swedish – A new variety ? ». In : *Journal of Multilingual and Multicultural Development*, n°1, 2, vol.9, p.129-140.
- LABOV W. (1976). *Sociolinguistique*. Paris : Editions de Minuit.
- LE PAGE R. et A. TABOURET-KELLER (1985). *Acts of identity. Creole-bases approaches to language and ethnicity*. Cambridge : Cambridge University Press.
- MADSEN L.M. (2008). *Fighters and Outsiders. Linguistic practices, social identities, and social relationships among youth in a martial arts club*. Thèse de doctorat en linguistique. Université de Copenhague.

NORTIER J. (2001). « “Fawaka, what’s up?” Language use among adolescents in Dutch monoethnic and ethnically mixed groups ». In : A. HVENEKILDE et J. NORTIER (eds) (2001), *Meetings at the crossroads. Studies of multilingualism and multiculturalism*. Oslo : Novus forlag, p.61-72.

NORTIER J. et B.A. SVENDSEN (2015). *Language youth and identity in the 21st century*. Cambridge : Cambridge University Press.

POTTER R.A. (1995). *Spectacular Vernaculars : Hip-hop and the Politics of Postmodernism*. Albany, NY : State University of New York Press.

OPSAHL T., U. RØYNELAND et B.A. SVENDSEN (2008). « ‘Syns du jallanorsk er lættis, eller ?’ – om taggen [lang = X] ». In : J.B. JOHANNESSEN et K. HAGEN (eds) (2008), *Språk i Oslo. Ny forskning omkring talespråk*. Oslo : Novus Forlag.

QUIST P. (2000). « Ny københavnsk ‘multietnolekt’ – om sprogbrug blandt unge i kulturelt heterogene miljøer. » In : *Danske talesprog*, tome 1. Copenhagen : Institut for Dansk Dialektforskning. C.A.Reitzels forlag, p.143-212.

QUIST P. (2008). « Sociolinguistic approaches to multiethnolect: Language variety and stylistic practice ». In : *International Journal of Bilingualism*, n°1 et 2, vol.12, p.43-61.

QUIST P. et B.A. SVENDSEN (2010). *Multilingual Urban Scandinavia. New Linguistic Practices*. Bristol : Multilingual Matters.

SVENDSEN B.A. et U. RØYNELAND (2008). « Mutiethnolectal facts and functions in Oslo, Norway ». In : *The International Journal of Bilingualism*, n°1 et 2, vol.12, p.63-83.

RAMPTON B. (2011). « From ‘Multi-ethnic adolescent heteroglossia’ to ‘Contemporary urban vernaculars’ ». In : *Language & Communication*, n°4, vol.31, p.276-294.

SSB, CENTRAL BUREAU OF STATISTICS, Norway (2008). [URL : <http://www.ssb.no/emner/02/>, consulté le 02/06/2014].

Politique linguistique de la Chine envers les dialectes : le débat autour du cantonais au début du 21^e siècle

Yufei Guo

INALCO, 65 rue des Grands Moulins 75 013 Paris
yfeigu@gmail.com

RÉSUMÉ

Cet article vise à étudier la politique linguistique chinoise, en prenant appui sur le cas du débat autour du cantonais depuis 2010. Cet incident qui a eu lieu à Guangzhou et à Hong Kong sous forme de manifestations durant l'été 2010 révèle un sujet longtemps mis à l'écart dans l'aménagement linguistique de la Chine : celui des « dialectes ». Nous aborderons d'abord la situation et la politique linguistique de la Chine aujourd'hui avant de nous focaliser sur une analyse portant sur la particularité du cantonais et sur notre enquête de terrain conduite pendant deux ans en 2013 et 2014 et qui est composée d'un sondage et d'entretiens.

ABSTRACT

Language policy of China: the debate over Cantonese at the beginning of the 21th century

This paper aims to study the language policy of China, building on the language controversy in Guangzhou since 2010. The incident, which happened in Guangzhou and Hong Kong in the form of demonstration during the summer of 2010, revealed a long-neglected subject in China's language policy and planning: the question of dialect. We begin with the China's language situation and its language policy since 1950. Then we will focus on the analysis of the particularity of Cantonese and of our field research conducted during 2013 and 2014, which combines a language survey and interviews.

MOTS-CLEFS : politique linguistique, cantonais, dialecte, diversité, Chine

KEYWORDS: language policy, Cantonese, dialect, diversity, China

1 Introduction

Cet article vise à étudier la politique linguistique chinoise en prenant appui sur le cas du débat autour du cantonais qui a lieu depuis 2010. La période étudiée est celle des années 1950 à aujourd'hui.

En Chine, le mandarin est désigné comme norme depuis les années 1950^{29, 30}. Mais il existe en parallèle de nombreux autres parlers sur le territoire. Face à l'influence de l'urbanisation, de la mondialisation et de la diffusion du mandarin, les dynamiques sociolinguistiques sont devenues de plus en plus complexes dans la société chinoise moderne.

Durant l'été 2010, un mouvement d'opinion sur « la défense du cantonais » a eu lieu à Guangzhou et à Hong Kong, provoquant de nombreuses retombées médiatiques^{31, 32} ainsi que des effets différés sur les plans académique^{33, 34} et politique^{35, 36} en Chine. Les manifestants y proclamaient la « défense du cantonais » contre la tendance dite hégémonique du mandarin dans les médias audiovisuels. Cette affaire a révélé un sujet

²⁹ Article 19, rédigé dans la Constitution depuis 1982 : « l'État promeut l'utilisation du mandarin à l'échelle nationale »

³⁰ Articles 9, 10, 12, 19, 20, rédigés dans la *loi sur la langue et l'écriture communes nationales*, promulgués en 2001

³¹ 推广普通话必灭粤语？广州激辩“粤语保卫战” (« Débat sur la défense du cantonais : la promotion du mandarin est-elle une condition préalable pour la répression du cantonais ? »), 新华网 (Site du Xinhua), consulté le 17/07/2010 sur http://news.xinhuanet.com/edu/2010-07/19/c_12345929.htm

³² 广州2000人上街集会“撑粤语” (« Deux mille habitants cantonais participent à la manifestation dans la ville du Guangzhou »), 香港明报 (*Journal quotidien Hongkongais Mingbao*), consulté le 26/07/2010, <http://www.afinance.cn/new/gncj/201007/282938.html>

³³ QU, S. (2011) 广州撑粤语运动引发的思考 (« Réflexion sur le mouvement sur « la défense du cantonais » en 2010 »), 云南师范大学学报 (*Journal académique de l'École normale de Yunan*), vol.43(1), 54-62

³⁴ PAN, W., R. YOU, Y. MAI, N. QIAN, H. ZHANG ET J. KONG (2010). 抢救日渐衰微的方言 (« Sauvegarder les dialectes en régression »), 上海社会科学报 (*Revue des sciences sociales de Shanghai*), vol.416.

³⁵ 广州市政府重申：“推普废粤”是根本不存在的伪命题 (« Le gouvernement du Guangzhou réitère : promouvoir le mandarin contre le cantonais est une fausse hypothèse »), 大洋网 (Dayang net), consulté le 28/07/2010 sur http://news.dayoo.com/guangzhou/201007/28/73437_13426378.htm

³⁶ 上海委员提出保护方言应成为国家战略 (« Un membre du CCPPC de Shanggai dépose une proposition écrite au gouvernement central : la protection des dialectes devrait devenir une stratégie de l'Etat »), 新浪网 (Xinlang net), consulté le 12/03/2013 sur <http://sh.sina.com.cn/news/b/2014-03-12/081685404.html>

longtemps mis à l'écart dans l'aménagement linguistique de la Chine : celui des « dialectes ». En effet, lorsque la campagne de promotion du mandarin fut lancée en 1955, peu d'attention fut accordée aux conditions d'existence des autres langues du groupe Han, conventionnellement nommées et regroupées sous le terme de « dialectes ». Avec l'évolution de la situation linguistique du pays et des consciences publiques, une remise en question sur la politique linguistique actuelle de la Chine est à l'ordre du jour.

Le débat autour du cantonais constitue le cas d'étude à partir duquel nous nous interrogeons : quelle est la situation sociolinguistique à Guangzhou ? Comment la question des langues interagit-elle avec les problèmes politiques, sociaux et culturels ?

2 Cadre théorique

Pour répondre à ces questions, nous adoptons un point de vue sociolinguistique, et plus précisément celui de l'approche dynamiste élaborée par Christiane Loubier (Loubier, 2008), schématisée dans la FIGURE 1. Cette approche suppose que l'analyse de la politique linguistique doit tenir compte de deux aspects : l'un sur l'autorégulation de la société (« politique linguistique *in vivo* » selon Calvet, 1996), et l'autre sur la régulation officielle (« politique linguistique *in vitro* », *ibid.*). L'aspect de l'autorégulation pourrait être analysé selon les forces sociolinguistiques de quatre catégories : la force symbolique qui est liée au système de représentations, la force fonctionnelle concernant l'utilisation effective des langues, la force démographique relative au nombre de locuteurs, et la force évolutive relative à l'évolution du temps.

Ce cadre d'analyse nous permet de repérer les forces sociolinguistiques liées à la régulation de l'usage des langues et d'en comprendre la dynamique d'interaction.

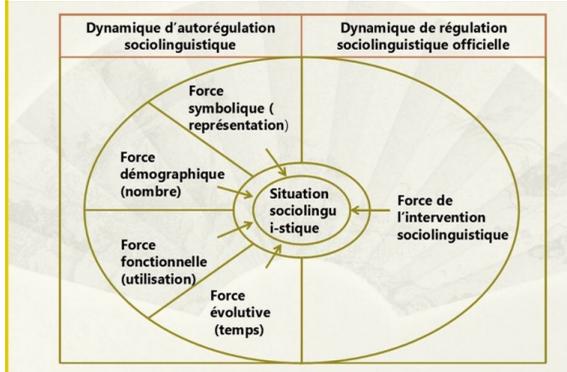


FIGURE 1 – L'aménagement linguistique : une double dynamique de forces sociolinguistiques (Loubier, 2008)

3 Dynamique de l'intervention sociolinguistique officielle en Chine depuis 1950

3.1 Contexte dans les années 1950

Map 4: Dialects of Mandarin and Southern Chinese

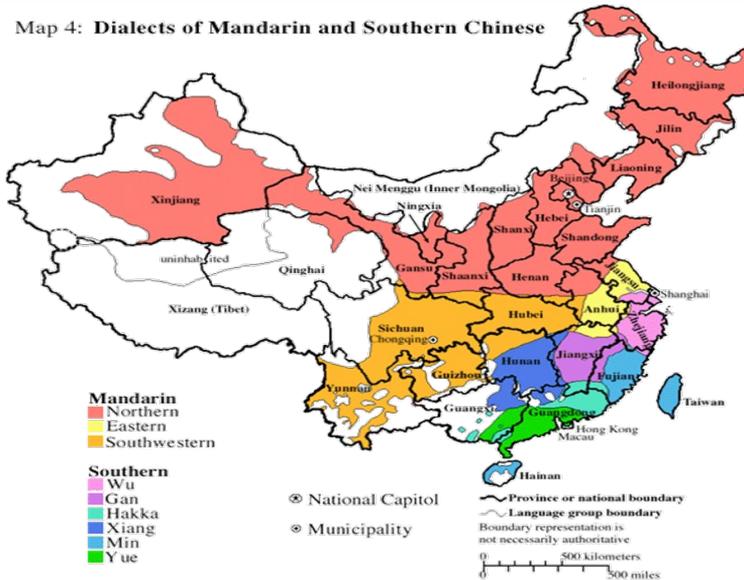


FIGURE 2 – Répartition géographiques des « dialectes » en Chine (Pennock, 2001)

La Chine est un pays multilingue avec des centaines de parlers sur son territoire. Les langues des groupes ethniques minoritaires mises à part, il existe sept ou huit variétés

linguistiques au sein du groupe majoritaire (Ramesey, 1987 ; Chen, 1999) : le mandarin (comprenant le mandarin du nord, de l'est et de sud-est), le wu, le xiang, le gan, le min, le hakka et le cantonais. Au sein de chaque variété, il existe de nombreuses sous-variétés. Le mandarin tel qu'il est parlé aujourd'hui dérive d'une sous-variété du mandarin, le pékinois. Selon la tradition, ces variétés et leurs sous-variétés sont regroupées sous le terme générique de « dialectes ». La FIGURE 2 montre la répartition géographique des groupes dialectaux.

La différence entre ces dialectes est très grande, à tel point qu'ils ne sont pas mutuellement intelligibles à l'oral. À l'écrit, il existe une écriture unifiée non interrompue depuis près de deux mille ans : le *wenyan*, ou chinois classique. C'est une langue sacrée utilisée dans les textes administratifs et les œuvres littéraires respectées, et elle n'a été maîtrisée que par les intellectuels, qui ne constituaient qu'un petit groupe de personnes à l'époque. Certaines variétés régionales (le mandarin, le wu et le cantonais) ont développé leur propre écrit vernaculaire entre le 14^e et le 17^e siècle. Ces écritures vernaculaires ont été utilisées dans les textes religieux, la littérature populaire, les chansons ou les opéras locaux (Ramesey, 1987 ; Chen, 1999 ; Liang, 2015). Au 16^e siècle, une langue orale, basée sur le pékinois, a été utilisée par convention tacite comme langue véhiculaire parmi les fonctionnaires. Les Chinois l'appelaient *langue des fonctionnaires* (Ramesey, 1987 ; Chen, 1999) et les missionnaires étrangers l'appelaient *mandarin*³⁷. Mais cette langue ne fut en aucun cas une langue standard : elle ne fut pas définie, standardisée ni diffusée parmi les gens ordinaires ; son influence fut encore plus restreinte dans les régions du sud.

Les premières interventions linguistiques officielles ne commencèrent pas avant la fin du 19^e siècle, quand les intellectuels et le gouvernement ressentirent le besoin d'établir une *langue nationale*. Mais à cette époque-là, on ne parvint pas à un consensus sur la norme de cette langue nationale (Li, 1924 ; Liang, 2015).

En 1932, la norme pour la langue nationale fut enfin clarifiée. Dans la *Circulaire numéro 3051 du Ministère de l'éducation*, publiée par la République de Chine, le pékinois et l'écriture vernaculaire du mandarin furent définis respectivement comme base pour la prononciation et l'écriture standard.

³⁷ Harper D. (2001). Mandarin, consultable sur <http://www.etymonline.com/index.php?term=mandarin>

3.2 Régulation officielle à l'égard du mandarin depuis 1950

C'est à partir des années 1950 que la politique linguistique chinoise fut établie et mise en place de façon systématique. En 1956, la langue de Pékin, choisie comme langue nationale dans les années 1930 (*Circulaire numéro 3051 du Ministère de l'éducation*, 1932), a été redéfinie, en Chine continentale, comme la langue commune ; l'écrit vernaculaire, basé sur la grammaire et le lexique des langues du nord (les variétés du mandarin), a été défini comme l'écriture standard du chinois moderne (*Directives sur la promotion du mandarin*, promulguées par le Conseil des affaires de l'État de la République populaire de Chine en 1956).

Des années 1950 jusqu'à nos jours, la priorité de la politique linguistique chinoise a été accordée à la standardisation et à la promotion du mandarin. En 1982, une clause ajoutée dans l'article 19 de la Constitution précisant que « l'État promeut l'usage du mandarin à l'échelle nationale » fournit la base juridique pour la promotion de cette langue comme langue officielle *de facto*. En 2001, la première loi linguistique de la Chine – la *Loi sur la langue et l'écriture communes nationales* – stipule que l'utilisation du mandarin est obligatoire dans l'administration publique (article 9), dans l'enseignement (article 10) et dans les médias (article 12). Parallèlement, une centaine de lois, de réglementations ou de circulaires à effet national publiées entre 1987 et 2005 contiennent des articles encourageant davantage l'usage du mandarin dans le domaine éducatif, médiatique, juridique et commercial³⁸.

De plus, une série d'activités de propagation a été créée pour renforcer le prestige du mandarin. Dans le domaine de l'éducation, par exemple, la promotion du mandarin est fixée comme un critère pour évaluer des « écoles modèles » (*Circulaire sur l'évaluation des écoles de modèle*, promulguée conjointement par le Ministère de l'Éducation et le Conseil des affaires de l'État en 2004).

Selon les chiffres de 2011, environ 70% de la population chinoise d'aujourd'hui savent parler mandarin, par rapport à 60% en 2010 et 50% en 1984 (Xie *et al.*, 2011).

3.3 Régulation officielle à l'égard des « dialectes »

Quant aux autres variétés linguistiques du groupe ethnique majoritaire, appelées « dialectes » selon la tradition, la position du gouvernement est à la fois ambiguë et

³⁸ Voir l'inventaire des textes juridiques sur la politique linguistique chinoise, consulté le 9 septembre 2014 sur http://www.china-language.gov.cn/8/2007_7_27/1_8_2931_0_1185508069552.html

contradictoire. Leur statut et leur usage ne sont pas précisés dans les textes juridiques.

En réalité, il existe de nombreuses restrictions, explicites ou implicites. Dans le domaine médiatique, l'usage des dialectes a été interdit progressivement dans le journal télévisé (1986), dans les films doublés (2004), dans les programmes éducatifs ou historiques (2005) et dans les programmes de divertissement (2014) (voir les règlements publiés par l'Administration Nationale de la Radiodiffusion, du Cinéma et de la Télévision, ANRCT par la suite). Dans le domaine éducatif, il est démontré que certaines écoles interdisent l'usage des dialectes dans le campus, même après les cours (voir les reportages médiatiques^{39, 40} et les travaux de Liang). En effet, le test sur le niveau du mandarin (système d'évaluation établi en 1998) exige désormais que les enseignants atteignent un certain niveau de mandarin pour avoir la qualification d'enseignant ; le système évaluatif des écoles modèles, mis en œuvre en 2004 (voir plus haut, partie 3.2), indique que la promotion du mandarin est un facteur important pour que les écoles soient sélectionnées comme modèles.

Tout cela implique une idéologie monolingue et une « violence symbolique » (terme emprunté à Bourdieu, 1982) à l'égard des « dialectes ».

3.4 Particularité à l'égard du cantonais

Le cantonais fait partie de sept variétés linguistiques au sein du groupe ethnique majoritaire chinois. Il est parlé principalement dans la région du Guangdong, dans le sud-est de la région du Guangxi, à Hong Kong et à Macao. En dehors de la Chine, les diasporas cantonophones sont dispersées à travers l'Europe, l'Amérique du Nord et l'Asie du Sud. Le nombre total des locuteurs du cantonais dans le monde entier est estimé à plus de 55 millions (Chen, 1999).

Même s'il est généralement considéré comme un « dialecte » en Chine continentale, le cantonais jouit d'un grand prestige. Quelques facteurs y contribuent : le lien historique, géographique et culturel entre Hong Kong et Guangzhou (capitale de la région du Guangdong), le développement économique des deux villes, la popularité des médias cantonophones, et une écriture vernaculaire bien établie (malgré une absence de norme

³⁹ 越秀区执信南路小学“要求”学生不管是上课还是课余时间都要讲普通话 (L'école primaire de Zhixinnanlu exige que leurs élèves parlent mandarin pendant et après les cours), publié le 09/07/2010, 羊城晚报 (Journal du soir de Yangcheng), p.A17.

⁴⁰ 好多广州细路唔识讲白话 (Beaucoup d'enfants natifs de Guangzhou ne savent pas parler cantonais), publié le 24/12/2008, 广州日报 (Journal quotidien de Guangzhou), p.A19

unifiée, il existe un haut niveau de consensus sur ce qui doit être représenté à l'écrit (Liang, 2015).

La différence dans l'intervention sociolinguistique officielle entre Guangzhou et d'autres villes chinoises réside dans l'usage médiatique. En 1988, l'ANRCT a donné l'autorisation aux chaînes télévisées locales de Guangzhou de diffuser des programmes en cantonais (Wu, 2004), ce qui a été une exception par rapport aux autres dialectes dont l'usage est strictement contrôlé (voir 3.3).

4 Enquête de terrain : essai de compréhension de l'interaction des dynamiques

Afin de mieux comprendre les dynamiques sociolinguistiques à Guangzhou, nous avons effectué une enquête de terrain pendant les années 2013 et 2014. L'enquête comprend deux grandes parties : un sondage effectué à Guangzhou, et des entretiens passés à Guangzhou, Hong Kong et Pékin.

4.1 Sondage : méthodologie et déroulement

Le sondage a été conduit sous la forme d'un questionnaire. L'objectif est de connaître la pratique et l'attitude langagières des habitants installés de façon permanente à Guangzhou.

À cause d'un trop grand écart dans la structure démographique entre les vieux et les nouveaux quartiers (la population migratoire occupe une large proportion dans les nouveaux quartiers, voir les recensements du Bureau de statistique de Guangzhou de 1991, 2001, 2010), nous avons choisi les vieux quartiers comme terrain, dans lesquels nous avons ciblé trois écoles comme lieux concrets pour la réalisation de l'enquête.

Ces trois écoles comprennent une école primaire et une école secondaire labellisées comme « modèles », et une école secondaire catégorisée comme « générale »⁴¹. Trois classes de chaque école ont été choisies de façon aléatoire.

Le sondage a été conduit sous forme d'un questionnaire conçu selon deux versions différentes : une pour les enfants et l'autre pour les adultes. Chaque version comprend des questions à choix unique ou multiple ainsi que des questions dont la réponse doit être rédigée en texte libre. Chaque version est composée de questions portant sur les

⁴¹ voir le classement des écoles par le Bureau de l'éducation à Guangzhou, consultable sur <http://www.gzedu.gov.cn/>

informations de base (lieu de naissance, âge, langue acquise en première langue vivante – L1 par la suite...), sur les pratiques langagières des sondés dans différentes situations (langue parlée dans la famille, à l'école, au travail...) et sur leur attitude envers le cantonais et envers le mandarin.

L'enquêteur est entré dans chaque classe, a fait remplir aux élèves les questionnaires dans leur version enfant et les a récupérés sur place. Ensuite, l'enquêteur a distribué les questionnaires dans leur version adultes aux élèves, leur demandé de les apporter à leurs parents, et les a récupérés le lendemain. Au total, un nombre de 659 questionnaires sont considérés comme validés.

Trois éléments sont pris en considération dans l'analyse : âge, lieu de naissance et catégorie sociale.

4.2 Entretiens : méthodologie et déroulement

Il s'agit d'entretiens avec des personnes en provenance de milieux sociaux différents, en relation plus ou moins directe avec la politique linguistique.

Les sujets comprennent des chercheurs en sociolinguistique, dialectologie ou sociologie en Chine continentale et à Hong Kong, des membres officiels du gouvernement central et local, des travailleurs des médias et de l'information, des membres d'associations universitaires ayant pour vocation de promouvoir la langue et la culture cantonaises. Les termes abordés comprennent leur opinion sur l'incident du cantonais en 2010, sur la vitalité du cantonais, sur le rapport entre le cantonais et le mandarin et sur la politique officielle envers les dialectes.

La plupart des entretiens ont été enregistrés, à l'exception d'un entretien avec un membre officiel du gouvernement travaillant pour l'aménagement linguistique du gouvernement central.

4.3 Analyse des résultats

Dynamique de la force évolutive et démographique

Pour les natifs, on a constaté un déclin de l'usage du cantonais et une assimilation linguistique (Weinreich, 1967 ; Macky, 1980 ; Giacalone Ramat, 1983 ; Sasse, 1992 ; Clyne, 2003) du cantonais vers le mandarin entre les deux générations dans la famille : chez les adultes, 82.5% ont appris le cantonais en L1 ; 69.8% pensent qu'ils sont encore plus à l'aise en parlant cette langue aujourd'hui, et 88.8% s'adressent à leurs parents en

cantonais la plupart du temps. Chez les enfants, 66.9% ont appris le cantonais en L1, mais ceux qui se sentent le plus à l'aise avec le mandarin (53.8%) sont plus nombreux que ceux qui prétendent que c'est le cantonais (40.6%) ; dans la conversation avec leurs parents, 55.5% utilisent le cantonais la plupart du temps, alors qu'il y en a 24.8% qui utilisent essentiellement le mandarin.

Pour les gens immigrés, le déclin de la transmission de leurs langues d'origine est plus saillant. Parallèlement, on a constaté une résistance contre le cantonais et une tendance au monolinguisme chez les enfants immigrés : 40.9% des adultes, contre 20% des enfants ont appris leur langue d'origine comme L1 ; 44.3% des adultes, contre 16.7% des enfants parlent encore la langue d'origine à leurs parents aujourd'hui. Chez les adultes, 74.6% prétendent savoir parler le cantonais, seulement 17.4% prétendent ne parler que le mandarin ; ceux qui aiment le cantonais (40.3%) sont plus nombreux à ceux qui ne l'aiment pas (18.8%). Chez les enfants, par contraste, seulement 46.6% disent qu'ils savent parler cantonais, alors que 43.3% prétendent ne savoir parler que mandarin aujourd'hui ; ceux qui n'aiment pas cette langue (48.4%) ont dépassé ceux qui l'aiment (25%).

Dynamique de la force fonctionnelle

Selon les résultats des questionnaires, le domaine d'usage du cantonais est essentiellement réservé à la famille. Cette langue est aussi fréquente dans les lieux dits « familiers » (l'usage du cantonais est fréquent dans la conversation avec des voisins et quand les gens vont au marché des viandes et des légumes). La langue occupe aussi une place dans le domaine socio-économique (66.7% des adultes rapportent qu'ils utilisent le cantonais ou un mélange de deux langues avec leurs clients au travail). Par contre, l'usage du mandarin prédomine dans le domaine sociopolitique et éducatif (chez les enfants, la plupart parlent essentiellement en mandarin avec leurs professeurs (71.14%), leurs camarades de classe (51.43%) et leurs meilleurs amis (48.57%) pendant et après cours), et cette langue a tendance de l'emporter de plus en plus dans la famille (voir partie 4.3.1).

Dynamique de la force symbolique

a) Parmi les gens ordinaires, la pratique langagière diffère selon les différentes catégories sociales. Trois aspects ont été pris en considération : la profession, le niveau d'étude et le type d'école (labellisé comme « modèle » ou « générale »).⁴²

⁴² En raison de la longueur limitée de l'article, nous ne pouvons pas montrer ici les statistiques en détails.

La plupart des enfants ont indiqué le métier de leurs parents dans le questionnaire. Selon leur réponse, les enfants dont les parents font un métier trouvé dans la base de la hiérarchie sociale parlent essentiellement en cantonais dans la famille ; par contraste, les enfants dont leurs parents exercent un métier socialement distingué parlent majoritairement en mandarin dans leur famille.

Comparé avec la pratique langagière dans les deux écoles modèles où le phénomène monolingue du mandarin domine, l'usage du cantonais dans l'école générale est assez fréquent dans la communication parmi les élèves ou entre les élèves et les enseignants après les cours.

Malgré notre intention de comparer la pratique langagière des adultes selon leur métier, la plupart n'ont pas répondu à la question de la profession dans le questionnaire. Heureusement, la plupart ont indiqué leur niveau d'étude. Selon leurs réponses, moins les adultes ont fait d'études, plus ils utilisent le cantonais ou un mélange de deux langues avec leurs collègues et leurs clients et dans les réunions de travail.

En ce qui concerne l'attitude langagière des gens ordinaires, le mandarin est toujours associé avec le caractère officiel et la fonction de communication, alors que le cantonais est toujours associé avec l'identité régionale, la représentation culturelle et la richesse en linguistique (par exemple, beaucoup ont mentionné que certains traits historiques du chinois ancien sont mieux préservés en cantonais qu'en mandarin). La grande majorité manifeste une volonté pour la transmission du cantonais, ce qui n'est pourtant pas sans contradiction avec leur pratique.

b) Chez les hommes politiques en relation avec l'aménagement linguistique, d'une part, ils adoptent une attitude affichée pour le plurilinguisme ; mais quand nous leur demandons en quoi consiste les mesures concrètes pour la promotion de ce plurilinguisme, ils déclarent que le gouvernement fera au mieux pour « sauvegarder et documenter les langues qui sont réellement en danger ». Mais « la régression de dialectes est le courant irrésistible de notre temps. (...) La politique linguistique du gouvernement ne peut pas faire autrement que de suivre cette tendance et de respecter la loi naturelle du développement des langues et écritures. (...) Toutes les tentatives qui essaient d'arrêter cette tendance sont vouées à l'échec » (extrait traduit de la transcription d'entretiens).

c) Pour les travailleurs des médias, on peut remarquer que, selon qu'ils sont de Guangzhou ou de Hong Kong, ils n'ont pas la même approche de l'incident du cantonais de 2010. Alors que les deux travailleurs à Guangzhou se focalisent plutôt sur la régression de la langue ou des problèmes sociaux entre les cantonais natifs et les immigrés, le journaliste hongkongais considère cet incident comme « une forme de résistance populaire contre le régime autoritaire du gouvernement de la Chine continentale ».

d) Chez les étudiants à la tête d'associations vouées à la promotion de la langue et culture cantonaises, la plupart considèrent la question des langues d'un point de vue culturel, avec une attitude « pacifique » et ouverte. Ils pensent que les différents parlers peuvent coexister sans conflits dans un espace social donné. Pourtant, on peut remarquer une certaine inquiétude sur la régression de la langue et culture locale dans leurs paroles.

e) Chez les chercheurs intéressés à l'incident du cantonais, chacun porte un point de vue différent. Certains associent le débat linguistique à des problèmes sociaux entre immigrés et natifs ; certains l'associent à la lutte pour la démocratie et les droits de l'homme ; certains pensent que le débat est dû à la conscience linguistique du peuple, renforcée par le développement des outils informatiques qui leur permettent de voir, échanger et diffuser des informations avec plus de rapidité et d'ampleur ; d'autres enfin en attribuent la cause principale au déclin de la vitalité des langues et cultures régionales.

5 Discussion

5.1 Situation sociolinguistique à Guangzhou

En la comparant avec les études précédentes sur l'usage et l'attitude du cantonais à Guangzhou (Kalmar *et al.*, 1987 ; Leung, 1993 ; Tang, 2006 ; Guo, 2004 ; Liang 2015), notre étude montre que Guangzhou est en train de connaître un changement sociolinguistique qui part d'une situation diglossique dominée par le cantonais il y a vingt ans vers une situation diglossique dominée par le mandarin aujourd'hui. La régression de la vitalité du cantonais se retrouve dans vie quotidienne des habitants à Guangzhou. Mais la régression des autres dialectes est encore plus évidente dans la ville. Le mandarin domine dans tous les domaines publics aujourd'hui, et il a tendance à pénétrer de plus en plus dans les domaines privés.

La pratique et l'attitude des gens peuvent varier selon les différentes catégories sociales. Plus les adultes occupent une place importante dans la hiérarchie sociale (ayant un travail socialement distingué ou possédant un diplôme de haut niveau) ou plus les enfants font leurs études dans une école sélective, plus ils ont tendance de parler uniquement mandarin dans la vie quotidienne.

Pourtant, la vitalité sociolinguistique du cantonais reste dynamique à Guangzhou. Dans notre recherche, la plupart des natifs mettent en lumière la valeur identitaire, culturelle et historique du cantonais et montrent une volonté pour la transmission de leur langue. De 1988 à nos jours, la place dominante du cantonais dans les médias locaux n'a pas connu de grand changement. D'ailleurs, la proximité avec Hong Kong où le cantonais est utilisé comme une langue officielle et véhiculaire influence l'usage et le prestige de cette langue à Guangzhou de manière positive.

5.2 Synergie entre la langue et la société à Guangzhou

La dynamique sociolinguistique interagit en permanence avec d'autres dynamiques sociales. Dans le cas de notre recherche, Guangzhou est toujours une ville dotée d'une certaine richesse économique ; elle est étroitement liée à Hong Kong sur les plans géographique, historique et culturel. Depuis les années 1980, la structure démographique et culturelle de la ville sont influencées par l'immigration et l'urbanisation ; l'industrie médiatique cantonophone, qui a connu une grande popularité dans toute la Chine depuis les années 1980-1990, est en train de vivre une certaine régression depuis le début du 21^e siècle face au développement médiatique de la Chine continentale⁴³.

⁴³ « Les chansons cantonophones : la noblesse du passé », Fenghuang net, consulté le 5 février 2012 sur <http://city.ifeng.com/fhzl/20120214/217218.shtml>

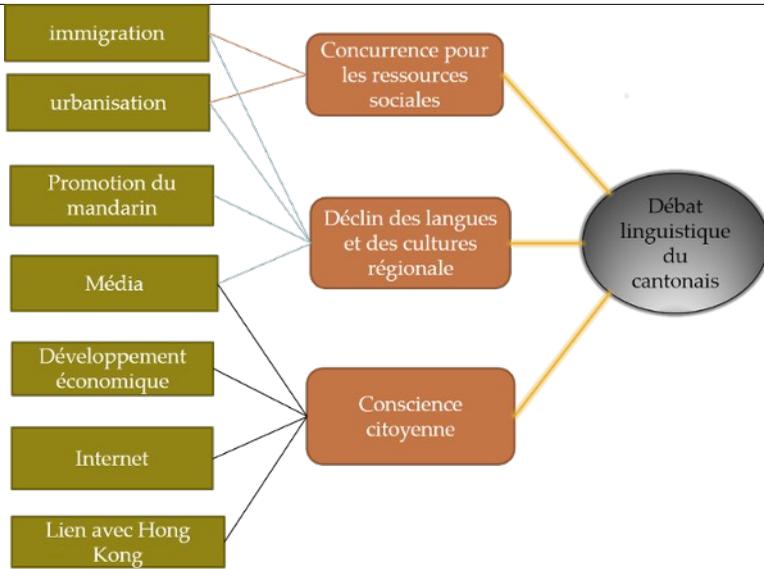


FIGURE 3 – Schéma montrant les causes pour débat linguistique du cantonais

À partir de notre recherche, nous avons développé un schéma montrant les synergies entre la langue et la société à Guangzhou (Figure 3) : comme illustré, l'immigration et l'urbanisation ont suscité une certaine concurrence pour les ressources sociales entre des habitants natifs et immigrés ; ces deux facteurs, accompagnés des effets de la politique linguistique et de la régression médiatique cantonophone, ont entraîné le déclin linguistique et culturel de la ville. Le pouvoir économique de la ville, le développement informatique de notre siècle, la proximité avec Hong Kong et la promotion de l'identité culturelle dans les médias ont conjointement contribué à l'émergence et au renforcement de la conscience citoyenne, y compris la défense des droits de l'homme, la volonté de participer activement à la construction de leur ville et le désir de préserver leur identité et une mémoire collective au cours de l'urbanisation. Ces trois éléments, à savoir la concurrence pour les ressources sociales, le déclin linguistique et culturel et la conscience citoyenne, interagissent avec l'un l'autre ; ils ont conjointement entraîné le débat linguistique autour du cantonais, et ils sont en voie de se reconstruire à partir de ce débat.

Références

BALDAUFF R.B. (1997). *Language planning from practice to theory*. Clevedon, England : Multilingual Matters.

- BOURDIEU P. (1982). *Ce que parler veut dire : l'économie des échanges linguistiques*. Paris : Fayard.
- CALVET L.-J. (1996). *Les politiques linguistiques*. Paris : Presses universitaires de France.
- CALVET L.-J. (2002). « Mondialisation, langues et politiques linguistiques », *Le français dans le monde*, n°323, p.39-42.
- CHEN P. (1999). *Modern Chinese-History and sociolinguistics*. Cambridge : Cambridge University Press.
- COOPER R.L. (1989). *Language Planning and social change*. Clevedon Hall : Cambridge University Press.
- COSTA J. (2010). *Revitalisation linguistique : discours, mythes et idéologies : approche critique de mouvement de revitalisation en Provence et en Écosse*. Thèse de doctorat en linguistique. Université Grenoble 3.
- FISHMAN J.A. (1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Clevedon, England : Multilingual Matters.
- GUO X. (2004). 广州市语言文字使用情况调查报告 (Rapport sur l'usage des langues et de l'écriture à Guangzhou), *中国社会语言学 (Journal de la sociolinguistique de Chine)*, n°2.
- LEE K. et W. LEUNG (2012). « The status of Cantonese in the education policy of Hongkong », *Multilingual Education*, vol.2, n°1.
- LI J. (1924). *国语运动史纲 (Histoire du Mouvement sur la langue nationale)*, Beijing : Shangwu yingshuguan.
- LIANG S. (2015). *Language Attitudes and Identities in Multilingual China: a linguistic ethnography*. Springer.
- LOUBIER C. (2008). *Langues au pouvoir : politique et symbolique*. Paris : L'Harmattan.
- MACKAY W. (1980). « The ecology of language shift ». In P.H. NELDE (ed.), *Sprachkontakt und sprachkonflikt*. Wiesbaden : Steiner, p.35-41.
- NORMAN, J. (1988). *Chinese*. Cambridge : Cambridge University Press.
- PENNYCOOK A. (2002). « Mother tongues, governmentality, and protectionism ». *International Journal of the sociology of language*, vol.154, p.11-28.

- POON A. (2010), Language use, and language policy and planning in Hong Kong, *Current issues in language planning* : 11:1, 1-60.
- RAMSEY R. (1987). *The Languages of China*. Princeton : Princeton University Press.
- RICENTO T. (ed.) (2006). *An introduction to language policy : Theory and method*. Oxford, UK : Wiley-Blackwell Publishing.
- SASSE H-J. (1992). Theory of language death. In M. BREZINGER (ed.), *Language death: factual and theoretical explorations with special reference to East Africa*. Berlin, New York : Mouton de Gruyter, p.7-30.
- TANG Y. (2006). 广州市中学生语言态度研究 (On language attitude towards middle school students in Guangzhou city). Mémoire de master. Université Jinan (暨南大学), Guangzhou.
- WEINRICH U. (1953). *Language in contact: Findings and problems*. New York : Publications of the linguistic circle of New York, n°1.
- WU C. (2004). 粤语南方卫视正式启播 (La chaîne télévisée par satellite en cantonais démarre officiellement), [URL : <http://xinxifb.act.tvscn.com/news/newscontent/4950/1.htm>, consulté le 20/12/2010].
- XIE J., W. LI, S. YAO et W. WEI. (2011). 普通话普及情况调查分析 (L'enquête sur le résultat de la promotion du mandarin en Chine), *语言文字应用 (Linguistique appliquée)*, vol. 3. Beijing : Ministère de l'éducation de la République populaire de Chine.
- XUE C. (2009). 上海市民语言生活状况调查 (An investigation into the language situation of Shanghai citizens), *语言文字应用 (Applied linguistics)*, vol.2, p.75-83.
- YANG J. (1985), « Problèmes de chinois contemporain ». In J. MAURIS (ed.), *La crise des langues*. Paris : Le Robert, p.421.
- ZHOU M. et H. SUN (2004). *Language Policy in the People's Republic of China: Theory and practice since 1949*, vol.4. Norwell, MA : Kluwer Academic Publishing.
- UNESCO (2009). « Investir dans la diversité culturelle et le dialogue interculturel », *Rapport mondiale de l'UNESCO*, Paris : Éditions UNESCO.
- 广州市政协建议 : 广州电视台改用普通话 (CCPPC de Guangzhou propose : les chaînes télévisées de Guangzhou devront changer la langue de diffusion en mandarin),

(publié le 09/06/2010). *Journal quotidien métropolitain de Nanfang* (南方都市报), p.A12

MINISTÈRE DE L'ÉDUCATION & COMMISSION D'ÉTAT SUR LA LANGUE ET ÉCRITURE (2004), 关于开展语言文字规范化示范校创建活动的意见 (Conseils pour établir des écoles modèles en matière d'usage de la langue et d'écriture standard), [URL : http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_5/200501/5438.html, consulté le 27/09/2014].

撑粤语的背后 (Derrière le débat du cantonais), [URL : <https://www.youtube.com/watch?v=NI-meu9u7Lg>, consulté le 21/05/2013].

Les verbes de mouvement italiens : stratégies de lexicalisation des natifs

Rosa RUSSO^{1,2}

(1) Université Paris VIII, 2 rue de la liberté, 92 526 Saint-Denis Cedex

(2) Université de Naples « Federico II », Corso Umberto I, 40bis, 80 138 Naples

rrosarusso@gmail.com

RÉSUMÉ

Dans cette étude nous présentons les résultats d'une recherche en cours sur les stratégies de lexicalisation des événements de mouvement en italien. En partant de la théorie de Talmy (2000) selon laquelle les langues peuvent être divisées en deux types, langues à satellite (*satellite-framed*) et langues à cadrage verbal (*verb-framed*) parmi lesquelles on classe l'italien. L'analyse montre que, contrairement à ce que Talmy a supposé, l'italien exprime la Trajectoire non seulement dans la racine verbale mais aussi dans des satellites (*fuggire via*, litt. « fuir loin/fuir en s'éloignant »). Par conséquent, notre étude de l'italien vise à montrer qu'il est souhaitable d'établir un continuum entre les langues à cadrage verbal et les langues à satellites, et sortir de la dichotomie talmyenne.

ABSTRACT

The Italian motion verbs: the lexicalization strategies in language of native speakers

The study discusses the results of a research, in progress, regarding the strategies of lexicalization of motion events in Italian based on the theory by Talmy (2000): the typological distinction between Satellite-framed and Verb-framed languages, Italian being part of the second mentioned. The analysis underlines that the Italian language, in contrast to the assertions made by Talmy, expresses the Trajectory using not only synthetic constructions but also relies on German based patterns (Verb + Particle: *fuggire via*). As a consequence, using the theory by Talmy as a platform from which extending the idiosyncrasies of the Italian, the research aims at illustrating that the picture of the Italian language offers a much bigger variety than supported by the Talmian dictonomy.

MOTS-CLÉS : typologie des langues, espace, italien l1, psycholinguistique, verbes de mouvement

KEYWORDS: linguistics typology, space, l1 italian, psycholinguistics, motion events.

1 Introduction

Les recherches consacrées à la sémantique de l'espace, ou bien aux moyens linguistiques employés pour exprimer les relations spatiales, suscitent beaucoup d'intérêt au sein de la communauté scientifique. En témoignent de multiples études menées dans des perspectives nombreuses et variées (Becker et Carroll, 1997 ; Talmy, 2000 ; Bowerman et Levinson, 2001 ; Slobin, 2004 ; Hickmann et Hendriks, 2011).

Dans ce travail on entend présenter les résultats préliminaires d'un projet plus large concernant l'expression de l'espace dynamique en L1 et en L2. Plus précisément, notre étude examine l'expression du déplacement par les verbes de mouvement en italien langue maternelle (L1) en adoptant comme base théorique de départ la classification sémantico-cognitive de Talmy (2000).

2 Le cadre théorique

Cette recherche s'inscrit dans le cadre théorique de la typologie lexicale, discipline qui vise à identifier le moyen avec lequel les langues lexicalisent les concepts, ou bien la façon dont les locuteurs des langues encodent certaines catégories sémantiques.

Selon le schéma universel proposé par Talmy (2000), un événement de mouvement réunit quatre éléments conceptuels distincts :

- *figure* (une entité qui se déplace ou qui est localisé),
- *motion* (l'expression du mouvement),
- *path* (la trajectoire du mouvement suivie par la Figure),
- *ground* (entité à laquelle la Figure est reliée dans l'espace).

Outre ces quatre composantes, Talmy propose d'inclure des traits secondaires (*coevents*), comme la Manière ou la Cause du mouvement qui transforment un événement simple en événement complexe. Selon la classification typologique de Talmy, c'est l'expression de la Trajectoire qui permet de différencier les langues en types, telles que les langues à satellites (*satellite-framed language*) vs langues à cadrage verbal (*verb-framed languages*). Les premières (par exemple les langues germaniques mais aussi les langues slaves) expriment la Trajectoire dans des satellites – une catégorie grammaticale représentée, par exemple, par un préfixe ou une particule verbale – et lexicalisent la Manière et la Cause du mouvement dans la racine verbale.

En revanche, les deuxièmes (par exemple les langues romanes) lexicalisent la Trajectoire du mouvement dans le verbe principal et laissent la Manière et la Cause du déplacement implicites ou périphériques.

Toutefois, cette classification typologique présente des limites, ce qui a amené beaucoup de chercheurs à formuler des propositions alternatives⁴⁴. Le débat entre les différents spécialistes montre que les langues choisissent des stratégies distinctes pour exprimer le mouvement. Donc, la question qui se pose est : quelles stratégies et quels éléments syntaxiques et sémantiques appartiennent au domaine spatial italien et dans quel mesure on peut considérer l'italien une langue *verb-framed* ? En effet, Schwarze (1985) remarque qu'en italien il y a des constructions dans lesquelles les verbes comme *cadere* (« tomber ») ou *entrare* (« entrer ») sont accompagnés par un satellite comme en (1). Bernini (2006 : 68) ajoute la composante déictique et distingue la trajectoire du mouvement et pourtant dans une phrase comme en (2), le verbe *viene* indique le mouvement et le déictique, *fuori* la trajectoire et *dalla casa* la relation spatiale. En outre, il montre qu'en italien cohabitent différents types de lexicalisation des événements de mouvement, par exemple *va giù/cade*⁴⁵ – *è andato via/è partito*⁴⁶.

(1) La mela cade giù dall'albero.⁴⁷

(2) Il gatto viene fuori dalla casa di corsa.⁴⁸

Ces perspectives nous permettent de focaliser sur la langue italienne.

3 L'expression de l'espace en italien

Dans les dernières années plusieurs travaux ont démontré que l'expression du mouvement en italien ne correspond pas du tout aux descriptions de Talmy. En effet, l'italien présente des traits du « type germanique », ou bien les constructions analytiques « verbe + particule » comme *venire fuori* ou *andare via* qui s'étendent de plus en plus (Schwarze, 1985 ; Simone, 1997 ; 2008, Bernini 2006a, 2006b, 2008). Ces types de construction sont appelés « verbes syntagmatiques » (VS), c'est-à-dire des

⁴⁴ Cf. Slobin (2004).

⁴⁵ Tomber.

⁴⁶ Partir.

⁴⁷ La pomme tombe de l'arbre.

⁴⁸ Le chat sort de la maison en vitesse.

« syntagmes formés par une tête verbale et un complément constitué par une “particule” (à l’origine un adverbe), unis par une cohésion syntaxique à un degré tel qu’il n’est pas possible de commuter le verbe syntagmatique entier avec un seul de ses éléments »⁴⁹. On peut illustrer ce point à l’aide des exemples en (3) :

- (3) *venire giù* > descendre ;
tirare fuori > faire sortir

La structure morphosyntaxique de ces constructions montre que l’italien possède des caractéristiques des langues à satellites, tout en conservant celles des langues à cadrage verbal (Schwarze 1985 : 361). Sans aucun doute, ces constructions rapprochent l’italien des langues à satellites, si on pense aux verbes à particule (*phrasal verbs*) de l’anglais. Par conséquent, parmi les langues romanes, qui feraient typiquement partie des langues à cadrage verbal, l’italien représenterait un cas à part puisque même si la présence de certains verbes syntagmatiques a été observée dans les autres langues romanes, la langue italienne présente une catégorie compacte et productive (Iacobini et Masini 2006 ; Calvo Rigual, 2008 ; Iacobini 2009).

Dans les paragraphes qui suivent, nous analyserons les productions de locuteurs natifs italophones et réexaminerons les combinaisons des verbes de mouvement et des prépositions spatiales.

4 Méthode

4.1 Les données

L’étude porte sur 24 récits oraux produits par des natifs italophones. Ici, nous analyserons seulement les données de 5 locuteurs – sachant que l’échantillon est à enrichir – qui ont produit 160 énoncés.

Les informateurs ont entre 20 et 30 ans et ont un niveau d’instruction élevé.

Les productions ont été transcrites grâce au logiciel CLAN⁵⁰ qui nous a permis de coder toutes les dimensions spécifiques qui font l’objet de l’analyse et de poser des repères pour des recherches rapides et automatisées sur ces composantes particulières.

⁴⁹ Simone (1996 :49).

⁵⁰ CLAN est un logiciel de transcription qui est associé une base de données (CHILDES). Il a été créé par Brian MacWhinney (cf. MacWhinney, 2000).

4.2 La tâche

Le corpus est constitué de récits produits à partir d'une tâche⁵¹ qui consiste à décrire 32 dessins animés en couleurs où un agent qui s'appelle Popi effectue une action (tirer, pousser) et provoque le déplacement d'objets selon une certaine trajectoire (entrer, traverser, monter, descendre) et d'une certaine manière (rouler, glisser), par exemple *un uomo attraversa la strada spingendo una ruota che rotola fino all'altro lato*⁵².

Les locuteurs visionnent individuellement les mini-clips et les racontent à l'expérimentateur en répondant à la question : « Que s'est-il passé ? ». La tâche expérimentale amène le locuteur à exprimer les différentes composantes du mouvement : la Cause, la Manière de la cause, la Manière du déplacement de l'objet, la Manière du déplacement de l'agent et la Trajectoire.

5 Analyse de données

L'analyse des productions nous permet de mieux reconnaître la distribution des composantes de l'événement du mouvement et de nuancer la dichotomie de Talmy (2000), à savoir que les langues à cadrage verbal expriment le mouvement exclusivement à travers les constructions synthétiques, ou bien à travers la racine verbale.

Les résultats de l'analyse portent sur 3 dimensions de réponses :

1. la densité sémantique ;
2. le focus de l'information ;
3. le locus de l'information.

En ce qui concerne la densité sémantique, l'analyse concerne le nombre des composantes sémantiques exprimées dans chaque réponse produite par les locuteurs. Vu que les langues à cadrage verbal nécessitent plusieurs énoncés pour exprimer plus d'une information, on s'attend à ce que l'italien présente un pourcentage bas.

Nous rappelons que la tâche employée permet d'indiquer cinq types de composantes : la Cause [C], la Trajectoire [T], la Manière du mouvement de l'agent [MAg], la Manière de la cause [Mac] et la Manière du mouvement de l'objet [Mob] illustrées dans

⁵¹ La tâche a été empruntée à des recherches antérieures en acquisition L1 et L2 (Hendriks, Hickmann et Demagny, 2008 ; Hickmann et Hendriks, 2011 ; Demagny, 2013).

⁵² Un homme traverse une route en poussant une roue qui roule jusqu'à l'autre côté.

l'exemple (4) :

- (4) Popi si trova in una foresta e trascinando [C+M] un carrellino di colore verde alle sue spalle entra [M+T] in una grotta.⁵³

Contrairement aux attentes, les résultats montrent que les réponses des italophones sont moyennement denses (42%), donc les résultats ne correspondent pas au profil typologique des langues à cadrage verbal.

Par rapport au focus de l'information, il s'agit d'identifier le type de composantes sémantiques le plus souvent exprimées. Ces composantes sont exprimées seules ou en association avec une ou plusieurs autres comme montre l'exemple (5) :

- (5) Popi si trova in una foresta e spinge [C+M] un grosso pneumatico all'interno di una grotta entrando [M+T] a sua volta nella grotta.⁵⁴

On note que les combinaisons sémantiques ne varient pas en fonction des types d'événements présentés. Dans notre corpus, les italophones préfèrent exprimer surtout la Cause (11%), la Manière de la Cause (41%) et la Trajectoire (48%) et ils ne sont pas sensibles à la Manière du mouvement de l'objet et à la Manière du mouvement de l'agent exprimées très rarement.

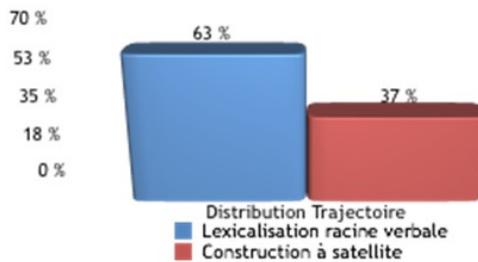


FIGURE 6 – Distribution de la Trajectoire

Enfin, le locus de l'information, qui fait référence aux stratégies employées pour exprimer l'information en deux locus : la racine verbale et les satellites. La Figure 1

⁵³ Popi se trouve dans une forêt et en traînant un chariot de couleur verte derrière soi il entre dans une grotte.

⁵⁴ Popi se trouve dans une forêt et il pousse un gros pneu à l'intérieur d'une grotte en entrant à son tour dans la grotte.

montre en termes de pourcentage la distribution de la Trajectoire dans notre échantillon.

En observant le graphique de la Figure 1, on note que la lexicalisation de la Trajectoire dans les constructions à satellite est tout à fait secondaire. En effet, dans notre échantillon dans 63% des cas la Trajectoire est lexicalisée dans la racine verbale et dans 37% des cas la lexicalise dans la périphérie.

Quant au type de verbes utilisés, nous observons qu'un des verbes qui se présente fréquemment est *spingere* (« pousser ») employé en deux modalités : tout seul comme illustré dans l'exemple (6) et en combinaisons avec un satellite (7) :

(6) Popi **spinge** una palla.⁵⁵

(7) Lui **spinge** in avanti la palla.⁵⁶

Dans l'exemple (6), la Trajectoire est lexicalisée dans la racine verbale, tandis qu'en (7) le verbe associé à l'adverbe de direction (*in avanti*) ne code pas la Trajectoire mais représente un intensificateur de mouvement.

En outre, souvent, la Trajectoire est lexicalisée dans le même énoncé à la fois dans la racine verbale et avec un syntagme prépositionnel (8) :

(8) Popi **entra** in una grotta.⁵⁷

L'exemple (8) montre que la localisation peut être exprimée et à travers le verbe *entrare* (« entrer ») et par la préposition *in* puisque la Trajectoire est lexicalisée dans le verbe mais le satellite exprime le fond avec un syntagme prépositionnel dont la préposition reprend le type de Trajectoire.

Un autre phénomène marginal mais existant est l'usage d'un verbe qui exprime la localisation et qui est suivi par un adverbe de direction qui produit une information redondante (9) :

(9) Popi da giù **sale verso** su.⁵⁸

⁵⁵ Popi pousse une boule.

⁵⁶ Il pousse en avant une boule.

⁵⁷ Popi entre dans une grotte.

⁵⁸ Popi du bas monte vers le haut.

Le verbe *salire* (« monter ») contient la Trajectoire mais il est suivi aussi d'un satellite qui complète la description de la Trajectoire et y ajoute une référence déictique.

En ce qui concerne les verbes de manière, leur présence dans les récits ne semble pas négligeable (38%) ce qui est plutôt en contradiction avec l'hypothèse de Talmy (2000), à savoir que dans les langues à cadrage verbal l'usage des verbes de manière n'est pas significatif.

De plus, en observant les données, on entrevoit la combinaison « verbe de manière + direction » comme le montre l'exemple (10) :

(10) Popi **trascina** una sedia **all' interno** della caverna verde.⁵⁹

Ce type de combinaisons, comme montré par Iacobini et Masini (2007), modifie l'*Aktionsart* du verbe en ajoutant des informations de nature aspectuelle comme la télélicité.

6 Conclusion

Les résultats préliminaires présentés dans cet article nous permettent d'affirmer que dans certains cas, les locuteurs choisissent les moyens linguistiques qui reflètent les caractéristiques des langues à cadrage verbal en exprimant la Trajectoire dans les verbes synthétiques. En revanche, dans d'autres cas, le choix des locuteurs renvoie au modèle des langues à satellite. Pour cette raison la langue italienne représente un cas de variation intratypologique. Au sein de la famille des langues romanes typiquement à cadrage verbal, l'italien suit le modèle des langues à satellites au moins en ce qui concerne les stratégies de lexicalisations des verbes de mouvement. Cette particularité rend l'italien difficile à classer en termes binaires et indique que son statut de langue à cadrage verbal dans la typologie de Talmy (2000) devrait être nuancé. Cela confirme l'idée qu'il faut réviser l'hypothèse de Talmy (2000) et que plutôt que parler d'une dichotomie très tranchée il vaut mieux parler d'un continuum des traits linguistiques dont disposent différentes langues ce qui permet de les classer comme langues à satellites ou comme langue à cadrage verbal, ou encore comme ayant les caractéristiques partagées.

⁵⁹ Popi traîne une chaise à l'intérieur de la grotte verte.

Références

- BECKER A. et M. CARROLL (1997). *The Expression of Spatial Relations in a Second Language*. John Benjamins : Amsterdam.
- BERNINI G. (2005). « Strategie di lessicalizzazione : tipologia e apprendimento. Il caso dei verbi di moto ». *Deutscher Romanistentag "Spracherwerb in und um die Romania"* (Saarbrücken ; 25-28 septembre 2005).
- BOWERMAN M. et C.S. LEVINSON (2001). *Language Acquisition and Conceptual Development*. Cambridge : Cambridge University Press.
- CALVO RIGUAL C. (2008). « I verbi sintagmatici italiani con appunti contrastivi con lo spagnolo e il catalano ». In : C. GÓNZALEZ ROYO et P. MOGORRÓN HUERTA (eds) (2008), *Estudios y análisis de fraseología contrastiva : lexicografía y traducción*. Universitat d'Alicant, p.37-66.
- HICKMANN M. (2007). « Static and dynamic location in French : Developmental and cross-linguistic perspectives ». In : M. AURNAGE (ed.) (2007), *The categorization of spatial entities in language and cognition*. Amsterdam : John Benjamins, p.205-231.
- HENDRIKS H. et M. HICKMANN (2011). « Space in second language acquisition ». In : V. COOK et B. BASSETTI (eds) (2011), *Language and Bilingual Cognition*. Hove, UK : Psychology Press, p.315-339.
- IACOBINI C. et F. MASINI (2007). « Verb-particle constructions and prefixed verbs in Italian: typology, diachrony and semantics ». In : G. BOOLJ, B. FRADIN, A. RALLI et S. SCALISE (eds) (2007), *On-line Proceedings of the Fifth Mediterranean Morphology Meeting (MMM5)* (Fréjus ; 15-18 septembre 2005). University of Bologna, p.157-184.
- KOPECKA A. (2006). « The semantic structure of motion verbs in French : Typological perspectives ». In : M. HICKMANN et S. ROBERT (eds) (2006), *Space acrosslanguages : Linguistic systems and cognitive categories*. Amsterdam : John Benjamins, p.83-101.
- MASINI F. (2006). « Diacronia dei verbi sintagmatici in italiano ». In : *Archivio Glottologico Italiano*, XCI, p.67-105.
- PORQUIER R. (2001). « 'Il m'a sauté dessus', 'je lui ai couru après': un cas de postposition en français ». In : *Journal of French Language Studies*, n°1, vol.11, p.123-134.

PORQUIER R. (2003). « ‘Gli corro dietro’ / ‘Je lui cours après’ : à propos d’une construction verbale spécifique en italien et en français ». In : M. GIACOMO-MARCELLESIE et A. ROCCHETTI (eds) (2003), *Il verbo italiano. studi diacronici, sincronici, contrastivi, didattici, Atti del XXXV Congresso internazionale di Studi della Società di Linguistica Italiana* (Paris ; 20-22 septembre 2001). Roma : Bulzoni, p.491-500.

SCHWARZE C. (1985). « ‘Uscire’ e ‘andare fuori’ : struttura sintattica e semantica lessicale ». In : *Sintassi e morfologia della lingua italiana d’uso. Teorie e applicazioni descrittive, Atti del XVII Congresso internazionale di studi di Urbino* (Urbino ;11-13 septembre 1983). Roma : Bulzoni, p.355-371.

SIMONE R. (2008). « Verbi sintagmatici come costruzione e come categoria ». In : M. CINI (ed.) (2008), *I verbi sintagmatici in italiano e nelle varietà dialettali. Stato dell’arte e prospettive di ricerca*. Frankfurt : Peter Lang, p.13-30.

SIMONE R. (1996). « Esistono verbi sintagmatici in italiano ? ». In : T. DE MAURO et V. LO CASCIO (eds) (1996), *Lessico e grammatica. Teorie linguistiche e applicazioni lessicografiche*. Rome : Bulzoni, p.155-170.

SLOBIN D.I. (1991). « Learning to think for speaking : Native language, cognition, and rhetorical style ». In : *Pragmatics*, 1, pages 7-26.

SLOBIN D.I. (1996). « From ‘Thought and language’ to ‘thinking for speaking’ ». In : J.J. GUMPERZ et S.C. LEVINSON (eds) (1996), *Rethinking Linguistic Relativity*. Cambridge University Press, p.70-96.

SLOBIN D.I. (1996). « Two ways to travel: Verbs of motion in English and Spanish ». In : M. SHIBATANI et S. THOMPSON (eds) (1996), *Grammatical constructions. Their Form and Meaning*. Oxford : Clarendon Press, p.195-317.

SLOBIN D.I. (2004). « The many ways to search for a frog: linguistic typology and the expression in motion events ». In : S. STRÖMQVIST et L. VERHOEVEN (eds) (2004), *Relating events in narrative : vol.2, Typological and contextual perspectives*. Mahwah, NJ : Lawrence Erlbaum Associates, p.219-257.

STUTTERHEIM C. Von (2002). « Différences translinguistiques dans la conceptualisation des événements ». In : *Revue française de linguistique appliquée*, VII, p.99-115.

TALMY L. (1983). « How language structures space ». In : H. PICK et L. ACREDOLO (eds) (1983), *Spatial Orientation : theory, research and application*. New York : Plenum, p.225-282.

TALMY L. (1985). « Lexicalization patterns: semantic structure in lexical forms ». In : SHOPEN T. (ed.) (1985), *Language typology and syntactic description, vol.III, Grammatical categories and the lexicon*. Cambridge : Cambridge University Press, p.57-149.

TALMY L. (2000). *Toward a cognitive semantics*. Cambridge, MA : MIT PRESS.

WATOREK M., S. BENAZZO et M. HICKMANN (2012). *Comparative Perspectives on Language Acquisition. A tribute to Clive Perdue*. Clevedon : Multilingual Matters.

Diversité dans l'expression de l'espace dynamique en L1 : étude contrastive entre italien, français et anglais

Simona Anastasio^{1,2,3}

(1) STL UMR 8164, Université Charles de Gaulle Lille 3, 59 650 Villeneuve d'Ascq

(2) Structures formelles du langage, UMR 7023, Université Paris 8, 93 526 Saint Denis

(3) Dipartimento Studi Umanistici, Université de Naples Federico II, 80 133 Naples
anastasiosimona@libero.it

RÉSUMÉ

Notre étude compare l'expression de l'espace dynamique dans des récits oraux de locuteurs natifs de l'italien, du français et de l'anglais, dans le but de vérifier dans quelle mesure la typologie de Talmy (1985, 2000) rend compte des usages réels des locuteurs de ces trois langues. Cette recherche se propose aussi d'évaluer la relation entre le langage et la cognition lors de l'expression de l'information spatiale. Les données confirment la présence d'une variation interlinguistique entre le français et l'anglais (angl. *They climb out of the jar* vs. fr. *il s'échappe de son bocal*), mais relèvent aussi une variation intratypologique, notamment la coexistence des deux stratégies encodant le mouvement qui semble être plus répandue en italien qu'en français (it. *Salta via dal barattolo*), malgré leur proximité génétique.

ABSTRACT

Differences in the expression of dynamic space in L1: contrastive study between Italian, French and English

This study compares the expression of dynamic space in the oral productions of Italian, French and English native speakers. Our aim is to test to what extent the Talmyan typology (1985, 2000) accounts for the real usage the speakers of these three languages make. This research is also designed to evaluate the relation between language and cognition during the linguistic productions. Our data confirm the existence of an interlinguistic variation (Eng. *They climb out of the jar* vs. Fr. *il s'échappe de son bocal*), but also show an intratypological variation, that is the coexistence of two strategies of motion encoding which seems to be more common in Italian than in French (It. *salta via dal barattolo*), in spite of their typological proximity.

MOTS-CLEFS : espace, variation intratypologique, perspective spécifique, italien L1, français L1, anglais L1

KEYWORDS: space, intratypological variation, specific perspective, L1 Italian, L1 French, L1 English

1 Introduction

Se situer et se déplacer dans l'espace constitue l'une des activités humaines les plus élémentaires : la référence au temps et à l'espace correspond aux catégories fondamentales de la cognition humaine. Par ailleurs, l'étude de ce domaine est particulièrement intéressante car elle met à jour une forte variabilité typologique entre les langues (Talmy, 1985, 2000). Au-delà des implications typologiques, une telle étude peut avoir une portée qui dépasse le domaine strictement linguistique pour s'inscrire dans un cadre plus vaste, à savoir celui de la relation entre la langue et la cognition.

Depuis quelques années, de nombreuses recherches interlinguistiques, comparant des productions natives dans différentes langues (Berman & Slobin, 1994), ont montré que dès l'enfance, le locuteur prête attention à différents aspects de la réalité, en fonction des catégories lexico-grammaticales que sa L1 met à disposition. La perspective spécifique adoptée par le locuteur semble être ainsi partiellement contrainte par les caractéristiques typologiques de la L1 (cf. l'hypothèse du *Thinking for Speaking*, Slobin, 1996). Notre recherche veut donc apporter une contribution à ce débat tout en tenant compte de la distance typologique entre les langues considérées pour cette étude.

Nous nous proposons d'évaluer la façon dont des locuteurs natifs de l'italien, du français et de l'anglais véhiculent l'information spatiale dans leur langue maternelle (L1), afin de voir si la dichotomie talmyenne est confirmée par les usages des locuteurs de ces trois langues. Plusieurs études ont montré que les locuteurs natifs des langues romanes et des langues germaniques se différencient dans la sélection et segmentation de l'information à verbaliser en raison des moyens lexico-grammaticaux qui sont spécifiques aux langues (Slobin, 2004 ; Benazzo et Dimroth, 2008). À ce propos, selon Talmy, l'italien et le français sont des langues à cadrage verbal (*verb-framed languages*), à savoir qu'elles préfèrent exprimer la Trajectoire (*Path*) dans le verbe principal tandis que la Manière et/ou la Cause sont encodées dans une proposition indépendante, voire subordonnée. En revanche, l'anglais en tant que langue germanique, est une langue à satellites (*satellite-framed language*). Les langues à satellites tendent à lexicaliser les autres composantes du mouvement (Manière et/ou Cause) dans la racine verbale tandis que la Trajectoire est postposée au verbe grâce à des particules dites satellites. Sur la base de ces observations, on présuppose l'existence d'une variation interlinguistique nette entre l'italien et le français d'un côté et l'anglais de l'autre.

Malgré cela, on s'interroge aussi sur la possibilité d'une variation intratypologique, à savoir des langues pouvant appartenir au même groupe (ce qui est le cas de l'italien et du français) sans qu'elles ne caractérisent la composante du déplacement de la même manière, comme cela a été montré par Berthele (2006), Iacobini & Masini (2007), Cini (2008), Iacobini (2009).

2 Les informateurs et les données

Dans cet article nous privilégions des productions orales en temps réel réalisées par des locuteurs natifs d'italien, français et anglais ayant entre 20 et 35 ans et disposant tous d'un niveau d'instruction universitaire. Les productions orales sont censées refléter la conceptualisation en temps réel du locuteur de manière plus fiable que les productions écrites, qui permettent de réfléchir plus longtemps sur l'organisation de l'information et de soigner la formulation de l'expression. Les données des italophones, des francophones et des anglophones ont été enregistrées respectivement en Italie, en France et en Irlande. Nous leur avons demandé d'accomplir deux tâches cognitivement complexes : un récit de fiction produit sur la base du support *Frog, where are you ?*⁶⁰ (Mayer, 1969) et un récit de film suite au visionnage d'un dessin animé muet de la série polonaise *Reksio*⁶¹ (Marzalek, 1967). Les deux supports, riches de mouvements dans l'espace, se prêtent donc à l'analyse du domaine référentiel choisi pour notre recherche. Cependant, il y a un écart pour ce qui est de l'accès aux supports lors de la production : pour le premier support, les informateurs ont pu regarder les images tandis que le second, en tant que récit de film, a demandé au locuteur de mémoriser l'histoire et de présenter les événements mémorisés selon une séquence temporellement organisée. Autrement dit, les informateurs ont dû faire un plus grand effort de mémoire pour raconter le dessin animé.

3 Phénomènes analysés

Dans les paragraphes suivants nous comparons les données recueillies en essayant de répondre aux questions suivantes :

⁶⁰ Il s'agit d'un livre de 24 images sans texte racontant l'histoire d'un enfant et de son chien qui partent à la recherche de leur grenouille.

⁶¹ L'épisode en question présente la journée typique d'un chien et de son maître, dont les activités sont contrariées par des obstacles : au cours d'une partie de patinage la glace se brise et l'enfant tombe à l'eau. Le chien intervient et tout rentre dans l'ordre.

1. Dans quelle mesure la dichotomie proposée par Talmy rend-elle compte des usages réels des locuteurs natifs de l'italien, du français et de l'anglais ?
2. Peut-on parler de variation intratypologique pour deux langues, telles que l'italien et le français, qui semblent encoder l'information spatiale de la même manière ?
3. Dans quelle mesure la L1 contraint-elle l'expression de l'information spatiale chez ses locuteurs ?

Pour ce qui est de l'analyse des données, nous nous sommes focalisés sur l'emploi des verbes de mouvement et des verbes encodant la manière du mouvement au sens large, ce qui constitue certainement la première catégorie syntaxique qui vient à l'esprit lorsqu'on s'interroge plus particulièrement sur l'expression du mouvement.

4 Analyse des données

Dans l'analyse des données on a tout d'abord remarqué un très large emploi des verbes de mouvement chez tous les informateurs interviewés. Le tableau suivant montre les résultats concernant les productions des locuteurs :

Locuteurs	Occurrences
Italophones	127
Francophones	153
Anglophones	75

TABLE 5 – Occurrences verbes de mouvement

D'après le tableau ci-dessus le mouvement semble être une information cruciale à véhiculer pour accomplir la tâche proposée. Toutefois, il faut quand même faire une distinction entre les stratégies linguistiques employées pour ce type d'information : comme on s'y attendait, les anglophones privilégient surtout les verbes avec un satellite qui lui est postposé pour indiquer la trajectoire tandis que les francophones et les italophones montrent une préférence pour les verbes dont la racine encode la direction du mouvement.

Les exemples ci-dessous illustrent ces résultats :

(1) FS-angl-01-03-12

The frog is trying to climb out of the jar (...) the deer is running away from the

rock the deer has pushed him [//]⁶² thrown him over the rock

[trad.fr : la grenouille essaie de s'échapper du bocal (...) le cerf court sur un escarpement le cerf le fait tomber de la falaise]

(2) FS-it-02-04-15

e il ranocchio scappa (...) e: e il cervo distrattamente [= ! rit] lo fa cadere da un #

[trad-fr : e la grenouille s'échappe (...) e: e le cerf distraitemment [= ! rit] le laisse tomber d'un #]

(3) FS-fr-04-06-24

Pendant ce temps-là la grenouille *s'échappe* de son bocal (...) le petit garçon rencontre un cerf qui le prend sur sa tête et l'amène plus loin # ahh le petit garçon *tombe* avec le chien dans une je ne sais quoi [= ! rit]

Les exemples ci-dessus illustrent bien la dichotomie talmyenne et donc la variation interlinguistique qu'on s'attendait à trouver dans les données recueillies.

Toutefois, notre analyse montre aussi qu'une telle distinction se base sur des tendances chez les locuteurs et non pas de tous les usages recensés. Autrement dit, aucune langue du monde ne fait partie ni d'un groupe ou de l'autre de façon absolue : il y a toujours des marges de liberté quand on s'exprime linguistiquement (Slobin, 1996), d'où la possibilité d'employer plusieurs stratégies visant à l'expression de la trajectoire, ce que nous avons remarqué surtout dans les récits produits par les anglophones. Les anglophones utilisent en effet des verbes intégrant la direction du mouvement, même si cela se vérifie très rarement dans nos récits (cf. ex. 4 et 5) :

(4) FS-angl-02-20

Eventually the dog enters his house

[trad.fr : enfin le chien entre dans sa niche]

(5) FS-angl-11-05

The frog has escaped

[trad-fr : la grenouille s'est échappée]

En outre, nos données relèvent aussi une variation intratypologique entre l'italien et le français, notamment la coexistence des deux stratégies encodant le mouvement qui semble être plus fréquente en italien qu'en français, nonobstant leur proximité

⁶² Voici quelques conventions de transcriptions adoptées dans cette étude : «[//]» = reprise avec correction ; «[=! rit]» = rire ; «e:» / «m:» = pause remplie ; # = pause brève.

typologique. Voyons de plus près quelques exemples :

(6) FS-it-01-04

La rana salta via dal barattolo

[trad-fr : la grenouille s'échappe du bocal]

(7) FS-it-02-16

il bambino cade giù cade nell'acqua insieme al suo cane

[trad-fr : l'enfant tombe en bas dans l'eau avec son chien]

(8) FS-it-03-05

Nel frattempo la rana esce fuori dal barattolo

[trad-fr : entre-temps la grenouille sort du bocal]

Les exemples ci-dessus montrent la présence en italien d'une nouvelle classe lexicale qui devient de plus en plus productive (Iacobini & Masini, 2006) : il s'agit de verbes à particule, dénommés verbes syntagmatiques (Simone, 1997). Ces derniers fonctionnent comme un item lexical, qui peut avoir un équivalent synthétique, par exemple *saltare via* = *scappare* (fr. s'échapper), *cadere giù* = *cadere* (fr. tomber). De telles constructions « à satellite » peuvent avoir plusieurs valeurs sémantiques en italien : comme le montre l'exemple (6) la particule adverbiale *via* ajoute une marque de direction à un verbe de mouvement, ou bien comme le soutient aussi Bernini (2005), ce type de construction sert surtout à lexicaliser l'éloignement du point de départ d'un parcours donné dans le déplacement.

La particule adverbiale peut également avoir une valeur pléonastique, qui n'intensifie qu'une information déjà présente dans le verbe, ce qui est le cas de *esce fuori* dans l'exemple (8) : comme le verbe *uscire* lexicalise déjà la trajectoire, la particule *fuori* ne fournit qu'une information redondante.

L'histogramme de la figure 1 montre que dans les productions langagières des francophones la stratégie à satellite n'a jamais été employée.

Pour cette raison, sur la base de nos données, nous soutenons que les verbes syntagmatiques représentent un cas de *variation intratypologique* entre l'italien et le français. Nous attribuons cette variation au fait que les verbes syntagmatiques constituent une classe verbale assez développée surtout dans l'italien parlé et plus un domaine donné est accessible dans une langue donnée plus il devient partie intégrante des habitudes linguistiques des locuteurs. Autrement dit, même si le français et l'italien

sont classifiés comme langues à cadrage verbal, il semble que l'italien détaille beaucoup plus la trajectoire, ce qui confirme l'idée selon laquelle des langues typologiquement proches peuvent avoir un différent degré d'élaboration des composantes sémantiques (Ibarretxe-Antunano & Hijazo-Gascón, 2013).

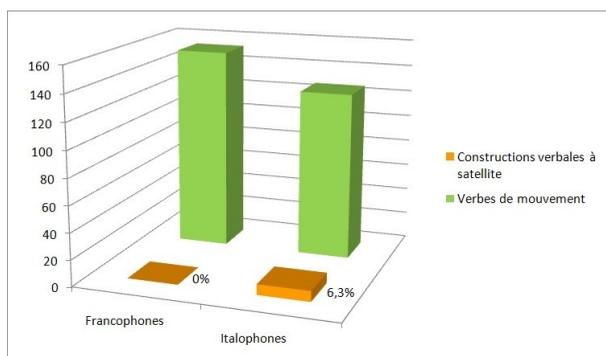


FIGURE 7 – Occurrences constructions verbales à satellite

5 La manière de se déplacer : une autre composante du mouvement

Comme le montre Slobin (2004), la manière du mouvement représente un autre critère de variation interlinguistique dans l'expression du déplacement : la manière et la cause ne doivent pas être considérées comme ayant un statut secondaire par rapport aux autres éléments qui composent un événement de déplacement. Pour cette raison, il propose un *continuum* allant des langues où la saillance de la manière est cruciale, ce qui est le cas des langues germaniques et des langues slaves (*high-manner salient languages*), à celles où elle est faible telles que les langues romanes (*low-manner salient languages*). À ce propos nous avons quantifié l'emploi des verbes de mouvement dans les trois langues considérées (cf. Figure 2).

Comme le montre l'histogramme ci-dessus, l'analyse de nos données confirme que l'anglais est une langue à haute saillance de manière, à savoir la disponibilité en anglais d'une classe verbale décrivant le mode de mouvement et sa facilité d'emploi ont pour effet que les anglophones prêtent beaucoup plus d'attention à la manière et soient ainsi préparés à encoder l'espace dynamique selon ce que la langue les prédispose à faire (cf. ex. 9, 10, 11) :

(9) FS-angl-04-23

The boy jumps on the top of a deer

[trad-fr : le garçon *monte* sur la tête du cerf]

(10) FS-angl-03-07

The frog is skipping out of the jar

[trad-fr : la grenouille *s'enfuit en sautant* du bocal]

(11) RK-angl-07-16

He slides across the ice and climbs the steps

[trad-fr : il *glisse* sur la glace et *monte* les escaliers]

En revanche le français et l'italien se comportent comme des langues à basse saillance de manière : la manière est secondaire par rapport à la composante trajectoire. Grammaticalement, ce caractère secondaire peut s'exprimer par la subordination des segments véhiculant la manière à ceux qui désignent la trajectoire.

Le français et l'italien ont le même comportement par rapport à l'expression de la manière et semblent élaborer cette composante de façon identique. Tous les francophones et les italoophones enregistrés se sont servis des verbes de manière tels que « patiner » (it. *pattinare*), « tomber » (it. *cadere*) et « glisser » (it. *scivolare*) surtout pour une scène de Reksio où, au cours d'une partie de patinage, la glace craque et l'enfant tombe à l'eau. C'est la scène la plus cruciale pour la cohérence de l'histoire :

(12) RK- fr- 05-37

Le maître *glisse* sur le lac glacé se brise et du coup le maître tombe dans l'eau

(13) RK-fr-01-25

La maîtresse récupère ses deux patins et *glisse* sur le lac gelé

(14) RK-it-04-18

Il bambino comincia a pattinare pattina

[trad-fr : le petit garçon commence à patiner *patine*]

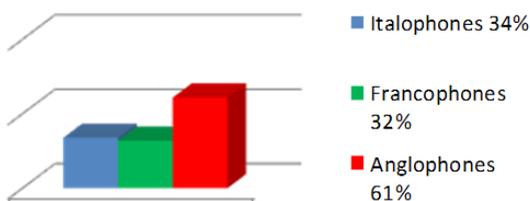


FIGURE 8 – Occurrences verbes de manière de mouvement

Nos informateurs ont utilisé très rarement des verbes de manière dans le récit de *Frog*

tandis qu'une fréquence plus élevée a été attestée dans les récits de film, ce qui montre aussi la pertinence dans l'usage des deux tâches. En d'autres termes, les locuteurs des langues romanes analysées dans notre étude semblent éviter d'utiliser les verbes détaillant la manière du mouvement s'ils n'y sont pas contraints pour la cohérence de l'histoire. Contrairement aux anglophones, ils n'ont pas tendance à détailler la manière du déplacement mais ils semblent prêter beaucoup plus d'attention à la direction. Probablement en termes cognitifs, il serait trop coûteux chez les italo-phones et les franco-phones d'exprimer le mode du mouvement vu que le répertoire de verbes lexicalisant la manière est moins productif par rapport à celui de l'anglais (Slobin, 2004). L'analyse de nos données montre qu'il y a une plus forte influence typologique des structures de langues dans les productions des locuteurs en ce qui concerne l'expression de la manière du déplacement plutôt que dans l'encodage de la trajectoire, même si cette influence n'est pas déterministe.

6 Conclusions

Dans cette étude nous avons analysé la façon dont l'espace dynamique est encodé dans les récits oraux produits par des locuteurs natifs d'italien, français et anglais. Globalement nos résultats confirment qu'il existe une variation interlinguistique entre les langues romanes et les langues germaniques : les anglophones se servent surtout de verbes de manière du mouvement et expriment la direction dans le satellite tandis que les italo-phones et les franco-phones privilégient des verbes encodant la direction du mouvement dans la racine, ce qui confirmerait la dichotomie talmyenne et sa reformulation proposée par Slobin. Toutefois, nous avons aussi relevé la présence d'une variation intratypologique dans les usages des locuteurs italo-phones vs. franco-phones, variation qui tend à nuancer la dichotomie talmyenne : bien que l'italien et le français partagent le même modèle de lexicalisation dans l'expression du mouvement, l'emploi des constructions à satellite (*cadere giù, saltare via*) semble être plus fréquente en italien, qui exploite l'existence d'une classe verbale, représentée par les verbes syntagmatiques. Notre analyse confirme donc que, d'une part, les productions langagières des locuteurs natifs sont fortement contraintes par la façon de penser pour parler acquise lors de l'acquisition de la L1 et que, d'autre part, il faut toujours tenir compte du fait que les langues ne sont pas des systèmes fermés et fixes, au contraire elles évoluent au cours du temps et par conséquent elles peuvent même évoluer d'une typologie à l'autre.

Enfin nos résultats permettent aussi de préciser que, même si très rarement, les italophones et les francophones ont recours aux verbes encodant la manière du mouvement surtout pour les scènes permettant de comprendre la suite des événements dans l'histoire, ce qui est le cas d'une scène de *Reksio* : le fait que le garçon glisse sur le lac fait briser la glace et le fait tomber à l'eau. La cohérence du discours semble être un principe qui pousserait les locuteurs natifs d'italien et de français à exprimer le mode du mouvement.

Références

- BENAZZO S. & C. DIMROTH (2008). « The information structure of learner language ». Eurosla 18 (Aix-en-Provence ; 11-13 septembre).
- BERMAN R.A. & D.A. SLOBIN (1994). *Relating Events in narratives : A Crosslinguistic developmental Study*. Hillsdale, NJ : Lawrence Erlbaum Associates.
- BERNINI G. (2005). « Strategie di lessicalizzazione: tipologia e apprendimento. Il caso dei verbi di moto ». *Deutscher Romanistentag Spracher in und um die Romania* (Saarbrücken ; 25-28 septembre).
- BERTHELE R. (2006). *Ort und Weg. Die Sprachliche Raumreferenz In Varietäten des Deutschen, Rätomanischen und Französischen*. Berlin : Walter de Gruyter.
- CINI M. (2008). « I verbi sintagmatici in italiano: una proposta di descrizione ». In : M. CINI (ed.), *I verbi sintagmatici in italiano e nelle varietà dialettali. Stato dell'arte e prospettive di ricerca. Atti delle giornate di studio* (Torino ; 19-20 febbraio 2007). Frankfurt am Main : Peter Lang, p.41-59.
- HENDRIKS M. & M. WATOREK (2011). The role of conceptual complexity in the acquisition of the spatial domain by L1 and L2 learners of French, English and Polish. In : M. WATOREK, S. BENAZZO & M. HICKMANN (eds) (2011), *Comparative Perspectives on Language Acquisition: A tribute to Clive Perdue*. Dublin : Multilingual Matters, p.400-418.
- IACOBINI C. & F. MASINI (2006). « The emergence of verb-particle constructions in Italian: locative and actional meanings ». In : *Morphology*, n°2, vol.16, p.155-188.
- IACOBINI C. & F. MASINI (2007). Verb-particle constructions and prefixed verbs in Italian: typology, diachrony and semantics. In : G. BOOIJ, B. FRADIN, A. RALLI & S. SCALISE (eds), *On-line Proceedings of the Fifth Mediterranean Morphology Meeting (MMM5)* (Fréjus ; 15-18 septembre 2005) [en ligne]. University of Bologna, p.157-184.

IACOBINI C. (2009). « Les verbes parasynthétiques : de l'expression de l'espace à l'expression de l'action ». In : *Revue de linguistique latine du centre Alfred Ernout*, vol.2, p.35-58.

IBARRETXE-ANTUNANO I. et A. HIJAZO GASCON (2013). « Same family, different paths ». In : J. GOSCHLER & A. STEFANOWITSCH (eds) (2013), *Variation and Change in the Encoding of Motion Events*. Amsterdam : John Benjamins Publishing Company, p.39-54.

SIMONE R. (1997). « Esistono verbi sintagmatici in italiano? ». In : T. DE MAURO & V. LO CASCIO (eds) (1997), *Lessico e grammatica. Teorie linguistiche e applicazioni lessicografiche*. Roma : Bulzoni, p.155-170.

SLOBIN D.I. (1996). « From 'Thought and language' to 'thinking for speaking' ». In : J.J. GUMPERZ & S.C. LEVINSON (eds) (1996), *Rethinking Linguistic Relativity*. Cambridge : Cambridge University Press, p.70-96.

SLOBIN D.I. (2004). « The many ways to search for a frog: linguistic typology and the expression in motion events ». In : S. STRÖMQVIST & L. VERHOEVEN (eds) (2004), *Relating events in narrative: Vol. 2. Typological and contextual perspectives*. Mahwah, NJ : Lawrence Erlbaum Associates, p.219-257.

SLOBIN D.I. (2006). « What makes manner of motion salient? Explorations in linguistic typology, discourse and cognition ». In : M. HICKMANN & S. ROBERT (eds) (2006), *Space in languages: Linguistic systems and cognitive categories*. Amsterdam/Philadelphia : John Benjamins, p.59-81.

TALMY L. (1983). « How language structures space ». In : H. PICK & L. ACREDOLO (eds) (1983), *Spatial Orientation: theory, research and application*. New York : Plenum, p.225-282.

TALMY L. (2000). *Toward a cognitive semantics*. Cambridge MA : MIT Press.

Les *phrasal verbs* dans les productions écrites des apprenants de l'anglais langue seconde

Emilie RIGUEL

ED 514 (EDEAGE) et EA 4398 (PRISMES), Université Paris 3 – Sorbonne Nouvelle
emilieriguel@live.fr

RÉSUMÉ

Véritable fléau de l'apprenant, les *phrasal verbs* représentent une difficulté notable pour les locuteurs non-anglophones. Une étude quantitative et qualitative sur corpus a été menée afin de comparer l'usage des *phrasal verbs* dans les productions d'étudiants natifs et non-natifs, à l'écrit. Les erreurs et contresens de *phrasal verbs* dans les productions écrites des apprenants non-anglophones ont notamment été relevés. Leur analyse confirme que le transfert d'une langue à l'autre est tributaire des contraintes induites par la langue source (L1) sur la langue cible (L2) et de la complexité intrinsèque de la langue cible.

ABSTRACT

Phrasal verbs in English second-language learners' written productions

Considered as the scourge of the learner, phrasal verbs are significantly challenging for non-English speakers. A quantitative and qualitative corpus study was carried out in order to compare the use of phrasal verbs in native and non-native students' written productions. In particular, phrasal verb errors and misinterpretations in non-English speaking learners' written productions were identified. Their analysis confirms that the transfer from one language to another depends on the constraints induced by the source language (L1) on the target language (L2) and the inherent complexity of the target language.

MOTS-CLEFS : *phrasal verbs*, acquisition de la langue seconde, linguistique de corpus, erreurs, enseignement

KEYWORDS: phrasal verbs, second language acquisition, corpus linguistics, errors, teaching

1 Introduction

1.1 Généralités sur les phrasal verbs

Les phrasal verbs en quelques chiffres

Les *phrasal verbs* représentent une catégorie fascinante de la langue anglaise et sont ainsi considérés comme une spécificité de l'anglais (Fraser, 1976 ; Moon, 2005, cité dans Macmillan, 2005). Comme le souligne McArthur (1898 : 38), ils ont toujours représenté une partie dynamique de la langue anglaise⁶³. Les *phrasal verbs* constituent en effet un tiers du vocabulaire verbal anglais (Li, Zhang, Niu, Jiang et Srihari, 2003). Ils sont davantage employés à l'oral qu'à l'écrit. Par ailleurs, il existe environ 3000 *phrasal verbs* en anglais, dont 700 d'usage courant (Bywater, 1969 ; McArthur et Atkins, 1974 ; Cornell, 1985). En plus du grand nombre de *phrasal verbs* existants, de nouveaux sont constamment inventés. Comme le fait remarquer Bolinger (1971 : xi), ils constituent ainsi une classe très productive et correspondent alors à « une explosion de créativité lexicale qui surpasse toute autre chose dans notre langue »⁶⁴. En outre, les *phrasal verbs* sont d'autant plus productifs qu'ils évoluent et s'adaptent à l'air du temps ; les *phrasal verbs* nouvellement créés s'imprégnant ainsi des évolutions de la société (par exemple, *Google up* dérive de *look up*).

Définition et caractéristiques du *phrasal verb*

Il n'existe pas de définition universelle du *phrasal verb*. En effet, comme le soulignent Gardner et Davies (2007 : 341), « les linguistes et les grammairiens luttent avec les nuances des définitions du *phrasal verb* »⁶⁵. Une des raisons justifiant ce manque de consensus (Darwin et Gray, 1999 ; Sawyer, 2000) repose sur le fait que certains linguistes définissent le *phrasal verb* comme l'association d'un verbe et d'une préposition ou d'une particule adverbiale, tandis que d'autres ne considèrent le *phrasal verb* que comme étant un verbe suivi d'une particule adverbiale uniquement. Il faut toutefois souligner que, les *phrasal verbs* ont traditionnellement toujours été compris comme se composant d'un verbe et d'une particule adverbiale. Le verbe support constituant le *phrasal verb* est monosyllabique, d'origine germanique et exprime un mouvement

⁶³ Traduction de la citation originale anglaise : « a vigorous part of English ».

⁶⁴ Traduction de la citation originale anglaise : « an explosion of lexical creativeness that surpasses anything else in our language ».

⁶⁵ Traduction de la citation originale anglaise : « linguists and grammarians struggle with nuances of phrasal verb definitions ».

concret ou abstrait. Il s'agit essentiellement de verbes dits légers car possédant un sémantisme très large et dont la fréquence d'emploi est très importante dans le langage courant. Les principaux candidats aux verbes supports au sein des *phrasal verbs* sont ainsi *bring, come, get, give, go, make, put, set, take*, etc. Quant aux caractéristiques de la particule adverbiale, celle-ci est morphologiquement invariable, véhicule un sémantisme plutôt vague et décrit un mouvement ou un résultat. Il existe un noyau de particules qui se sont révélées plutôt stables au fil du temps et qui figurent comme étant les plus fréquentes dans le langage courant : *up, down, in, out, on, off, back, away* et *over* (Akimoto, 1999 ; Claridge, 2000).

En ce qui concerne les significations des *phrasal verbs*, elles peuvent s'étendre de valeurs directionnelles, ou littérales, ou transparentes (ex. *stand up, take away*) à des valeurs aspectuelles, ou complétives (ex. *burn down, eat up*), jusqu'à des valeurs non-compositionnelles, ou idiomatiques, ou opaques (ex. *face off, figure out*) (Live, 1965 ; Fraser, 1965, 1966 ; Bolinger, 1971 ; Makkai, 1972 ; König, 1973 ; Moon, 1997 ; Celce-Murcia et Larsen-Freeman, 1999). Les classes sémantiques des *phrasal verbs* peuvent être ainsi représentées sur un vaste continuum s'étalant des significations les plus compositionnelles (directionnelles, puis aspectuelles) aux sens non-compositionnels (idiomatiques) (Bolinger, 1971 ; Moon, 1998) (voir Figure 1).

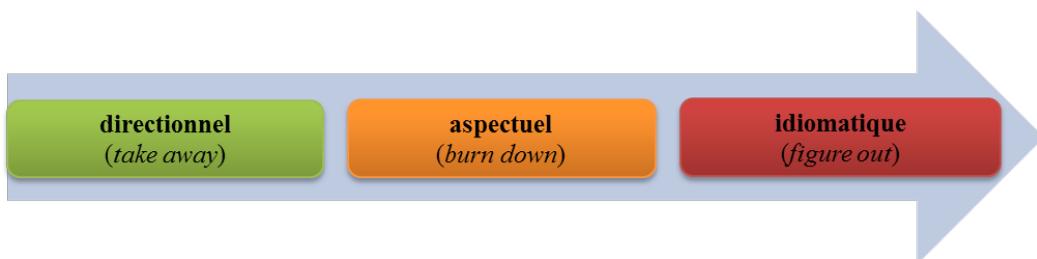


FIGURE 9 – Continuum représentant les classes sémantiques des *phrasal verbs*

1.2 Les *phrasal verbs* et les apprenants non-anglophones

Les *phrasal verbs*, révélateurs du niveau de compétence en anglais

De nombreux linguistes et chercheurs ont reconnu l'importance des expressions constituées de plusieurs mots, et particulièrement des *phrasal verbs*, car elles attestent de la maîtrise de l'anglais et sont le gage de son authenticité (Klein, 1989 ; Folse, 2004 ; Wood, 2004). Les *phrasal verbs* permettent ainsi d'évaluer le niveau de compétence en anglais des apprenants, comme en témoignent des examens tels que le TOEFL ou les tests de Cambridge. Cowie (1993 : 38) les considère métaphoriquement comme une

ortie qui doit être empoignée afin que les étudiants puissent atteindre des compétences proches de celles des locuteurs natifs à l'oral et à l'écrit⁶⁶. Quant à Cullen et Sargeant (1996 : vii), ils expliquent que comprendre et être capable d'utiliser correctement ces constructions à l'oral et à l'écrit est indispensable afin que l'apprenant puisse développer une parfaite maîtrise de la langue⁶⁷.

Les *phrasal verbs*, véritable « fléau de l'apprenant »

Il a beaucoup été question des difficultés posées par les *phrasal verbs* pour les apprenants non-anglophones. En effet, non seulement les *phrasal verbs* peuvent avoir une flexibilité syntaxique réduite, mais ils peuvent également être davantage figurés sémantiquement. Par conséquent, dans certains cas, la signification d'un *phrasal verb* s'avère être difficile à déduire de ses constituants. Par exemple, le *phrasal verb* « *to play something down* » n'a aucunement rapport avec un événement sportif ou de jeu, mais signifie plutôt « *to minimize the importance of something* » (« minimiser l'importance de quelque chose »), comme l'illustre l'exemple suivant, extrait du British National Corpus (BNC) (Davies, 2004-) :

1. *The European Commission sought to play down fears yesterday that new European Community rules limiting imports of cheaper bananas from Latin America would force up prices for consumers.* K59_1005 (BNC)

Par ailleurs, de nombreux *phrasal verbs* sont fortement polysémiques (par exemple, *make up*), ce qui rend la compréhension de leurs différentes significations d'autant plus difficile pour les apprenants. Par conséquent, le recours au contexte s'avère indispensable afin de lever les ambiguïtés. Les énoncés ci-dessous illustrent le cas de *make up*, un verbe à particule à très forte polysémie :

2. *Come on, Annie. Let's make up.* ALJ 2705 (BNC)
3. *Full of cynical amusement, she continued to stare at herself until, inspired, she started to make up her face carefully, emphasising her brown eyes with liner, and smoky eyeshadow, and dusting her high cheekbones with blusher.* HGM 934 (BNC)

⁶⁶ Traduction de la citation originale anglaise : « *a nettle that has to be grasped if students are to achieve native-like proficiency in speech and writing* ».

⁶⁷ Traduction de la citation originale anglaise : « *understanding and being able to use these constructions correctly in spoken and written English is essential if the learner is to develop a complete command of the language* ».

4. *You could make up a whole story. On no real evidence. It would change all sorts of things.* APR 1125 (BNC)
5. *The girl in the chemist's shop said the chemist would make up the prescription the minute he got back from the bank.* H9G 2630 (BNC)
6. *I understand life, and the family ties that make up almost all of it, much less than I ever did.* AEO 2910 (BNC)
7. *Give me time to make up my mind. I promise I'll do everything I can to help the rest of you.* AEB 1717 (BNC)
8. *I'd be ever so appreciative if you could, lass. And as I'm putting you out on your half-day I'll make it up to you, there will be something extra by way of a thank you in your pay packet on Friday.* AN7 304 (BNC)
9. *Since the plant manager was never able to make up a day's loss of output which pulled down his monthly overall efficiency figures on which he was judged, it was never difficult for Clasper to prove his point.* AC2 530 (BNC)
10. *Here's your chance to make up for the naughty things you've done to me.* B0B 2568 (BNC)
11. *You hypocrite, stop making up to my sisters and playing the shining knight, I saw you go to communion today, and it made me sick. How could you? When you don't even... You looked like... I saw you coming back from the communion rails, with your eyes down and your hands folded, as if you weren't putrid inside, but I know. I know.* GUX 107 (BNC)

Tous ces exemples issus du BNC montrent clairement que le contexte contribue pleinement à éliminer les ambiguïtés et à déchiffrer correctement les différentes significations de *make up* : (2) se réconcilier ; (3) (se) maquiller ; (4) inventer une histoire, souvent dans le but de tromper ; (5) faire, préparer ; (6) composer, constituer, former ; (7) se décider, prendre une décision ; (8) revaloir quelque chose à quelqu'un, avec reconnaissance ; (9) compenser, rattraper son retard, rattraper ce qui a été perdu ; (10) se faire pardonner, se rattraper, se racheter ; et (11) faire l'éloge de quelqu'un, se faire bien voir par quelqu'un, faire des courbettes à quelqu'un, notamment dans le but d'en tirer un avantage.

Compte-tenu de leur complexité et de leur nature imprévisible, les expressions

composées de plusieurs mots, et en particulier les *phrasal verbs*, peuvent être difficiles à comprendre et à mémoriser pour les locuteurs non-anglophones dans l'expérience langagière courante (Coady, 1997). Ils sont source de confusion et d'ambiguïté – notamment en matière d'idiomaticité et de polysémie (Cornell, 1985 ; Side, 1990 ; Moon, 1997 ; Celce-Murcia et Larsen-Freeman, 1999 ; Rudzka-Ostyn, 2003) – à un tel point que Sinclair (1996 : 78) va jusqu'à les qualifier de « fléau de l'apprenant »⁶⁸.

2 Objectif

Ce travail a pour but d'analyser quantitativement et qualitativement l'usage des *phrasal verbs* dans les productions écrites d'étudiants natifs et non-natifs. La présente étude se propose ainsi de rendre compte de l'évitement et de la sous-représentation⁶⁹ (Levenston, 1971 : 115) des *phrasal verbs* chez les apprenants non-anglophones. L'objectif de ma recherche consiste également à identifier et interpréter les divers types d'erreurs et de contresens commis par les apprenants. Je montrerai enfin que le « transfert négatif »⁷⁰ de la L1 à la L2 est dû à l'influence de la L1 de l'apprenant et à la complexité inhérente de la L2.

3 Corpus et méthode

Afin de mettre en évidence l'évitement des *phrasal verbs* chez les locuteurs non-anglophones, une étude expérimentale a été menée auprès d'étudiants francophones (132 participants inscrits en première année de Licence LLCE ou LEA anglais à l'Université Paris 3 – La Sorbonne Nouvelle) et anglophones (29 participants issus de la New York University) auxquels j'ai soumis des questionnaires à choix multiples. A partir d'une étude quantitative sur corpus, je montrerai ensuite que les *phrasal verbs* sont sous-utilisés dans les productions écrites des apprenants non-natifs. Tous les *phrasal verbs* ont ainsi été extraits de l'*International Corpus of Learner English, version 2* (ICLEv2) (Granger et al., 2009), composé de dissertations d'étudiants (niveau intermédiaire à avancé) issus de douze langues maternelles différentes (bulgare, tchèque, néerlandais, finnois, français, allemand, italien, norvégien, polonais, russe, espagnol et suédois), et du *Louvain Corpus of Native English Essays* (LOCNESS) (Granger

⁶⁸ Traduction de la citation originale anglaise : « *the scourge of the learner* ».

⁶⁹ Traduction de la citation originale anglaise : « *the under-representation* ».

⁷⁰ Ou « interférence » (Weinrich, 1953 ; Lado, 1957 ; Selinker, 1969 ; Ellis, 1997). Il s'agit des erreurs d'usage de la L1 à la L2.

et al., 1998), le corpus-contrôle complétant ICLEv2. Enfin, à partir du corpus ICLEv2, je relèverai et analyserai qualitativement les erreurs et contresens commis par les apprenants.

4 Les *phrasal verbs* dans les productions écrites d'apprenants non-anglophones

4.1 Étude quantitative

Évitement des *phrasal verbs*

Les apprenants de l'anglais langue seconde ont tendance à adopter une stratégie d'évitement à l'égard des *phrasal verbs*, préférant la plupart du temps utiliser un verbe simple d'origine latine comme équivalent. Cette idée d'évitement a été clairement soulignée par Bywater (1969 : 97) :

Ce qui distingue à l'écrit, et surtout à l'oral, un bon étudiant étranger d'un anglophone est que ce que le locuteur natif écrit ou dit est rempli de ces expressions, tandis que la plupart des étrangers les craignent et les évitent soigneusement ; leur discours manquant alors d'authenticité. Les étudiants étrangers qui aiment être flattés sur leur anglais peuvent le mieux atteindre cet objectif en utilisant correctement ces verbes composés.⁷¹

Cet évitement est confirmé par les résultats des tests à choix multiples que j'ai soumis aux étudiants natifs et non-natifs (voir Table 1).

Groupe	Choix d'un phrasal verb	Choix d'un verbe simple	Bonnes réponses (phrasal verb)	Bonnes réponses (verbe simple)	Mauvaises réponses (réponses pièges)
Natifs	61,28%	38,72%	61,28%	38,72%	0%
Non-natifs	34,91%	65,09%	25,82%	59,64%	14,54%

TABLE 6 – Résultats des tests à choix multiples chez les étudiants natifs et non-natifs (en %)

⁷¹ Traduction de la citation originale anglaise : « *The plain fact is that what distinguishes the writing and, above all, the speech of a good foreign student from those of an Englishman is that what an Englishman writes or says is full of these expressions, whereas most foreigners are frightened of them, carefully avoid them, and sound stilted in consequence. Foreign students who enjoy being flattered on their English can best achieve this by correctly using masses of these compound verbs.* »

Contrairement aux locuteurs natifs, les étudiants non-anglophones « évitent » les *phrasal verbs* et affichent une forte préférence pour les verbes simples d'origine latine. D'origine francophone, les apprenants ont choisi spontanément la structure leur étant familière d'un point de vue morphologique. Ainsi, l'absence d'un trait linguistique similaire entre la L1 et la L2 empêche sa production et son utilisation en L2.

Sous-représentation des *phrasal verbs* dans les productions des apprenants

Après extraction et comptage des *phrasal verbs* de LOCNESS et de tous les sous-corpus d'ICLEv2, j'ai comparé le nombre de *phrasal verbs* produits par les apprenants non-anglophones par rapport aux locuteurs natifs (voir Figure 2). J'ai ensuite calculé le pourcentage d'utilisation des *phrasal verbs* dans les productions des étudiants non-natifs par rapport aux natifs (voir Figure 3). Enfin, les résultats ont été réorganisés et classés par familles de langues afin d'obtenir la production moyenne de *phrasal verbs* par familles de langues (voir Figure 4).

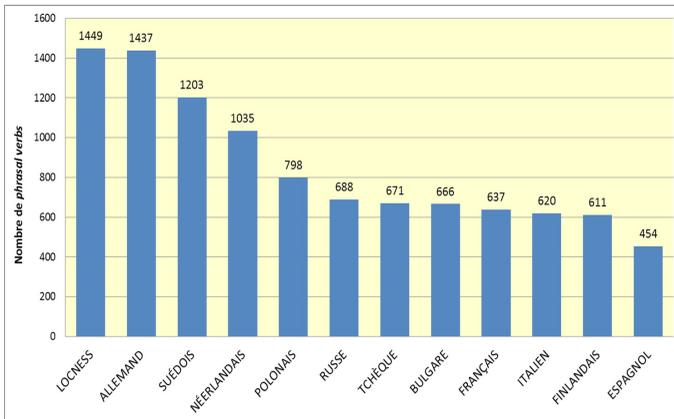


FIGURE 10 – Nombre d'occurrences de *phrasal verbs* dans LOCNESS et ICLEv2

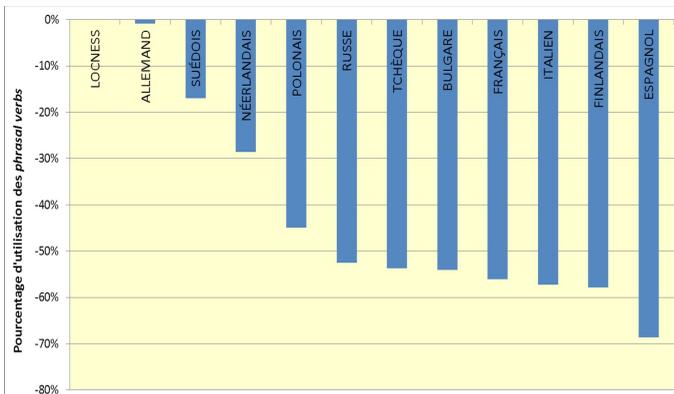


FIGURE 11 – Pourcentage d'utilisation des *phrasal verbs* dans LOCNESS et ICLEv2

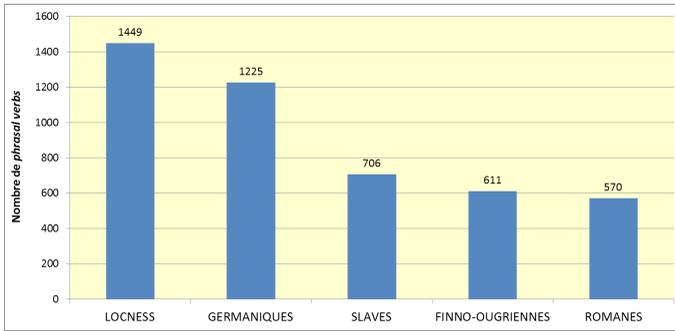


FIGURE 12 – Production moyenne de *phrasal verbs* par familles de langues

Les étudiants dont la L1 appartient à la famille des langues germaniques (allemand, suédois, néerlandais) produisent presque autant de *phrasal verbs* que les locuteurs natifs. Ceci n'est absolument pas surprenant puisque les *phrasal verbs* sont une spécificité des langues germaniques. Il y a donc un trait linguistique similaire entre la L1 et la L2. Lorsque la L1 appartient à la famille des langues slaves (polonais, russe, tchèque, bulgare), la « sous-représentation » des *phrasal verbs* dans les productions des apprenants non-anglophones est moyennement prononcée ; l'aspect étant marqué par la présence de préfixe ou de suffixe dans les langues slaves. Les apprenants dont la L1 fait partie de la famille des langues finno-ougriennes (finnois) « sous-utilisent » considérablement les *phrasal verbs* dans leurs productions. Ceci s'explique par le fait que les verbes à particule anglaise sont exprimés par des verbes simples en finnois. Il existe des verbes composés de deux mots en finnois ; cependant, ils sont peu nombreux et correspondent au registre informel ou familier. Enfin, la « sous-représentation » des *phrasal verbs* est très fortement prononcée dans les productions des apprenants dont la L1 appartient à la famille des langues romanes (français, italien, espagnol). Comme nous l'avons vu dans la section 4.1.1, cette « sous-utilisation » des *phrasal verbs* est due à la différence structurale majeure qu'il existe entre la L1 et la L2.

4.2 Étude qualitative : erreurs commises par les apprenants

Erreurs de style

Les apprenants non-anglophones sont peu conscients des différences entre discours informel et écrit formel. Par conséquent, ils ont tendance à employer des *phrasal verbs* correspondant au registre informel ou familier, voire argotique dans des contextes et/ou écrits formels (et inversement), comme l'illustrent les énoncés suivants, extraits de ICLeV2 :

12. *Try to knock back a few glasses some evening watching the lights in an opposite prefabricated house.* ICLE-CZ-PRAG-0040.3
13. *At that moment I did not know at all what to do, what would be better, how they would react: if I bumped off one of these boys I could have problems as a teacher because physical punishment is not permitted or I could be hit with any punch as well.* ICLE-CZ-PURK-0014.1
14. *So Tony helps the couple Hastings-Neville, or tries to do so, for he mucks things up.* ICLE-SP-UCM-0005.8

Erreurs sémantiques

Les erreurs les plus courantes commises par les apprenants sont les erreurs sémantiques, responsables de contresens majeurs, comme l'attestent les exemples ci-dessous.

Erreurs de verbe. Les apprenants utilisent la bonne particule mais la combinent au mauvais verbe support.

15. *Butter went out⁷² in the course of the week and new one not bought.* ICLE-GE- AUG-0057.1
16. *You can get your energy from peas, beans, cheese, and bread in the same way as from steaks. there are lots of people that feel better since they have put away with⁷³ meat.* ICLE-GE-AUG-0049.2

Erreurs de particule. Les apprenants emploient le bon verbe support mais l'associent à la mauvaise particule.

17. *Every time I read the newspaper I learn more about the terrible actions against people that should be our friends. Their houses are set on fire, they're beaten down⁷⁴ with bottles or baseball rags or they are even killed in fights.* ICLE-GE-AUG-0069.1
18. *It will then be much easier and rapid to find a job that suits you if you can search for it in different countries at the same time, and if you do not have to fill up⁷⁵ many forms and to do a lot of administration work.* ICLE-FR-UCL-0064.3

⁷² *go out* (sortir) ≠ run out (s'épuiser, venir à manquer)

⁷³ *put away with* n'existe pas ; *put away* (mettre de côté, dans le sens économiser) ≠ do away with (abolir, se débarrasser de, faire disparaître)

⁷⁴ *beat down* (enfoncez une porte, faire baisser le prix) ≠ beat up (tabasser)

⁷⁵ *fill up* (remplir un contenant, une salle, etc.) ≠ fill in/out (remplir un formulaire, un questionnaire, écrire)

Erreurs de collocations

Le mot *collocation* vient du latin *collocare* qui signifie « placer ensemble ». Les apprenants ignorent cependant les relations privilégiées qui existent naturellement entre certains mots au sein d'un énoncé et ils ont tendance à associer maladroitement et de façon inappropriée certains *phrasal verbs* avec d'autres mots, comme l'illustrent les énoncés suivants :

19. *In former days girls did not often have the chance to enjoy education or to get a job beside traditional professions like chambermaids or housemaids. Usually they had to marry and to set up⁷⁶ a family.* ICLE-GE-SAL-0013.3
20. *This well established middle-class still didn't give women the same opportunities as men, so in the seventies due to the ferments developed feminism broke out⁷⁷.* ICLE-IT-ROMS-0034.2
21. *Anyway, it is also true that others problems have showed up as consequences of the fights that have been carried out⁷⁸.* ICLE-IT-ROMS-0004.2

Phrasal verbs « inventés »

Pour suppléer un manque, un déficit dans la langue cible, les apprenants éprouvent le besoin de créer de nouveaux verbes à particule qui bien souvent n'existent pas en anglais.

22. *The insiders, that is her family including me of course, know that she has got a fancy about "Freundin" not because of the latest hair-cuts, about "Brigitte" not because of the instructions to fashion your jeans up⁷⁹ by stone-washing and colouring them.* ICLE-

⁷⁶ *set up* s'associe naturellement aux mots relatifs aux affaires (monter une entreprise, s'établir, s'installer) ≠ start a family (fonder une famille)

⁷⁷ *break out* se combine aux mots correspondant à la guerre, au combat, à la violence, à la révolte, à la rébellion, etc. (éclater). Ici, l'étudiant aurait dû employer le verbe emerge (naître, apparaître) car il s'agit de la naissance du mouvement féministe.

⁷⁸ *carry out* s'associe aux mots renvoyant aux expériences, recherches, enquêtes, investigations, etc. (mener, procéder, accomplir, conduire) ≠ put up a fight (mener le combat, lutter, se défendre, résister)

⁷⁹ Bien que les verbes à particule *fashion up* (22) et *mushroom up* (23) n'existent pas en anglais, on ne peut pas parler d'erreurs à strictement parler. En effet, dans ces énoncés, la particule *up* signifie « plus », « davantage » ; elle a une valeur intensificatrice. Ceci montre donc que les apprenants non-anglophones ont bien compris et intégré le sémantisme aspectuel des *phrasal verbs* formés avec *up* et qu'ils appliquent la règle correctement et de façon créative. Le verbe à particule *mushroom up* peut être corrigé par *pop up* (pousser, champignonner) ou par l'expression idiomatique *spring up like mushrooms* (pousser comme des

GE-AUG-0048.3

23. *Tennis courts and clubs mushroomed up¹⁷ all over the Place.* ICLE-GE-AUG-0012.4
24. *It is obvious that we cannot reach absolute equality but we sure can level down⁸⁰ social differences.* ICLE-FIN-JYV-0062.1

Erreurs syntaxiques

Les apprenants non-natifs ignorent les propriétés syntaxiques des verbes à particule et emploient certains *phrasal verbs* intransitifs de manière transitive, et vice versa.

25. *Although parents use light or heavy manners in growing up⁸¹ their children, they are not able to control the future and the idea of having to pay a large sum of money for their children's offences cannot change the situation.* ICLE-IT-TOR-0015.1
26. *Then, the hormones having ceased to be excessively produced, which is only after two or three years, he or she begins to look for another love, splitting up⁸² the relationship.* ICLE-PO-POZ-0031.5

4.3 Influence de la langue maternelle

Le *transfert négatif* ou *interférence* de la L1 à la L2 s'explique principalement par la grande influence de la L1 de l'apprenant mais aussi par la complexité intrinsèque de la L2.

27. *The ideal environment to grow up⁸³ children is the traditional family made up of mother, father and one or more children.* ICLE-IT-TOR-0017.3

champignons).

⁸⁰ Le verbe à particule *level down* n'existe pas dans le dictionnaire. L'apprenant a volontairement associé le verbe support *level* (aplanir, niveler, égaliser) avec la particule *down* (diminution) puisqu'il parle de « réduire les différences sociales » dans son énoncé. Son intention est intéressante et montre clairement qu'il a conscience du sémantisme de *down*. Sa production est cependant incorrecte ; le verbe à particule *level down* n'existant pas. La correction la plus judicieuse est *iron out* (aplanir, réduire, éliminer, faire disparaître) ou *smooth out* (aplanir, faire disparaître).

⁸¹ *grow up* (intransitif) mais *bring up* (transitif)

⁸² *split up* (intransitif) mais *end* (transitif)

⁸³ L'énoncé (27) est issu du corpus d'apprenants italiens. À première vue, nous pourrions penser qu'il s'agit d'une confusion de la part de l'apprenant entre *bring up* (transitif) et *grow up* (intransitif). Cependant, en y regardant de plus près, nous constatons que *grow up* est employé 13 fois de manière transitive et qu'il y a 4 occurrences de *grow a child*. Or, *raise a child* en anglais se traduit par *creocere un bambino* en italien, et *creocere* en italien se traduit par *grow* en anglais, d'où l'emploi incorrect de *grow up* dans l'exemple (27).

28. *Finally as I had decided to leave my bed, to stand up⁸⁴ in order to take off the receiver, I heard the well-known voice of my mother saying: "Good morning, darling".* ICLE-GE-AUG-0024.1
29. *Butter went out²² in the course of the week and new one not bought.* ICLE-GE-AUG-0057.1
30. *How to come up again when you were fallen in the snow without having to make your skis off²².* ICLE-GE-AUG-0057.1

5 Conclusion

La présente étude a permis de rendre compte des difficultés rencontrées par les apprenants non-anglophones à l'égard des *phrasal verbs*. Ce travail a notamment montré que les erreurs d'usage des *phrasal verbs* dans les productions des apprenants sont dues aux contraintes induites par la L1 sur la L2 ainsi qu'à la complexité intrinsèque de la L2. Pour prolonger cette étude, il serait intéressant d'analyser l'emploi des *phrasal verbs* dans les productions d'étudiants non-natifs en considérant d'autres facteurs tels que l'influence du niveau de compétence des apprenants, la durée de la période d'immersion en pays anglophone, la valeur sémantique des *phrasal verbs*, etc.

6 Bibliographie sélective

- BOLINGER D. (1971). *The Phrasal Verb in English*. Cambridge : Harvard University Press.
- CORNELL A. (1985). « Realistic Goals in Teaching and Learning Phrasal Verbs ». In : *IRAL*, n°4, vol.23, p.269-280.
- DAVIES M. (2004). *The British National Corpus (BNC). 100 million words, 1980-1993*. [URL : <http://corpus.byu.edu/bnc/>, consulté le 27 septembre 2014]
- GRANGER S. (ed.) (1998). *Learner English on Computer*. Londres/New York : Addison Wesley Longman.
- GRANGER S., E. DAGNEAUX, F. MEUNIER et M. PAQUOT (eds) (2009). *International Corpus of*

⁸⁴ Les énoncés (28), (29) et (30) sont extraits du corpus d'apprenants allemands. Les *phrasal verbs* anglais *get up*, *run out* et *take off* se traduisent respectivement en allemand par *aufstehen*, *ausgehen* et *abmachen*, lesquels sont formés respectivement de l'association de *auf* + *stehen* (*up* + *stand* en anglais), *aus* + *gehen* (*out* + *go* en anglais) et *ab* + *machen* (*off* + *make* en anglais), d'où les emplois incorrects de *stand up*, *go out* et *make off* respectivement en (28), (29) et (30).

Learner English. Handbook and CD-ROM. Version 2. Louvain-la-Neuve : Presses Universitaires de Louvain.

LEVENSTON E.A. (1971). « Over-indulgence and under-representation – aspects of mother-tongue interference ». In : G. NICKEL (ed.) (1971), *Papers in contrastive linguistics*. New York : Cambridge University Press, p.115-121.

SINCLAIR J. (1996). « The search for units of meaning ». In : *TEXTUS*, n°1, vol.9, p.75-106.

Universaux et acquisition du lexique chez les jeunes enfants de migrants allophones en France

Emmanuelle EGGINTON

Université Paris X Nanterre, 200 avenue de la République, 92 001 Nanterre
manu_egginton@yahoo.fr

RÉSUMÉ

La notion d'universaux de l'acquisition du lexique apparaît dans de nombreuses recherches. Les travaux innéistes, comme ceux de McNamara (1972) et de Gentner (1982) avec l'hypothèse de la partition naturelle, défendent l'hypothèse selon laquelle les noms sont appris avant les prédicats. Les approches fonctionnalistes et constructivistes de Kern (2005) et de Bassano (2007) identifient également plusieurs étapes dans l'acquisition du lexique. Nous confronterons à ces postulats les premiers résultats de notre étude sur le bilinguisme et l'acquisition du français chez les enfants de migrants après les avoir présentés, pour ensuite déterminer comment le lexique des enfants nous renseignerait sur d'éventuels universaux et quels seraient les facteurs de variation.

ABSTRACT

Universals and lexical acquisition among young children of migrant allophone parents in France

The notion of universals in the acquisition of the lexicon surfaces in a number of studies. Research done by tenants of the innate/nativist theory, such as McNamara (1972) and Gentner (1982) who advances the theory of natural partition/division, postulate that nouns are learnt before verbs. The functionalist and constructivist approaches adopted by Kern (2005) and Bassano (2007) also identify several stages involved in the acquisition of the lexicon. We will confront these premises, with the first results of our study on bilingualism and the acquisition of French by children of migrants after having presented them. Then we will determine how the lexicon of children would inform us about putative universals and what factors might be responsible for variation.

MOTS-CLEFS : acquisition du lexique, bilinguisme, universaux, français, parents allophones, input

KEYWORDS: lexical acquisition, bilingualism, universals, French, allophone parents, input

1 Les universaux dans l'acquisition du langage

Les universaux de l'acquisition du langage prêtent à discussion. Nous ne nous intéresserons ici qu'à ce qui concerne l'acquisition du lexique. Dans ce domaine, les travaux innéistes, comme ceux de McNamara (1972, 1984) et de Gentner (1982) avec l'hypothèse de la partition naturelle, démontrent que l'acquisition du lexique suit un schéma fixe, à savoir les noms apparaissent avant les prédicats.

D'autres chercheurs en psycholinguistique qui se penchent sur l'acquisition et le développement du langage mettent en relief des approches fonctionnalistes et constructivistes se proposant, elles aussi, de trouver la part d'universalité dans le développement langagier sans nier une part d'innée. À titre d'exemple, Bassano (2007) nous rappelle, en se référant à l'étude de Caselli, Casadio et Bates (1999), que « Trois étapes ont été ainsi identifiées dans l'évolution de la composition du vocabulaire entre 16 et 30 mois : l'accent se porterait d'abord sur les éléments à valeur référentielle – les noms –, puis sur les éléments à valeur prédicative – les verbes en particulier –, et enfin sur les éléments grammaticaux. Ces tendances de développement ont été retrouvées de façon largement similaire pour bon nombre d'autres langues examinées à partir d'adaptations du CDI [MacArthur-Bates Communicative Development Inventories], en particulier pour l'italien, pour lequel la comparaison avec l'anglais a été explicitement faite ». Pour Kern (2005), il faudrait atteindre un lexique de 300 à 500 mots pour arriver au niveau de la réorganisation du lexique avec une augmentation du nombre de mots grammaticaux et ce dans de nombreuses langues (voir la notion de masse lexicale critique chez Marchman et Bates, 1994).

Cette évolution de la structuration du lexique considérée comme universelle, mais mise en porte-à-faux par des études sur le coréen (Gopnik et Shoi, 1995), peut se schématiser comme le propose (Kail, 2012 : 33) :

routines → référence → prédication → grammaire

Le terme routine fait référence aux expressions associées aux rituels familiaux et sociaux comme : *merci, au revoir, bravo*. Nous aurons l'occasion d'y revenir dans la partie 3.

Nous nous proposons donc d'observer si l'on retrouve ce schéma acquisitionnel chez les enfants de migrants.

2 Présentation des enfants

Nous avons commencé notre observation en novembre 2013 auprès d'enfants âgés de 1 an à 3 ans et 10 mois. Ils vivent dans les environnements linguistiques variés suivants : des familles bilingues français-autre langue, des familles non francophones (monolingues ou multilingues) et des familles monolingues-français (groupe de contrôle). Nous nous proposons de présenter les enfants répartis dans les trois groupes d'âge qu'ils forment.

2.1 Le groupe des plus jeunes

Le groupe des plus jeunes est composé de trois garçons nés entre novembre 2012 et février 2013 : Ad, Am et Da. L'âge moyen du groupe est de 10 mois et 29 jours au début de l'enquête. Nous avons anonymisé les enfants par respect de l'intimité des familles.

Ad est le second enfant d'une famille française. Il a une sœur de 17 ans et entend uniquement du français autour de lui. Am, lui, est le benjamin des trois enfants d'une famille d'origine marocaine. Son père et sa mère, tous deux enfants de migrants, ont gardé la langue de leurs parents : berbère du Rif marocain, pour les échanges du couple et ceux avec les membres plus âgés de la famille, mais ils parlent français avec leurs enfants. Da est le fils d'un couple mixte franco-dominicain. Au quotidien, ses parents parlent français et espagnol. Espérant qu'il devienne bilingue, sa mère lui parle essentiellement en espagnol et son père utilise les deux langues dans ses échanges avec Da.

2.2 Le groupe intermédiaire

Ce groupe est constitué de cinq enfants : Ha, Id, It, Na et Ya. Ils sont nés entre octobre 2011 et janvier 2012. Pour ce groupe, l'âge moyen au début de l'enquête est de 1 an 11 mois et 10 jours.

Ha est le premier enfant d'un couple de Goranis du Kosovo arrivés en France en 2008. Ses parents parlent de nombreuses langues dont le français mais Ha baigne essentiellement dans le goranski. Id est la cadette d'une fratrie de quatre enfants. Ses parents parlent principalement yoruba, anglais et français. Id fréquente la crèche à partir de l'âge de 6 mois, à raison de 9h par jour, cinq jours par semaine. It est la fille unique d'un couple franco-mexicain. Ses parents tentent de favoriser l'espagnol (langue

de la mère) dans les échanges à la maison et d'y exposer au maximum It. Elle fréquente la crèche depuis l'âge d'un an et dix mois, au même rythme qu'Id. Na a une grande sœur de 10 ans. Depuis sa naissance, sa grand-mère fait office de nourrice. Seul le français, parfois teinté de picard, berce ses oreilles. Ya est la petite sœur de Ma qui fait partie du troisième groupe. Ils sont arrivés d'Espagne en France en août 2013 avec leur mère. La langue qu'elle entend majoritairement est l'espagnol.

2.3 Le groupe des plus âgés

Trois enfants, nés entre février et juin 2010, forment le troisième groupe : Al, Ys et Ma. L'âge moyen du groupe en début d'enquête est de 3 ans 6 mois et 17 jours.

Dans la fratrie, Al est née juste avant Id qui fait partie du groupe intermédiaire. Elle a fréquenté la crèche comme sa cadette avant d'intégrer l'école maternelle à 2 ans et demi. Ys est l'aîné de deux enfants. En famille, les échanges se font aussi bien en arabe qu'en français. Ys ne parle quasiment qu'en français mais comprend parfaitement l'arabe. Il est entré à l'école à 3 ans et 2 mois. Ma est, lui aussi, l'aîné de la famille. Il a commencé à fréquenter l'école maternelle en septembre 2013.

3 Observations des premiers résultats et remarques

Notre recherche doctorale porte sur trois facettes de l'acquisition du langage : l'apparition des formes verbales, l'évolution du répertoire lexical des enfants et l'utilisation des *fillers*. Dans cette présente étude, nous ne nous intéressons qu'à la facette lexicale des universaux, même si l'âge avancé des plus âgés pourrait nous renseigner sur bien d'autres facettes du développement langagier. Nous observons dans quelle mesure le lexique du premier répertoire des enfants confirme l'existence des universaux et quels pourraient être les facteurs de variation. Nous utiliserons pour ce travail les premiers résultats exploitables de notre recherche : les questionnaires renseignés par les parents chaque trimestre portant sur les connaissances lexicales de leur enfant dans les langues de leur entourage. Ces questionnaires sont l'adaptation des versions courtes du CDI par Kern *et al.* (12 mois et 18 mois) ainsi qu'un questionnaire largement inspiré du Développement du Langage de Production en Français (DLPF) version 4 élaboré dans le cadre du Programme « École et Sciences cognitives 2001 : Apprentissages des langues : dysfonctionnements et remédiations » (Bassano *et al.*, 2005). Nos observations portent sur les questionnaires correspondant aux âges 1;03, 2;06 et 4;3. Ces trois âges représentent trois étapes de l'acquisition du langage : 1;03

correspond à l'âge du début des productions lexicales des enfants, à 2;06, l'explosion lexicale a eu lieu et il s'agit également de l'âge associé au passage de la phase 1 à la phase 2 de l'échelle de longueur moyenne de l'énoncé (LME) (Brown, 1973) et à 4;03, le lexique et la morphosyntaxe sont considérés comme mis en place. Notre attention a tout d'abord été attirée par le nombre de mots produits par les enfants.

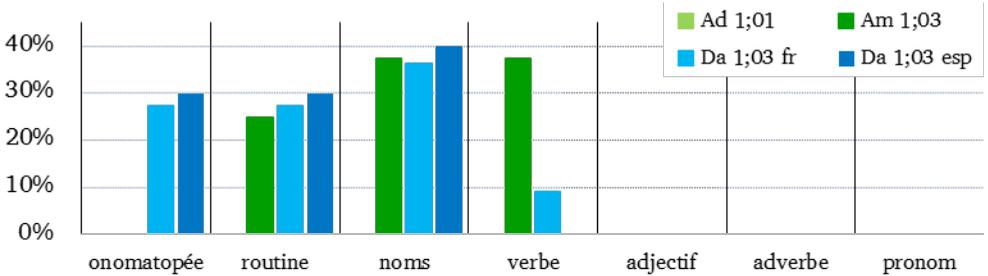


FIGURE 13 – Lexique produit par les enfants de 1;03

Dans la catégorie des plus jeunes, Da, qui est le seul à posséder un répertoire supérieur à 10 mots, rentre dans les statistiques produites par (Bassano, 2007) selon lesquelles « chez les enfants qui ont un vocabulaire entre 10 et 50 mots, la proportion de noms varie de 20% à 60% ». Pour Da, il s'agit de 36,36% de noms dans son lexique, soit 4 noms. Am, dont le répertoire est de 8 mots, a 3 noms à son actif, soit 37,5% de son lexique. Notons ici que les pourcentages pour des lexiques si restreints (une dizaine de mots) ne sont pas très parlants et peuvent leurrer. Nous pouvons également observer que les adverbes, les adjectifs et les pronoms sont absents du lexique d'Am et de celui de Da, ce qui recoupe la théorie de l'ordre acquisitionnel du lexique, selon laquelle les enfants ne développent que dans un deuxième temps ces catégories de mots qui rejoignent les lexiques prédicatifs et grammaticaux selon la classification de Bassano (2007 : 25). Or, il est important de noter que les adverbes (de temps, lieu, quantité, manière) sont, du point de vue linguistique, des mots lexicaux tout comme les noms, les verbes et les adjectifs. Ad, pour sa part, n'avait pas encore produit de mot.

Les enfants du groupe intermédiaire (de 2;06) produisent entre 39 et 583 mots en français avec une moyenne de 257,6 mots (pour 354 en moyenne chez les enfants francophones, Kern, 2013) et une médiane de 223 mots. Ha et Ya, peu exposés au français, sont les deux enfants produisant le moins de mots en français (39 et 93) mais ils sont également ceux qui produisent le moins de mots dans leur langue maternelle (179 et 64). Id et It, pour leur part, sont celles qui produisent le plus de mots en français (350 et 583). Notons que Na, seul enfant issu d'une famille francophone,

produit 223 mots.

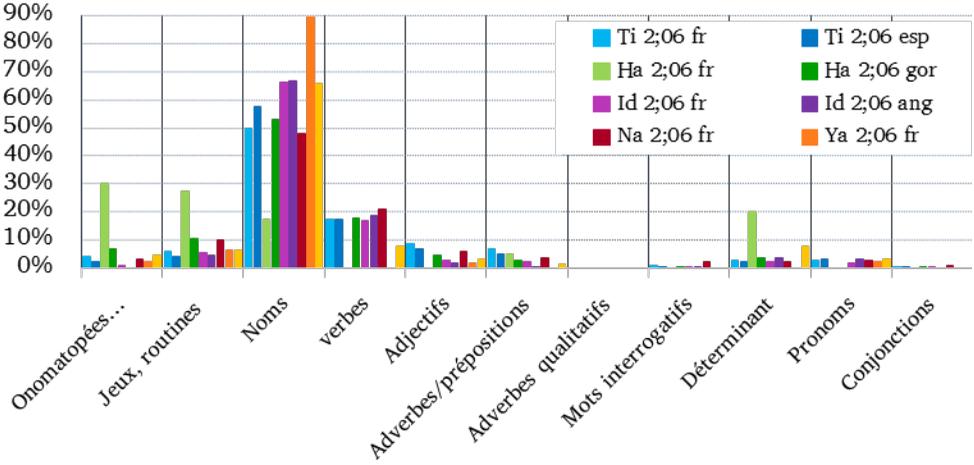


FIGURE 14 – Lexique produit par les enfants de 2;06

Après traitement des données recueillies, nous avons établi des graphiques traduisant la répartition du lexique connu par chaque enfant selon les classes grammaticales et la répartition des noms selon les catégories. Ces graphiques montrent une certaine uniformité. Pour chaque groupe, on observe un pic significatif au niveau des noms : ils représentent de 17,50% à 89,26% du lexique des enfants de 30 et 51 mois, avec une moyenne de 57,23% avec une médiane égale à 56,49%, le premier quartile à 49,09% et le quatrième quartile à 69,97%. Il faut noter ici que le pourcentage de noms de 17,50% est obtenu par Ha qui n'est quasiment pas exposé au français et qui présente une courbe pour le lexique français atypique par rapport aux autres enfants de sa tranche d'âge, avec un fort pourcentage d'onomatopées, interjections et bruits d'animaux (30%), de routines (*super, allô, coucou...*) (27,5%) et de déterminants (20%). Cette courbe semble plus correspondre à la première des trois vagues de structuration du vocabulaire rappelées dans Kail (2012:33) : vague de prévalence des éléments référentiels avec dans un premier temps, une faveur certaine vers les termes de routines.

Le fort pourcentage de déterminants chez Ha s'explique grâce à deux facteurs. Comme nous en avons fait la remarque pour les pourcentages des plus jeunes, le nombre relativement peu élevé de mots produits par Ha donne des valeurs proportionnelles élevées pour des quantités relativement faibles. Ha produit 7 déterminants alors qu'It en produit 16 qui représentent 2,35% de son lexique. Le second facteur est que Ha compte jusqu'à 10 en français, en goranski et en anglais et les chiffres font partie de la

classe « déterminants » dans nos données. De même, Ya a une proportion importante de noms dans son lexique. Elle aussi est peu exposée au français mais plus que Ha : les différents membres de la famille parlent français et sa mère qui vient de terminer une formation en français, exerce beaucoup sa langue et tente d'initier Ya au français en vue de sa rentrée prochaine à l'école. Malgré ce pourcentage élevé, la courbe de son lexique français reste proche de celle des autres. Ces courbes nous semblent corroborer les universaux relatifs à l'acquisition du lexique avec quelques variations personnelles. La courbe de Ha semble appuyer le postulat selon lequel l'acquisition d'une langue est conditionnée par l'exposition de l'enfant à cette langue surtout que la courbe correspondant à son lexique goranski est similaire à celles des autres enfants intermédiaires. Il sera intéressant de voir l'évolution de celle-ci avec l'entrée à la maternelle de Ha d'ici peu. Les courbes des lexiques en langues autres que le français chez les enfants du groupe intermédiaire sont également conformes aux universaux que nous avons soulignés en partie 1.

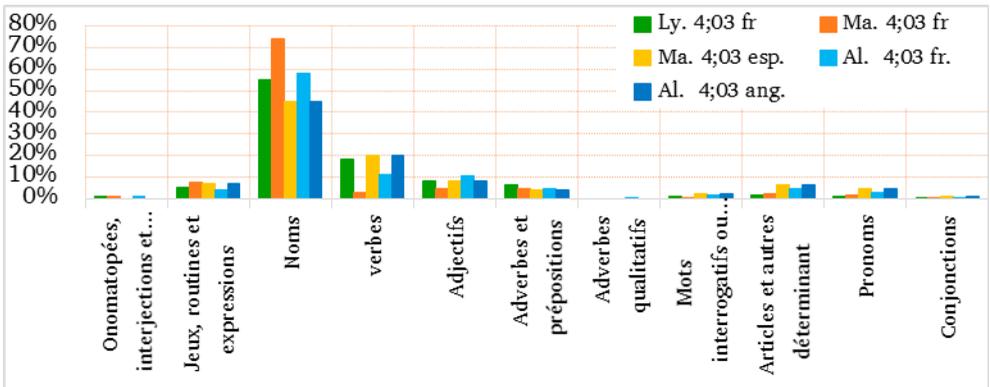


FIGURE 15 – Lexique produit par les enfants de 4;03

Chez les plus âgés, le nombre de mots dans les lexiques français est compris entre 510 et 824, la moyenne est de 678,3 mots et la médiane est de 701 mots. Nous observons, conforme à la progression du développement lexical prévue dans la partie 1, que la proportion de noms diminue (56,49% en moyenne) à la faveur des prédicats (24,27% en moyenne) et des mots grammaticaux (13,25% en moyenne) surtout chez Ys et Al aussi bien en français que dans l'autre langue. Nous pouvons noter que si ces chiffres vont dans le sens de l'évolution de la structuration du lexique, ils n'égalent pas les données moyennes fournies par Bassano (2005) pour les enfants de 30 mois. Il faut néanmoins remarquer que ce phénomène n'est pas flagrant chez Ma. Le taux de noms dans son lexique français est élevé (73,73%) alors que les prédicats représentent 7,65%

et les mots grammaticaux 9,61%. Ces rapports se retrouvent également dans son lexique espagnol. Ceci peut probablement s'expliquer par le développement tardif de sa production langagière. En effet, Ma a très longtemps communiqué en n'utilisant que quelques mots : à 4 ans, le lexique produit est de 99 mots en français comparé à 510 à 4;03 et 78 mots en espagnol (680 à 4;03).

Nous pouvons également pointer le fait que Ha et Ma ont des lexiques français nettement moins nourris que leur lexique goranski et espagnol, respectivement. À l'inverse, pour les autres enfants chez qui le plurilinguisme est présent à la maison et le français prévalant dans leur environnement social (crèche, école), le lexique français est plus développé que celui des autres langues. Ces observations alimentent la thèse selon laquelle l'enfant acquiert les langues de son environnement immédiat quel que soit le pays où il vit.

4 Conclusions

Les données observées sont en adéquation avec les universaux de l'acquisition du lexique que nous avons soulignés : un développement lexical suivant un schéma précis avec, dans un premier temps, une forte proportion de routines et de noms qui, progressivement, cédera du terrain, tout d'abord aux prédicats puis aux mots grammaticaux (Kail, 2012). Néanmoins, nous pouvons noter que les variations individuelles (caractère, sociabilité de l'enfant, etc.) et le temps d'exposition à la langue (l'input) jouent un rôle non négligeable dans l'acquisition du lexique.

Contrairement aux conclusions des observations de Gopnik et Shoi (1995) selon lesquelles l'acquisition naturelle du coréen ne répond pas à l'hypothèse du schéma acquisitionnel, nous pouvons avancer que dans l'état actuel de notre recherche, l'ensemble des langues berçant les enfants de notre étude semblent l'appuyer.

Il ne semble pas y avoir de différence notable dans le développement du lexique entre monolingues et bilingues quant au volume lexical et l'ordre d'apparition dans les différentes catégories lexicales. Notre étude étant longitudinale, nous attendons d'avoir réuni davantage de données pour évaluer cette hypothèse.

Références

- BASSANO D., F. LABRELL, C. CHAMPAUD, F. LEMÉTAYER et P. BONNET (2005). « Le DLPF : un nouvel outil pour l'évaluation du développement du langage de production en français. ». In : *Enfance*, n°2, vol.57, p.171-208.
- BASSANO D. (2005). « Production naturelle précoce et acquisition du langage », *Lidil*, n°31 [en ligne]. [URL : <http://lidil.revues.org/136>, consulté le 26/07/2014].
- BASSANO D. (2007). « Emergence et développement du langage : enjeux et approches fonctionnalistes ». In : E. DEMONT et M.N. METZ-LUTZ (eds) (2007). *L'acquisition du langage et ses troubles*, Marseille : SOLAL Éditeurs, Collection Psychologie, p.13-46.
- BROWN R. (1973). *A first language : the early stages*. Cambridge, Mass. : Harvard University Press.
- CASELLI M.C., P. CASADIO et E. BATES (1999). « A comparison of the transition from first words to grammar in English and Italian ». In : *Journal of Child Language*, n°1, vol.26, p.69-111.
- GENTNER D. (1982). « Why nouns are learned before before verbs: Linguistic relativity versus natural partitioning ». In : S. KUCZAJ (ed.) (1982), *Language development, vol.2: Language, thought and culture*. Hillsdale, NJ : Lawrence Erlbaum Associates, p.301-334.
- GOPNIK A. et S. CHOI (1995). Names, relational words, and cognitive development in English and Korean speakers : nouns aren't always learned before verbs. In : TOMASELLO M. et W. MERRIMAN (eds) (1995), *Beyond names for things : Young children's acquisition of verbs*. New Jersey : Erlbaum.
- HOMBERT J.M. (ed.) (2005). *Aux origines du langage et des langues*. Paris : Fayard.
- KAIL M. (2012). *L'acquisition du langage*. Paris : Presses Universitaires de France.
- KERN S. (2013). *De l'universel au spécifique dans l'acquisition de la langue maternelle : Une approche émergentiste de la diversité*. Mémoire présenté pour l'obtention de la HDR. Université Lyon 2. [URL : http://www.researchgate.net/publication/261913159_Kern_HDR__vol1, consulté le 15/08/2014].
- KERN S. (2005). De l'universalité et des spécificités du développement langagier précoce. In : J.M. HOMBERT (ed.) (2005), *Aux origines du langage et des langues*. Paris : Fayard.
- MACNAMARA J. (1972). « Cognitive basis of language learning in infants ». In : *Psychological Review*, vol.79, p.1-13.

MACNAMARA J. (1984). *Names for Things: A Study in Human Learning*. MIT Press.

MARCHMAN V. et E. Bates (1994). « Continuity in lexical and morphological development: a test of the critical mass hypothesis ». In : *Journal of Child Language*, vol.21, p.339-366.

Apprentissage du français L3 : spécificités temporelles de la langue française du point de vue d'apprenants japonais en milieu hétéroglotte

Najima KERROU

Laboratoire Savoirs Textes Langage, Université Lille 3

najima.kerrou@univ-lille3.fr

RÉSUMÉ

Cet article présente les premiers résultats d'une étude sur l'acquisition du français L3 chez des étudiants (japonais L1, anglais L2), en milieu guidé et contexte hétéroglotte, en se focalisant sur leurs productions langagières (corpus oral constitué de récits de film + entretiens, recueilli selon la méthode longitudinale, pendant une année à 6 mois d'intervalle). L'étude se concentre sur l'expression de la temporalité. Les récits en français montrent l'émergence d'une morphologie verbale temporelle et aspectuelle et la variabilité de l'impact de la L1 et la L2 sur le français L3. Les entretiens d'explicitation mettent en évidence les différences individuelles dans le rôle joué par la L1 et la L2 et explicitent la perception personnelle des spécificités temporelles françaises par les étudiants.

ABSTRACT

Learning French as a Third Language : Temporal specificity of French for Japanese Learners in guided acquisition

This paper aims to present the first results of a study on the Acquisition of French as a Third Language for instructed informants (L1 Japanese, L2 English) in Japan. We analyze their language productions (oral corpus : short-film retelling + interviews, longitudinal data collected during 1 year, 6 months apart). This study focuses on the Expression of Temporality. The narrative discourse in French shows the emergence of a temporal-aspectual verbal morphology and the varying impact of L1 and L2 on L3 French. The explicitation interviews highlight individual differences in the role played by L1 and L2 and the personal perception of French temporal specificities for students.

MOTS-CLEFS : production orale, japonais L1, français L3, temporalité, discours narratif, entretien d'explicitation

KEYWORDS: oral production, L1 Japanese, L3 French, temporality, narrative discourse, explicitation interview

1 Introduction

Cette étude porte sur le processus d'acquisition d'une L3 (le français) lorsque celle-ci est particulièrement éloignée de la L1 (le japonais) d'un point de vue typologique. Nous nous plaçons dans le courant fonctionnaliste. Les locuteurs ont appris cette langue 3 en milieu guidé, dans le milieu universitaire en l'occurrence. Il est question de l'expression de la temporalité. Nos sujets d'observation sont des étudiants japonais en cursus universitaire de français au Japon, que nous avons suivis durant une année et observés à différents stades de leur acquisition du français. Nous nous intéressons à leurs productions orales spontanées (récits de film) et à leur ressenti sur cette tâche narrative (à travers les entretiens d'explicitation). Nous étudions particulièrement l'expression de la temporalité dans les récits en français à travers l'émergence de la morphologie temporelle et aspectuelle et l'impact de la L1 japonais et de la L2 anglais sur la production en français L3, ainsi que la manière dont les étudiants ont accompli leur tâche langagière et leur perception personnelle des spécificités temporelles d'une langue comme le français telles qu'elles sont exprimées dans les entretiens d'explicitation en japonais.

Dans cette étude longitudinale, nous nous posons donc les questions suivantes : comment se développe l'expression des références temporelles en français L3 chez des apprenants japonais en milieu hétéroglotte guidé ? Quels sont les rôles joués par la L1 japonais et la L2 anglais dans la production orale spontanée en français L3 ? Comment les apprenants ont-ils accompli cette tâche langagière et quel est leur ressenti par rapport à cette production orale et à l'expression de la temporalité en langue française ?

Nous avons mené une enquête linguistique et culturelle d'une année au Japon pour obtenir une réponse à ces questions. Les sujets observés au cours d'une année universitaire sont des étudiants de l'université du Tohoku, en cursus de Littérature et Linguistique Françaises. Nous avons observé les étudiants de Licence 2 à Licence 4 et, dans cet article, nous traitons des résultats pour les étudiants de Licence 4, soit le niveau le plus avancé, nos données étant encore en cours d'analyse pour les autres niveaux. Ces étudiants ont reçu un enseignement très formel de la grammaire du français et sont peu exposés au français oral. Nous leur avons fait visionner un dessin animé muet (« Reksio », projet APN), support souvent utilisé dans les études en Acquisition de L2, qui raconte l'histoire d'un petit chien et de son maître qui vont patiner sur un lac gelé. Les apprenants devaient ensuite raconter spontanément cette

histoire en français. Puis nous avons mené un entretien d'explicitation (Vermersch 1994) en langue japonaise pour obtenir leurs réflexions métalinguistiques sur la tâche accomplie. Précisons que nous ne sommes pas de langue maternelle japonaise, mais nous tenions à ce que les entretiens se déroulassent dans la langue natale des apprenants afin qu'aucune barrière linguistique ne gênât l'expression immédiate du ressenti. Puis nous leur avons fait faire une série d'exercices de grammaire sur les temps linguistiques afin de pouvoir mesurer les différences entre connaissances procédurales et connaissances déclaratives. Nous avons ainsi recueilli un corpus de récits + entretiens (+ exercices) selon la méthode longitudinale, 3 recueils de données en début, milieu et fin d'année universitaire avec 6 mois d'intervalle entre chaque recueil, et 4-5 sujets par année universitaire (Licence 2, Licence 3, Licence 4). Nous avons également obtenu des récits du film Reksio produits par des locuteurs natifs en français L1 et en japonais L1, qui vont nous servir de groupes de contrôle. Voici le corpus final recueilli :

	Récits en fr	Entretiens en jap	Exercices en fr	Récits en fr LM	Récits en jap LM
Avril 2012	14	14	14		
Sept 2012	14	14	14		
Fév 2013	14	14	14		
Total	42	42	42	8	8

TABLE 7 – Corpus de données recueilli en un an au Japon

Dénomination	Âge	Année d'études universitaires	Nombre d'années d'anglais	Nombre d'années de français	Total d'heures en français
H	21	4e	8	3	300
M	21	4e	8	3	300
R	22	4e	9	3	300
Y	22	4e	8	3	300

TABLE 8 – Présentation des apprenants de L4 lors du 1er recueil de données

Les étudiants de Licence 4 dont nous analysons ici les productions sont au nombre de 4 et n'ont jamais séjourné en France. Ils suivaient majoritairement des cours sur la langue française donnés en japonais (grammaire/traduction, textes littéraires, etc.) et un cours en français sur la littérature française assuré par un professeur français. Au moment du premier recueil de données en avril, ils avaient 3 ans d'étude de la langue française à leur actif. Notons que, lors de la Licence 1 (qui n'est pas une Licence de français car la

spécialisation commence en Licence 2), les étudiants ont le choix d'une langue vivante nouvelle (en plus de l'anglais) et tous nos sujets ont choisi le français comme LV2.

2 La temporalité dans les récits

2.1 Séquence acquisitionnelle temporelle

Afin de communiquer efficacement, un locuteur se doit de connaître la manière de faire référence au temps. D'ailleurs, dans l'absolu, tout énoncé contient une référence à un intervalle temporel et spatial ainsi qu'à un procès/situation. Prenons l'exemple suivant : un enfant français est chez lui et, soudainement, il se précipite à la fenêtre et dit à sa mère : *il neige !* Si, au même moment, nous sommes dans une classe de français langue étrangère, nous invitons nos apprenants débutants (disons de niveau A1.1 du CECR) à regarder par la fenêtre et leur demandons de décrire ce qui se passe, il y a de grandes chances qu'ils répondent : *maintenant neiger* (véridique, nous en avons fait l'expérience). Cet exemple nous permet de souligner l'existence de moyens explicites et implicites, lexicaux et grammaticaux pour exprimer la référence temporelle. En effet, dans l'énoncé *il neige*, produit en langue maternelle, le sens de *maintenant* est encodé grammaticalement par la morphologie verbale, ici le présent de l'indicatif, alors que dans l'énoncé *maintenant neiger*, produit en langue étrangère, c'est l'adverbe *maintenant* qui encode de manière lexicale la référence temporelle. Nous pouvons donc comprendre à quel point l'acquisition des moyens de la langue-cible pour exprimer ces références temporelles peut être un problème majeur pour les apprenants d'une langue étrangère car les moyens pour exprimer ces références diffèrent bien plus qu'ils ne se ressemblent (Perdue, 1995).

Cependant, les langues sur lesquelles se base notre recherche (japonais, français, anglais) permettent bien d'exprimer les deux spécifications temporelles les plus importantes des événements, « événements » pris au sens large du terme : la temporalité et l'aspect. La référence temporelle – ou temporalité *stricto sensu* – désigne la localisation des événements par rapport à un certain temps de référence. Pour le dire autrement, la temporalité décrit la manière dont le temps d'une situation est lié à celui d'une autre situation (Comrie, 1976). Il s'agit donc de la relation entre un événement et un point de référence donné – ce point de référence pouvant être donné soit par le contexte, soit par les connaissances du monde partagées par le locuteur et l'interlocuteur. L'aspect grammatical, à ne pas confondre avec l'aspect lexical, renvoie

quant à lui aux différentes perspectives que le locuteur peut adopter pour décrire le déroulement d'un événement : il permet de le présenter comme étant « en cours de déroulement », « accompli », « habituel », etc. Selon les langues en présence, le rôle de l'aspect varie : en japonais l'aspect est encodé grammaticalement au présent et au passé, à travers la distinction accompli/inaccompli. C'est également le cas pour l'anglais. Par contre, en français, l'expression de l'aspect n'est obligatoire qu'au passé, et elle est grammaticalisée à travers l'opposition passé composé/imparfait à l'oral et passé simple/imparfait pour l'écrit.

Nous partons du présupposé que la séquence acquisitionnelle temporelle de nos apprenants japonophones sera similaire à celle observée chez les apprenants en milieu non guidé du projet ESF (*Second Language Acquisition by Adults Immigrants*, étude interlinguistique et longitudinale sur une quarantaine d'adultes migrants, qui ont acquis de manière non guidée une L2) ainsi qu'à celle relevée chez des apprenants japonais et chinois en milieu guidé (Trévisiol, 2003 pour les japonophones et Sun, 2003 pour les sinophones). Dans ces études, nous avons remarqué une progression en trois étapes : dans les variétés initiales (*variété prébasique* : énoncés essentiellement nominaux reliés par le principe (topic) focus ; et *variété de base* : énoncés organisés autour d'un verbe mais pas de morphologie flexionnelle ni de subordination, cf. Klein & Perdue, 1997), la temporalité est exprimée essentiellement par des moyens pragmatico-lexicaux (principes d'organisation discursive, adverbiaux, etc.) alors que les moyens morphologiques (flexion verbale) n'apparaissent que tardivement (*variété postbasique* : énoncés organisés autour d'éléments verbaux fléchis).

2.2 Analyse des données en français L3

Analysons maintenant l'émergence de la morphologie temporelle et aspectuelle dans les récits des apprenants en suivant une *approche générale du groupe*.

Lors du 1er recueil en avril 2012, nous constatons une certaine sensibilité des apprenants dans la recherche des formes temporelles du français : sur 44 énoncés, nous recensons 43 formes verbales ; il y a très peu d'infinitifs, mais nous trouvons de la variété dans les formes fléchies, *V-0* (ex : il /glis/), *V-é* (ex : il /EdE/), *V-i* (ex : il /semi/). C'est surtout la forme de base *V-0* qui est utilisée. Cependant, aucune fonction n'est attribuée à cette variation à ce stade, il s'agit de la variété de base. Notons que nous ne nous trouvons pas de récit relevant du stade pré-basique car nos apprenants ont tous environ 3 ans d'étude du français. De même, nous relevons peu d'adverbiaux temporels

et ceux-ci apparaissent exclusivement en japonais ; un usage fréquent de la conjonction de coordination « et » entre les énoncés de la trame montre également que la succession des événements est plus discursive que temporelle.

Lors du 2e recueil en septembre 2012, nous remarquons toujours une grande attention apportée à la flexion verbale : sur 62 énoncés, nous recensons 62 formes verbales ; la majorité des formes sont des formes *V-0* (donc du présent), mais nous trouvons une plus grande variété dans les formes fléchies avec l'utilisation de formes composées avec un auxiliaire *aux + V-0*, *aux + pc*. Nous constatons ici l'apparition de récits racontés au passé. En outre, les moyens adverbiaux utilisés augmentent et sont presque seulement énoncés en français ; « et » est toujours fréquemment utilisé entre les énoncés de la trame mais certaines occurrences apparaissent clairement comme des équivalents de « ensuite » et la relation « après » est davantage explicitée (*après, ensuite, enfin*).

Lors du 3e recueil en février 2013, nous retrouvons un grand souci de la flexion verbale chez les apprenants : sur 68 énoncés, nous recensons 67 formes verbales ; nous remarquons une variété maximale dans les formes fléchies avec une utilisation massive de formes composées avec un auxiliaire, certains récits relèvent de la variété post-basique. Tous les apprenants racontent leur récit au passé. De même, le répertoire d'adverbiaux est davantage diversifié et plus aucun adverbial n'apparaît en japonais. Nous constatons également un développement des moyens pour exprimer le caractère immédiat (*soudain*) et la simultanéité (*à ce moment-là*) et l'utilisation des formes verbales dépasse considérablement celle des adverbiaux.

Comme pour les apprenants en milieu non guidé du projet ESF, chez nos sujets plusieurs formes morphologiques coexistent, certaines sont correctes, d'autres sont idiosyncrasiques. Toutefois, cette variation morphologique est non-fonctionnelle lors de la 1ère série d'entretiens, puis elle émerge timidement de manière fonctionnelle lors de la 2e série d'entretiens pour enfin apparaître clairement lors de la 3e série d'entretiens. L'itinéraire acquisitionnel sur une durée de 1 an est donc le suivant :

V-é/i ⇒ **V-0** (présent) ⇒ **aux + V-0** ⇒ **aux + V-é/i/u** (passé composé)

Cette séquence acquisitionnelle est ainsi similaire à celles observées par Trévisiol (2003) et Sun (2003) chez leurs apprenants japonais et chinois en milieu institutionnel.

Voyons maintenant comment les apprenants ont vécu l'accomplissement de cette tâche narrative orale.

3 Analyse des entretiens d'explicitation

3.1 Difficultés rencontrées lors de la tâche langagière

Lors de chaque recueil de données, nous avons immédiatement fait suivre le récit de film en français d'un entretien d'explicitation mené en japonais. L'entretien d'explicitation est un entretien qui vise une description très fine d'une activité passée, réalisée par une personne engagée dans la réalisation d'une tâche (Vermersch, 1994). Dans le cas de notre recherche, il s'agissait d'obtenir des réflexions métalinguistiques sur la manière dont les sujets observés ont accompli la tâche narrative orale. Nous voulions plus particulièrement obtenir une explicitation des stratégies mises en place pour raconter un récit de film. Nous leur avons donc posé dans l'ordre les questions suivantes en japonais (pour avoir des réponses précises) :

1. *Lorsque vous avez fait votre récit en français, qu'est-ce qui a été le plus difficile ?*
2. *Lorsque vous avez fait votre récit en français, avez-vous pensé à la conjugaison des verbes ?*
3. *Lorsque vous avez fait votre récit en français, avez-vous pensé un peu en anglais ?*
4. *Lorsque vous avez fait votre récit en français, y a-t-il des expressions ou des phrases que vous vouliez dire en français, mais vous n'avez pas réussi à les dire ?*
5. *Pourriez-vous me raconter un apprentissage sur la temporalité qui vous a fortement intéressé ou marqué ?*

Dans le cadre de cet article court, nous allons traiter la 1^{ère} question ici et la 5^e question dans la sous-partie suivante.

Notre première demande concernait les difficultés majeures rencontrées dans l'accomplissement de la tâche de production orale. D'après les réponses obtenues chez nos sujets de Licence 4, nous avons élaboré un classement de ces difficultés :

1. le « vocabulaire » (insuffisance et/ou manque d'activation) : 95 %
2. la syntaxe (difficultés à suivre la syntaxe française et/ou manque de temps) : 81 %
3. la conjugaison des verbes : 75 %
4. la prise de parole à l'oral en continu : 62 %
5. l'absence de support écrit : 45 %

Nous constatons que le point perçu comme le plus difficile est le lexique. Effectivement,

lors des récits, les demandes d'aide lexicale étaient récurrentes, sous la forme d'une élicitation explicite à l'aide d'expressions métalinguistiques « Comment on dit ... ? » ou sous forme d'élicitation implicite avec une intonation montante *hashigo ?* (= échelle). Ces demandes se faisaient à 80% en japonais (car ils savaient que l'enquêtrice était japonophone) et à 20% en anglais. Et suite aux réponses de l'enquêtrice, les apprenants répétaient puis réutilisaient les termes donnés, ou alors ils réagissaient avec une expression du type *ah oui, c'est vrai, ça se dit comme ça*. D'ailleurs, dans les entretiens ils revenaient sur le fait « qu'ils connaissaient les termes en français, *mais ça ne sort pas au moment du récit*. » Le deuxième point le plus difficile, la syntaxe, était très souvent lié chez nos apprenants au fait qu'en japonais, d'une part, le verbe se place tout à la fin de la phrase (SOV) alors qu'en français il est en 2e position (SVO) et d'autre part le sujet n'est pas obligatoirement exprimé dans la phrase japonaise mais il l'est en français. Cette difficulté est directement mise en relation par nos sujets avec le point n°3 du tableau : *non seulement il faut conjuguer le verbe au bon temps et se souvenir de la bonne conjugaison pour chaque verbe, mais il faut aussi accorder chaque verbe avec son sujet*. Aussi certains de nos sujets font-ils des choix délibérés concernant les temps linguistiques, comme l'explique M (sujet 2) dans son entretien : « j'ai pensé que je devais faire le récit au passé, mais c'était trop dur de me souvenir de tous les participes passés du passé composé, donc j'ai parlé au présent, désolée ». Cette remarque est d'autant plus intéressante que dans les récits en français LM, les locuteurs natifs font tous leur récit de film au présent de narration.

3.2 Spécificités de la temporalité en français selon les apprenants

Lors des entretiens d'explicitation, la dernière question « Pourriez-vous me raconter un apprentissage sur la temporalité qui vous a fortement intéressé ou marqué ? » avait pour objectif d'accéder à des dimensions du vécu de l'action qui ne sont pas immédiatement présentes à la conscience de la personne (Vermersch, 1994). Nous désirions ainsi à la fois obtenir quelques réflexions métalinguistiques temporelles supplémentaires sur la tâche accomplie et permettre aux souvenirs d'apprentissage d'émerger, afin d'explicitier les spécificités temporelles du français (ou du japonais.) de chaque apprenant. Selon la loquacité de chaque sujet, les entretiens d'explicitation ont duré entre 7 et 28 min. Notons un détail qui nous semble très intéressant : lors des pré-tests pour les entretiens d'explicitation (avec d'autres étudiants francophones, mais non spécialistes), nous avons remarqué que lorsque nous posons la 5e question, tous les

apprenants manifestaient le besoin de voir, de visualiser le terme « temporalité » en kanji (idéogrammes japonais) pour pouvoir livrer leurs réflexions à ce sujet, allant jusqu'à dessiner le terme sur la paume de leur main. Pour notre recueil de données, nous avons par conséquent une feuille calligraphiée avec le terme « temporalité » en japonais prête pour cet usage, ce qui a été apprécié de nos sujets. De retour à nos analyses, nous avons établi un classement des éléments qui revenaient le plus souvent dans les réponses de nos apprenants de Licence 4 :

1. la « multiplicité » des temps linguistiques en français
2. le principe de la conjugaison en français
3. la distinction entre passé composé et imparfait

L'élément le plus récurrent dans les réponses de nos sujets, à travers les 3 séries d'entretiens sur une année, est également l'élément qui leur semblait le plus spécifique de la langue française : le nombre de temps linguistiques différents que nous possédons en français. Les apprenants citent tous au moins les temps suivants : indicatif présent, passé composé, imparfait, plus-que-parfait, futur simple, futur proche, subjonctif présent, conditionnel présent. De cette multiplicité de temps, les sujets observés en déduisent que *les Français découpent le temps de manière très détaillée, le morcellent en petites parts, alors que pour nous Japonais le temps est une seule entité, un bloc*. Le deuxième point d'étonnement pour nos apprenants est la conjugaison des verbes : le fait que *la terminaison se modifie pour je, tu, il*, etc. et d'autre part, le fait que *la conjugaison change en fonction du temps*, autrement dit, que la valeur temporelle soit grammaticalement encodée par la morphologie verbale en français (alors qu'en japonais, elle est plus souvent encodée de manière lexicale par des adverbiaux). Curieusement, l'anglais n'est jamais évoqué lors de cette remarque alors que les apprenants ont étudié cette langue pendant 8 ans. Le troisième élément qui leur semblait très spécifique du français est la distinction entre le passé composé et l'imparfait. Cette distinction est bien entendu une distinction d'ordre aspectuel. Nous pouvons cependant exprimer un certain étonnement concernant cette remarque dans la mesure où l'aspect est grammaticalement encodé en japonais au présent comme au passé, contrairement au français où seule la distinction passé composé/imparfait est faite (à l'oral).

4 Conclusion

Notre étude sur l'expression de la temporalité dans les récits en français s'est donc ici focalisée sur le niveau le plus avancé de notre corpus de données. Nous avons pu constater l'émergence d'une morphologie temporelle et aspectuelle (avec un itinéraire acquisitionnel présent → passé composé) et nous avons aussi évoqué sans le développer l'impact de la L1 japonais et de la L2 anglais sur la production en français L3. L'analyse des entretiens d'explicitation nous a permis de comprendre comment nos sujets ont accompli leur tâche langagière, les stratégies qu'ils ont développées (explicitation explicite ou implicite, évitement, etc.) ainsi que leur perception personnelle des spécificités temporelles de la langue française (multiplicité de temps linguistiques, conjugaison, etc.) Des analyses plus poussées impliquant les autres niveaux sont en cours et nous permettront d'apporter quelques réponses à la question passionnante de l'expression de la temporalité en langue étrangère.

Références

- BENAZZO S. et M. STARREN (2007). « L'émergence de moyens grammaticaux pour exprimer les relations temporelles en L2 », *Acquisition et Interaction en Langue étrangère*, n° 25, 129-158.
- COMRIE B. (1976). *Aspect*, Cambridge : Cambridge University Press.
- CUQ J-P. (2003). *Dictionnaire de didactique du français langue étrangère et seconde*. Paris : CLE international.
- DHORNE F. (2005). *Aspect et temps en japonais*. Paris : Ophrys.
- GARNIER C. (1985). *La phrase japonaise : structures complexes en japonais moderne*. Paris : Publications Orientalistes de France.
- KATÔ S. (2009). *Le temps et l'espace dans la culture japonaise*. Paris : CNRS Éditions.
- KLEIN W. (1989). *L'acquisition de langue étrangère*. Paris : Armand Colin.
- KLEIN W., R. DIETRICH et C. NOYAU (1995). *The Acquisition of Temporality in Second Language*, Amsterdams : John Benjamins.
- PERDUE C. (1993). *Adult language acquisition : cross linguistics perspectives, volume II The Results*, Cambridge, Cambridge University Press.
- PERDUE C. (1995). *L'acquisition du français et de l'anglais par des adultes, former des énoncés*, Paris, CNRS Édition.

- RIEGEL M., J-C. PELLAT et R. RIOUL (2006). *Grammaire méthodique du français*, Paris : Presses Universitaires de France.
- SHIMAMORI R. (1997). *Grammaire japonaise systématique*. Paris :Maisonneuve.
- SHIMAMORI R. (2001). *Grammaire japonaise systématique, volume 2 : les expressions verbales et les expressions de la politesse*. Paris : Maisonneuve.
- SOGA M. (1983). *Tense and aspect in modern colloquial Japanese*. Vancouver : University of British Columbia Press.
- SUN J. (2003). « Comment se développe la morphologie verbale en français L2 chez les sinophones ? ». In : *Marges linguistiques*, vol.5, p.93-105.
- TREVISIOL P. (2003). *Problème de référence dans la construction du discours par des apprenants japonais du français langue 3*, Thèse de doctorat en sciences du langage, Université Paris VIII.
- TSUJIMURA N. (1996). *An Introduction to Japanese Linguistics*. Cambridge : Blackwell Publishers.
- VERMERSCH P. (1994). *L'entretien d'explicitation en formation continue et initiale*, Paris : ESF.
- VERONIQUE D. (1984). « Apprentissage naturel et apprentissage guidé ». In : *Le Français dans le monde*, vol.185, p.45-52.
- VERONIQUE D. (2009). *L'acquisition de la grammaire du français langue étrangère*. Paris : Didier.