

Colloque international

**Description de l'oral et méthodes d'analyse en
linguistique : perspectives actuelles**

Hommage à Claire Blanche-Benveniste (1935-2010)

Nancy, 20-22 octobre 2021



Textes provisoires. Merci de ne pas citer.

Auteurs/Autrices	Titre	Page
André Virginie, Etienne Carole	Décrire la syntaxe de l'oral dans les interactions du quotidien. Pourquoi et pour quoi ?	5
Avanzi Mathieu	Où en sont les études sur la variation régionale de la grammaire du français ?	12
Berrendonner Alain	Approche pronominale et sémantique référentielle : l'opposition /lui (leur)/ vs /y/	14
Bilger Mireille, Cappeau Paul	La question des unités d'analyse en syntaxe	20
Corminboeuf Gilles	L'adjectif adverbial : délimitation de l'objet, données, héritages épistémologiques	35
Cousinard Clara	Exploitation des descriptions du français parlé : l'exemple de <i>pendant</i>	43
Creissels Denis	À propos de la typologie de la relativisation	52
Da Cunha Yanis, Abeillé Anne	L'alternance actif/passif en français : un modèle statistique sur corpus écrits et oraux	62
Debaisieux Jeanne-Marie	Apport des études sur corpus : le domaine nominal	74
Delahaie Juliette, Canut Emmanuelle, Rouanne Laurence, Sadoun Driss	Les verbes recteurs faibles dans l'expression de l'opinion en français et en espagnol	78
Dister Anne, Branca-Rosoff Sonia	Le masculin et le féminin dans la désignation des référents humains. Etude à partir de l'exploitation de CFPP et CFPB	80
Duvallon Outi, Yatziv-Malibert II-II	Prolongements de la notion de paradigme : outil descriptif et théorique au-delà du français parlé	87
Ferreira Auphémie	Enrichir la description du français parlé grâce à de nouveaux corpus : les constructions [CV. Ø CV.] dans le MPF	100
Fiedler Sophia	Une reconsidération du discours rapporté en langue parlée avec <i>se dire, faire (à quelqu'un) et être là</i>	114
Gadet Françoise, Guerin Emmanuelle	Description du 'français tout court' : quelques considérations sur les pronoms	125
Horlacher Anne-Sylvie, Pekarek Doehler Simona	<i>Si vous avez quelqu'un sous la main</i> : les <i>si</i> -indépendantes en tant que format de requête	134
Johnsen Laure Anne	Corpus oraux et enseignement du français : présentation du projet « Enseignement » de la base OFROM	145
Kunisue Kaoru	<i>Tu comprends / vous comprenez</i> : les fonctions en tant que marqueur discursif dans le corpus du français parlé	149
Lee Chaeyoung, Tutin Agnès	Formules parenthétiques avec <i>dire</i> dans le discours scientifique oral	161
Auteurs/Autrices	Titre	Page

Legallois Dominique	La syntaxe au prisme des propriétés lexicales	173
Manno Chiara	Description et analyses de l'oral : le cas de régulateurs discursifs en français et en japonais	176
Martin Philippe	De la naissance du syntagme au groupe accentuel et à la structure prosodique	179
Mazziotta Nicolas	Syntaxe en n dimensions : choisir et représenter les espaces d'analyse	186
Ploog Katja	Syntaxe de l'oral et activité de construction en temps réel : retour sur les grilles	191
Roubaud Marie-Noëlle, Sabio Frédéric	Les relatives en <i>où</i> en français parlé	202
Schnedecker Catherine	Grammaire de l'oral vs grammaire de l'écrit ? Le cas des pronoms dits "indéfinis"	204
Surcouf Christian, Ausoni Alain	L'apprenant de FLE, bleu de méthylène de l'état de la description du français parlé ?	207
Ukhova Polina	Corpus bilingue multimodal du parler jeune : enjeux méthodologiques et avantages pour l'étude des particularités syntaxico-discursives	213
Vergez-Couret Marianne, Carruthers Janice	Temps verbaux et degrés d'oralité en français et en occitan	226
Verstraete Jean-Christophe	Impersonal constructions in Paman : two puzzles in pronominal morphosyntax	231

Décrire la syntaxe de l'oral dans les interactions du quotidien. Pourquoi et pour quoi ?

Virginie André, ATILF Université de Lorraine et CNRS
Carole Etienne, ICAR ENS Lyon

Les perspectives ouvertes par la description de l'oral, impulsée par Claire Blanche Benveniste et son équipe (notamment Blanche-Benveniste, 1990a ; Blanche-Benveniste, Jeanjean, 1987), ont encouragé les linguistes à élargir leurs études à ce matériel trop longtemps perçu comme chaotique. Néanmoins, au-delà des nettes avancées en linguistique descriptive, les implications et applications des recherches sur le français parlé restent encore sporadiques alors qu'elles sont essentielles dans différentes disciplines. Notre communication cherchera à montrer pourquoi et pour quoi l'exploitation des descriptions de l'oral est aujourd'hui en mesure de répondre à de vrais enjeux sociaux.

Dans le cadre de l'analyse des interactions verbales, les études sur le français parlé prennent une dimension dynamique, collaborative et située (Kerbrat-Orecchioni, 2005 ; Traverso, 2016). Le discours se construit au fur et à mesure des tours de parole des interactants (Sacks *et al.*, 1974) et en fonction des éléments de la situation de communication. Pour ce faire, les ressources lexicales et syntaxiques sont sélectionnées et reconfigurées par les locuteurs, par leurs ajustements réciproques et leur coopération à l'élaboration d'un discours commun. Les analyses de corpus d'interactions permettent de saisir comment les locuteurs utilisent ces ressources pour accomplir des activités et atteindre leurs objectifs dans une variété de contextes et de configurations. Ces analyses intègrent alors d'autres dimensions (phonétique, pragmatique, prosodique, gestuelle, etc.) pour décrire la langue réellement mise en œuvre *in situ*.

Les études sur le français parlé *en interaction* offrent ainsi des perspectives considérables dans différents champs scientifiques. Nous nous intéresserons à ce qu'elles apportent dans le domaine de l'enseignement et de l'apprentissage de la langue. En effet, l'exploitation didactique de ces travaux de recherche propose des méthodologies afin de faciliter l'acquisition de compétences interactionnelles. Ces compétences sont souvent délaissées par les enseignements de français, quels que soient les objectifs d'enseignement ou le niveau des apprenants, parce qu'elles restent les plus difficiles à enseigner ou parce qu'ils estiment que ces compétences sont incompatibles avec un cours en format plénière. Les documents de référence, les manuels de langue notamment, même les plus récents et les plus avancés scientifiquement n'intègrent pas les descriptions de la langue orale leur préférant une approche normative bien rodée mais également bien éloignée de la réalité. Nous pouvons le constater dans le dialogue suivant, extrait du manuel *Le français parlé* paru en 2017 :

Exemple 1 : « Manon arrive à la fête d'anniversaire de Mélanie » (p.213).

Manon :	Salut Hélène, la forme ?
Mélanie :	Oui, et toi ? Vas-y, rentre, fais comme chez toi.
Manon :	Tiens, c'est pour toi ! Bon anniversaire !
Mélanie :	Oh merci mais il fallait pas.
Manon :	Mais si, vingt ans, ça se fête.
Mélanie :	Oh, c'est le sac à main que j'avais vu ! Merci, je pouvais pas rêver mieux comme cadeau !
Manon :	Pour ne rien te cacher, j'ai hésité à le prendre pour moi.

Bien que ce manuel soit spécialisé dans le français parlé et qu'il souhaite montrer aux apprenants « le français oral de la vie quotidienne » (4^{ème} de couverture), ce dialogue est tout de même éloigné d'une interaction authentique. Nous pouvons, par exemple, comparer les tours de parole consacrés à l'accueil (les deux premiers tours) avec une interaction authentique dans une situation similaire :

Exemple 2 : E1 accueille chez elle son amie E2 (FLEURON).

E1: salut

E2: salut

E1: ça va

E2: ça va et toi

E1: bien super vas-y enlève tes chaussures

E2: j'enlève mes chaussures

E2: oh quel temps

E1: tu veux quelque chose à boire

E2: ben écoute un verre d'eau ouais

E1: ok super ben tout est dans le salon si tu veux

E2: où est-ce que je peux mettre ma veste

E1: euh mets la là

E2: ok parfait

(elles se dirigent vers le salon)

Les documents de référence ne décrivent pas non plus les spécificités des interactions verbales alors qu'elles sont au cœur du quotidien, qu'il soit privé ou professionnel, et de la vie sociale, ne préparant pas l'apprenant à les appréhender dans leur diversité. On peut citer les travaux de Giroud et Surcouf (Giroud & Surcouf, 2016) qui ont relevé l'absence de phénomènes fréquents de l'oral comme la chute du e, le passage de *nous* à *on* ou la négation simple dans les enregistrements de plus de dix manuels.

Conçus à partir de corpus d'interactions, les ressources CLAPI-FLE (<http://clapi.icar.cnrs.fr/FLE/>) et FLEURON (<https://fleuron.atilf.fr/>) proposent une exploitation des méthodologies et des résultats des recherches sur le français parlé en interaction (André, 2020 ; Ravazzolo & Etienne, 2019). Nous montrerons de quelles façons ces résultats sont utilisés à des fins didactiques, par des enseignants et des apprenants de français langue étrangère. Nous expliquerons plus précisément comment la description de la syntaxe de l'oral dans des interactions du quotidien permet de comprendre les procédés utilisés en relevant également leurs dimensions prosodiques et multimodales afin de les identifier par un faisceau d'indices.. Pour l'illustrer, nous détaillerons les activités **poser une question** et **raconter une anecdote** d'un point de vue syntaxique lorsque les locuteurs ajustent leurs constructions, s'interinfluencent, collaborent et prennent en compte le contexte, voire se projettent à nouveau dans un événement passé.

1) Poser une question

Les dispositifs Clapi-FLE et FLEURON offrent avec la diversité des situations recueillies et documentées un vivier de questions qui permettront de d'illustrer la variété de constructions lexicales et syntaxiques au-delà des exemples normés des manuels.

Nous proposons quelques exemples, ci-dessous, extraits de ces deux dispositifs.

Des questions « inachevées » d'un point de vue syntaxique mais achevées d'un point de vue pragmatique. Par ailleurs, on relève l'expression "on va dire" qui suit la réponse de l'agent :

Exemple 3 : A la scolarité de l'université, l'étudiant (E) demande à l'agent (A) quand est-ce qu'il sera autorisé à s'inscrire (FLEURON).

E **et la réponse se sera d'ici une semaine à peu près ou**
A deux semaines on va dire
E deux semaines

Des constructions syntaxiques avec des dislocations à gauche et à droite autour d'un segment et de sa reformulation, puis d'une répétition "neuvième↑" pour confirmer la réponse "neuf"^f:

Exemple 3 : Lors d'une discussion privée (CLAPI-FLE).

AGN **toi tu es en↑ tu es en quelle classe toi↓** (0.2)
IRM euh neuf (0.6)
AGN neuvième↑

Des questions dans un temps qui peut déconcerter un apprenant.

- Un passé pour exprimer une action future :

Exemple 4 : Lors d'une discussion au téléphone entre deux amis (CLAPI-FLE).

SAL ouais ça va super dis-moi euh tu fais quoi ce soir↑
ALI ce soir
SAL ouais parce que en fait **on allait sortir** moi Emmanuel et Manon **on allait sortir** pour euh
boire un coup et tout en ville donc je sais pas **si ça te dirait de venir**

- Un passé pour une action immédiate :

Exemple 5 : Une enseignante explique le fonctionnement des associations étudiantes (FLEURON).

E bon voilà on a fait un peu le tour bon même si le tour n'est jamais fait **mais je voulais**
aussi parler des associations citoyennes
A oui
B donc là qui sont qui ne sont pas étudiantes donc euh donc là pareil il y a beaucoup de de de
types d'association

- Un futur pour une action présente :

Exemple 6 : Dans une pharmacie, la pharmacienne (P) demande à la cliente de lui donner sa carte vitale (FLEURON).

P **faudra la carte vitale** s'il vous plait
E oui ((E cherche la carte dans son portefeuille))

Dans la plateforme Clapi-FLE, la fiche explicative "La question d'un point de vue pragmatique : les fonctions, les formulations, les réalisations linguistiques" recense, catégorise et documente les différents types de question à l'oral en proposant différentes attestations de complexité variable.

Ces questions prennent des formes non décrites dans les manuels. Si nous observons les interrogations dans le manuel *Le français parlé*, déjà cité précédemment, nous rencontrons les formes suivantes (2017 : 25-31) :

1. L'interrogation directe

- Inversion du sujet (« rarement utilisé en français parlé », « mais la terminaison [ti] est restée en français populaire pour marquer l'interrogation sans inversion » : *Ça va ti ?*)
- Est-ce que

- C'est... que/qui (en « français familier » ou « très familier » : *C'est où qu'elle est partie ?* ou *C'est qui donc qui frappe à la porte ?*)
- L'intonation

2. L'interrogation indirecte

- Qu'est-ce que (« en français familier » : *J'aimerais bien savoir qu'est-ce qui lui a pris ?*)
- Que (« en français très familier » : *Je te demande pourquoi que tu as fait ça ?*)
- Concordance des temps (pas toujours respectée en français parlé familier : *Il m'a expliqué qu'il a raté son examen.*)

2) Raconter une anecdote

Si le récit écrit est largement abordé dans les manuels, sa réalisation à l'oral reste bien peu expliquée alors qu'il est pourtant très fréquent et présente des caractéristiques spécifiques.

Les exemples suivants montrent notamment que les récits sont élaborés avec des marqueurs, s'appuient sur la multimodalité et font appel à des constructions syntaxiques particulières éloignées de l'écrit. De plus, les récits se co-construisent en interaction, les locuteurs ajustent leurs interventions au fur et à mesure des réactions de leurs interlocuteurs. Ils peuvent également recourir aux connaissances partagées en introduisant des parenthèses ou des incises et sont souvent suivis ou précédés d'évaluations. On trouve tous ces éléments dans cet extrait d'environ 1 minute 30.

Exemple 7 : Extrait d'un récit d'une rencontre amoureuse. ELI raconte la rencontre amoureuse de sa sœur à BEA. MAR était déjà informée de cette rencontre (CLAPI-FLE).

ELI **al|ors tu sais**
MAR [(en riant)] un truc de dingue ((rires))
BEA un truc énorme
ELI ben ouais et tout l' monde est d'accord↑ pour dire qu' c'est énorme
BEA oh
ELI **bon alors en fait**↑ (0.5) ben dis donc j' vais monopoliser la conversation ((rires)) .h alors en fait **le truc c'est que**
MAR c'est ta sœur qui monopolise tout ((rires))
BEA ((en riant)) ouais hein↑ ((rires)) (0.7)
ELI **le truc c'est que en fait euh** (0.7) donc c't été elle est allée travailler↓↓ ben normal
BEA hm hm
ELI pour la saison et elle est allée (0.7) aux impôts↑
BEA ouais↑
ELI **j' t'ai dit elle a été prise aux impôts** (elle a quand même) une chance elle euh j' te dis
BEA hm hm
ELI un truc international quand même
MAR ((rires))
ELI **donc elle a été prise aux impôts**↑ donc trop contente donc **super dur hein les impôts**
BEA hm [hm
ELI **[mise en passant mais bon** (0.8) trop sale ambiance en plus (0.2) tu m'étonnes↑
MAR ouais y a moyen j'ima[gine pas
BEA [quand tu dois d' mander d' l'argent
MAR ((rires))
ELI ((rires)) **tu m'êt[onnes**
MAR [vous devez deux-mille-cinq-cent euros hm
ELI ((rires)) **oh putain grave**

- BEA ((rires))
MAR ((rires))
ELI **et donc du coup** euh elle elle était là bas↑ et euh m~ et j` crois qu` c'est une ou deux s`maines après euh sa rentrée d'elle↑ **y a un gars qu` arrive en fait (0.3) et du coup c'est c'est lui [quoi]↓↓**
BEA [c'est lui]↓↓
ELI et i`s s` sont euh~ C'EST lui ((rires)) c'était lui↑ ((rires))
MAR (inaud.)
ELI **et du coup bon ben voilà** i`s s` sont entendus i`s ont commencé à sortir et **puis voilà↑** (0.3)
BEA ah c'est cool

Les récits à l'oral peuvent également être plus courts et intégrés dans des macro-activités langagières et interactionnelles, telles que décrire, argumenter, expliquer ou trouver un accord. Les récits sont également accompagnés de regards pour montrer l'étonnement ou d'autres émotions et de gestes qui viennent accompagner ou appuyer le discours comme l'illustre l'exemple suivant.

Exemple 8 : Un étudiant (E) explique son problème à l'agente de la scolarité (A).

- A vous êtes en quoi ↑
E alors en socio ((s'assoit)) **en fait** j'étais en L2 de socio de sociologie l'année dernière **((croise les doigts les mains posées sur le bureau de l'agente))**
A oui
E et j'aimerais me réinscrire sauf que je pensais que j'avais jusqu'au 15 octobre pour m'inscrire
A oui
E et euh j'ai été sur le site et le site ça marche pas et je suis venu ici **du coup ((lève les mains - les doigts croisés))** ce matin
A oui vous avez pas fait votre inscription pédagogique↑
E et on m'a dit justement voilà que l'inscription était jusqu'au 5 octobre [et du coup
A [et l'inscription
pédagogique↑ les cours↑
E et ben **les cours en fait j'y je les j'y suis j'y assiste ((mains sur les genoux, coudes écartés))** mais justement j'ai pas eu le enfin j'ai pas l'emploi du temps **((écarte les bras sous la table))**
E vu que je me suis pas encore inscrit parce qu'en fait [j'attendais ma paie de ce m- **((petit doigt de la main droite sur la table - fait un mouvement temporel de gauche à droite))**
A [comment vous faites pour suivre
les cours alors que vous avez pas l'emploi du temps↑
E ben parce que j'ai des copains qui y sont donc euh **((hausse les épaules))** je suis [leur
emploi du temps leur cours
A **[d'accord]↓ ((se tourne vers son ordinateur))** alors donnez-moi votre numéro d'étudiant

Cet exemple montre que le récit est une séquence dans une autre plus longue, qu'il est co-construit par les interactants, qu'il est structuré par des marqueurs interactionnels et que la multimodalité est mobilisée pendant tout le récit.

De telles réalisations syntaxiques recueillies *in situ* sont bien entendu réellement attestées (sans script ni consigne) et analysées dans les travaux sur la langue parlée en interaction, mais elles ne sont quasiment jamais enseignées dans les cours traditionnels

élaborés à partir de dialogues construits, normalisés, scriptés et joués. Par ailleurs, Clapi-FLE met à disposition des enseignants des fiches explicatives qui documentent certains de ces phénomènes dans leurs dimensions lexicales, syntaxiques, prosodiques, multimodales et interculturelles ainsi que des fiches « En Pratique » directement utilisables en salle de classe pour expliciter l'organisation de cet oral en interaction. Dans une séquence pédagogique, ces ressources peuvent s'articuler au concordancier de Fleuron pour identifier et comprendre en contexte le français parlé auquel les apprenants seront réellement confrontés.

Nous proposons donc d'exposer régulièrement les apprenants aux interactions recueillies par les linguistes de l'oral et aux ressources qu'ils ont construites en collaboration avec les didacticiens afin qu'ils puissent acquérir les mécanismes de l'oral comme s'ils étaient en immersion et y prendre part à leur tour de façon efficace (André, 2019 ; Alberdi *et al.*, 2020).

En outre, nous démontrerons que le modèle d'analyse syntaxique « en grille » (Blanche-Benveniste, 1990b), lorsqu'il est expliqué à des apprenants, permet de les rassurer, de les décomplexer et de faciliter leur prise de parole en interaction.

Exemple 9 : Un enseignant qui présente un dictionnaire illustré (FLEURON).

E ça va **de de de** la petite cuillère jusque à l'architecture à la plomberie la mécanique↑ tout↓
tous les thèmes

Exemple 10 : Un étudiant (E1) qui « piétine » lors d'un goûter chez E2 (FLEURON).

E2 **non ben euh on va mang- mang- manger du gâteau** et puis je prendrai une bière après

E1 ouais après ok n'hésitez pas à me demander

En grille

non	ben euh				
	on	va	mang-		
			mang-		
			manger	du gâteau	

Exemple 11 : Des colocataires font une liste de courses (FLEURON).

E1 après on peut prendre des des oui des légumes quoi des fruits

E3 oui mais ça le reste le c- comme d'hab légumes fruits

En grille

E1	après on peut prendre	des			
		des oui			
		des légumes	quoi		
			des fruits		

E3	oui mais	ça le	reste		
		le c-			
				comme d'hab légumes fruits	

Les enjeux sont importants dans la mesure où c'est en interagissant de façon efficace et appropriée, que les locuteurs - qu'ils soient scolaires, étudiants, migrants, salariés ou en insertion - réussissent à s'intégrer, se socialiser, étudier, travailler et s'émanciper.

Bibliographie

- Alberdi, C., Etienne, C., Jouin-chardon, E., (2020). Comprendre les spécificités du français oral par l’immersion virtuelle : un défi possible pour les apprenants. In *Analyser, écrire, traduire et enseigner les langues à l’ère du numérique, Études linguistiques - didactique des langues*, p 19-38.
- André, V. (2020). Corpus d’interactions et apprentissage du français langue étrangère. In Benzitoun C., Rebuschi M. *Les corpus en sciences humaines et sociales*. Nancy : Presses Universitaires de Nancy, p.101-121.
- André, V. (2019). Pourquoi faire de la sociolinguistique des interactions verbales avec des enseignants et des apprenants de Français Langue Étrangère. *LINX* 79. <https://doi.org/10.4000/linx.3694>
- Blanche-Benveniste, C. (1990a). *Le français parlé. Études grammaticales*. CNRS Éditions.
- Blanche-Benveniste, C. (1990b). Un modèle d’analyse syntaxique « en grille » pour les productions orales. *Anuario de Psicologia* (47), p.11-28.
- Blanche-Benveniste, C., Jeanjean, C. (1987). *Le français parlé. Transcription et édition*. Didier Érudition.
- Giroud, A. & Surcouf, C. (2016). « De “Pierre, combien de membres avez-vous ? ” à “Nous nous appelons Marc et Christian” : réflexions autour de l’authenticité dans les documents oraux des manuels de FLE pour débutants », *Congrès Mondial de Linguistique Française*, p. 1-18.
- Kerbrat-Orecchioni, C. (2005). *Le discours en interaction*. Armand Colin.
- Ravazzolo, E., Etienne, C., (2019), Nouvelles ressources pour le FLE à partir des études en interaction. *LINX* 79. <https://doi.org/10.4000/linx.3454>
- Sacks, H., Schegloff, E. A., Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language* 50, p.696-735.
- Traverso, V. (2016). *Décrire le français parlé en interaction*. Éditions Ophrys.

Où en sont les études sur la variation régionale de la grammaire du français ?

Mathieu Avanzi, STIH Université Sorbonne

Les régionalismes grammaticaux du français que l'on parle en Europe ont longtemps été le parent pauvre des travaux sur la variation diatopique. Plusieurs raisons permettent d'expliquer cet état de fait :

- Les dialectologues qui ont été les premiers à s'être intéressés aux français régionaux ont surtout travaillé sur le lexique ; les difficultés d'éliciter des tournures syntaxiques particulières tout en évitant les calques étant difficiles à contourner (Séguy 1973 : 81-83, à l'exception toutefois de Tuailon 1983).
- Les spécialistes du français parlé qui ont porté leur attention sur les phénomènes de variation *a priori* régionale en syntaxe ont vite été confrontés aux problèmes de différencier ce qui relevait de la variation diatopique des autres axes de la dia-variation (Blanche-Benveniste 1977, 1991).
- Les phénomènes de variation en grammaire apparaissent surtout dans les productions moins surveillées (Gadet 1997). À une époque où les corpus de français parlé n'étaient qu'à leurs balbutiements (Blanche-Benveniste 2005), ils étaient forcément difficiles à documenter.

On a cru qu'avec le développement des corpus de français parlé dans les années 2010, on aurait pu déterminer précisément la représentation géographique de phénomènes que l'on retrouve dans bon nombre d'études descriptives portant sur les régionalismes du français d'Europe (Avanzi 2018 ; Avanzi & Thibault 2020), phénomènes qui touchent à l'ordre des mots (1)-(3), la réalisation de certaines valences verbales (4)-(6), les modes et temps verbaux (7)-(9) ou à l'usage de certaines tournures grammaticales particulières (10)-(14)¹ :

- 1) oh ben j'ai personne vu [j'ai vu personne]
- 2) on m'a ça raconté [on m'a raconté ça]
- 3) allez, rends-me-le ! [rends-le-moi]
- 4) laisse, je vais y faire [je vais le faire]
- 5) j'en peux rien [j'y peux rien]
- 6) elle lui aide à faire ses devoirs [elle l'aide à faire ses devoirs]
- 7) je l'ai eu su faire [je savais le faire, mais je ne sais plus]
- 8) il veut pleuvoir [= il va pleuvoir]
- 9) les reins se bloquent, et le chat ne sait plus uriner [= le chat ne peut plus uriner]
- 10) je lui ai donné de l'argent pour lui s'acheter des bonbons [je lui ai donné de l'argent pour qu'il s'achète des bonbons]
- 11) au plus il mange, au plus il grossit [= plus il mange plus il grossit]
- 12) il a rien dormi [= il a pas dormi tu tout]
- 13) ils se gardent les enfants parmi [= ils gardent les enfants des uns et des autres]
- 14) tu veux d'autres pâtes [= tu veux encore des pâtes]

Mais bon nombre de chercheurs ont dû revoir leurs ambitions à la baisse. À témoins, malgré leur taille, les corpus OFROM (un million de mots de français romand, Avanzi et al. 2016) et VALIBEL (3 millions de mots de français wallon, Simon et al. 2014), ne comportent que quelques occurrences des régionalismes étiquetés sous (1)-(14).

¹ Les tournures sont glosées entre crochets.

Dans ce contexte, nous sommes nombreux à nous être tournés vers des techniques alternatives (*crowdsourcing* via smartphone, méthodes expérimentales) afin de pallier les problèmes de représentativité que posent les corpus. Dans cette communication, nous présenterons ces méthodes, les paramètres qu'elles permettent de contrôler (âge, sexe, origine géographique des participants), les statistiques qu'elles enjoignent à réaliser et les enseignements qu'on peut en tirer en regard de la vitalité et de l'aire d'extension des régionalismes grammaticaux en français d'Europe.

- AVANZI, Mathieu (2018). « Le pronom *y* accusatif en français régional et dans les dialectes gallo-romans : histoire et géographie », *Linguisticae Investigationes*, 41, 62-86.
- AVANZI, Mathieu, BÉGUELIN, Marie-José & Federica DIÉMOZ. (2016). « De l'archive de parole au corpus de référence : la base de données orales du français de Suisse romande (OFROM) », *Corpus* 16, 309-42.
- AVANZI, Mathieu & André THIBAUT (2020). « La variation régionale en grammaire », *L'information grammaticale*, 166, 3-6.
- BLANCHE-BENVENISTE, Claire (1977) « L'un chasse l'autre : le domaine des auxiliaires », *Recherches sur le français parlé*, 1, 100-148.
- BLANCHE-BENVENISTE, Claire (1991). « La difficulté à cerner les régionalismes en syntaxe », in G. Salmon (ed.), *Variété et variantes du français de villes. État de l'Est de la France – Alsace-Lorraine – Lyonnais – Franche-Comté – Belgique*. Champion-Slatkine : Paris/Genève, 211-220.
- BLANCHE-BENVENISTE Claire (2005). « L'étude grammaticale des corpus de langue parlée en français, in *La linguistique de corpus*, William, G. (sous la dir. de), Presses Universitaires de Rennes, 47-66.
- GADET, Françoise (1997). « La variation, plus qu'une écume », *Langue Française*, 115, 5-18.
- SÉGUY, Jean (1973). « L'accusatif prépositionnel en gascon et dans le français du sud-ouest », *Travaux de linguistique et de Littérature*, 11, 429-33.
- SIMON, Anne Catherine, CHRISTODOULIDES, Georges & Michel FRANCARD (2014). 1) Exploitation des données VALIBEL et PFC pour une étude sur les belgicisms de fréquence », communication tenue aux Journées PFC/FLORAL, Paris, 04 décembre 2014.
- TUAILLON, Gaston (1983). « Régionalismes grammaticaux », *Recherches sur le français parlé*, 5, 227-40.

Approche pronominale et sémantique référentielle : l'opposition /lui (leur)/ vs /y/

Alain Berrendonner, Université de Fribourg

• Mon but est de caractériser la valeur oppositive des deux pronoms clitiques /lui (leur)/ vs /y/ en emploi valenciel, et de rendre compte des effets de sens que C. Blanche-Benveniste a décrits sous les traits d'une opposition <+/- individualisant> :

(a) Tu lui survivras, à ta belle-mère / (b) Tu y survivras, à ta belle-mère.

En (a), « le lexique '+humain' (la belle-mère) se trouve construit avec un trait individualisant ». Mais en (b), « l'élément lexical se trouve 'désindividualisé' ». [Blanche-Benveniste & al. 1987 : 47]

• Mes données sont prises à l'écrit (Frantext contemporain, Google). La rareté des occurrences exploitables à l'oral ne permet pas de faire des comparaisons.

1. PRÉALABLES.

1.1. Théorie de la référence.

Par *objets-de-discours*, *dénotés* ou *référents*, j'entends non pas des choses, mais des entités cognitives qui sont les interprétants mentaux [Peirce 1978 : 21] des désignateurs linguistiques, et qui font partie du savoir partagé par les interlocuteurs [GP : 22 sq.]. Celui-ci est figuré comme un réseau évolutif : chaque objet-de-discours est le terme d'un faisceau de relations (ses *attributs*) qui s'incrémente au fil du discours. Ces attributs ne s'infèrent pas seulement du contenu des désignateurs utilisés pour nommer l'objet, mais aussi de tout ce qui est prédiqué sur eux. Il est bien connu, par exemple, qu'un SN indéfini peut recevoir une interprétation 'spécifique' ou non, selon le prédicat qui lui est appliqué :

- (1) (a) *Un avion* s'est écrasé hier dans les Vosges. [Kleiber 2001]
(b) *Un avion* émet énormément de gaz à effet de serre. [w]

En soi, le SN [*un avion*] dénote un objet-de-discours catégorisé mais *indéterminé* : tout ce qu'on en sait, c'est qu'il appartient à la classe des avions, mais ce peut être n'importe lequel de ceux-ci. En (1a) cependant, le prédicat, événementiel, vient ensuite conférer à cet objet une relation *s'écraser* (*hier, Vosges, x*) qui le distingue de tous les autres, et en fait un objet *déterminé*, doté d'une identité différentielle. En (1b) au contraire, le prédicat n'a pas cet effet, si bien que le référent reste à l'état indéterminé. Les objets-de-discours indéterminés de ce type n'ont pas d'identité définie, et équivalent à des variables parcourant une classe.

1.2. Sémantique des désignateurs.

Les traits qui composent le signifié des désignateurs linguistiques (SN, ProSN, pronoms clitiques) peuvent être classés en trois catégories : *essentiels*, *conjoncturels* et *casuels*.

1.2.1. Traits essentiels. Ils décrivent le dénoté par ses attributs intrinsèques : catégorie ontologique (+/- animé, +/- humain, +/- lieu, +/- action...), nombre (+/- comptable, +/- pluriel), personne (rôle dans l'interlocution), etc. Seront particulièrement concernés ici l'axe <+/- animé> et l'axe <+/- déterminé>, ex.

- (2) Ça lui ressemble, à ton chat. [+ déterminé]
Ça y ressemble, à un chat. [- déterminé] (≅ *un chat quelconque, n'importe quel chat*)

Chaque verbe recteur comporte des restrictions plus ou moins fortes sur le type d'actants qu'il admet pour termes, i.e. sélectionne des arguments porteurs de certains traits essentiels. Ex. pour le trait d'animation : *plaire à quelqu'un / *à quelque chose*.

1.2.2. Traits conjoncturels. Les objets-de-discours sont décrits d'autre part au moyen de traits qui codent leur statut cognitif momentané à l'intérieur du savoir partagé. Le seul axe d'opposition de ce genre concerné ici a trait au *degré de pertinence informationnelle*. Lorsqu'une relation est prédiquée entre plusieurs objets-de-discours, l'un de ceux-ci peut être signalé comme plus particulièrement informatif (= focus). À l'inverse, certains désignateurs marquent leur dénoté comme n'ayant ni plus ni moins de pertinence que le reste de l'information communiquée. Cf.

- (3) (a) – Qu'est-ce que tu regardes ?
- Je regarde *vous*, grand-père. [Gyp, cité par Sandfeld 1965 : 72]
(b) Je n'avais encore jamais vu de femme telle que vous. *Quand je vous regarde* je ne vois qu'un ange. [w]

Le ProSN *voüs* en (3a) marque son dénoté comme [+focus], alors que le clitique *voüs* (3b), forme de rappel, signale qu'il ne s'y attache aucune prégnance du même genre.

1.2.3. Traits casuels. Par ailleurs, chaque verbe recteur sélectionne des compléments porteurs de marques de fonction argumentale particulières, qui sont sans rapport avec leur contenu référentiel. Cf., à dénoté invariant :

- (4) Elle *lui* {plaît, nuit...} à son voisin.
Elle *en* {parle, rêve...} de son voisin.
Elle *y* {compte, insiste¹...} sur son voisin.
Elle *le* {déteste, ignore...} Ø son voisin.

Il a été suggéré de longue date que le paradigme des pronoms clitiques n'est autre qu'une déclinaison². Quant aux prépositions proportionnelles, elles ont la même fonction oppositive et il y a donc tout lieu de les analyser comme des marques de cas³.

Entre le cas et les traits essentiels d'un argument, il peut y avoir sinon implication stricte, du moins des affinités préférentielles.

2. /lui, leur/ VS /y/ : VALEURS OPPOSITIVES

L'approche pronominale des V qui construisent un régime [à SN] montre qu'ils se répartissent en trois classes :

2.1. Type A.

Ces V admettent /lui, leur/ et les pronoms de discours {*me, te, nous, vous*}, à l'exclusion de /y/. Ex. les composés causatifs en [faire + Vinf]¹ :

- (5) (a) Les Russes ont déclenché une offensive qui *leur* [*y] a fait prendre Varsovie. [Blocher-Saillens, f]
(b) Il prend un bout du chapeau d'Élomire, et *lui* [*y] fait faire un tour sur sa tête. [w]
(c) L'idée de cause mentale, c'est une cause toujours ad hoc, parce qu'en fait elle cause précisément la chose qu'on a envie de *lui* [*y] faire causer. [Castel, w]
(d) Le Ciel [...] donnera assez de force à mon amour pour *lui* [*y] faire entraîner mon esprit après le vôtre. [La Calprenède, *Herménégilde*, w]

Les causatifs peuvent être des objets animés ou non, et déterminés (5) ou non (6) :

- (6) (a) À un homme, on *lui* fait faire la guerre, on *lui* en fait faire des trucs, en lui expliquant que c'est obligatoire. [w]

¹ Ex. *J'insiste sur lui car c'est véritablement le bonhomme qui porte* [le film] à bout de bras. [w]

² Entre autres : [Damourette & Pichon 1911-1933 § 65, 2355] ; [Gaiffe & al. 1936 : 169] ; [Gougenheim 1963 : 157] ; [Pinchon 1972 : 216] ; [Moignet 1981 : 38] ; [Togebay 1983 : 212] ; [Arrivé & al. 1986 : 97]...

³ Entre autres : [Meigret 1550 : 21.17] ; [Arnauld & Lancelot 1660 : 36] ; [Régnier-Desmarais 1706 : 566]...

¹ Le clitique /lui/ a un allomorphe bref [i], connoté 'populaire', qui peut apparaître devant les verbes de type A, en commutation avec le pluriel /leur/. À ne pas confondre avec le clitique homonyme /y/.

(b) Lyncher : Mettre à mort sommairement *quelqu'un* ou *lui* faire subir des violences sans jugement régulier, en parlant d'une foule, d'un groupe. [Larousse]

Les V [*faire+Vinf*] ne comportent donc pas de restrictions sur les traits essentiels du complément causataire. S'ils excluent /y/, c'est en vertu d'une contrainte casuelle : la proportionnalité /à SN/ ÷ /lui, leur/ est la marque d'une fonction argumentale spécifique, qui est leur régime propre (= cas *datif*).

- Les occurrences de ProSN qui revêtent ce cas sont le plus souvent doublées d'un clitique, et s'interprètent comme contrastives, c'est-à-dire sont marquées [+ focus] :

(7) Pour s'expliquer à elle-même et pour *lui* faire comprendre à *lui* ses péchés perpétuels, la mère Chapdelaine s'était façonné une sorte de polythéisme compliqué. [Hémon, f]

2.2. Type B.

D'autres verbes, à l'inverse, sélectionnent /y/ comme seul clitique possible. P. ex. les *verba cogitandi* V^{co}g = {*penser, réfléchir, songer, faire attention...*}. **Je lui pense* est clairement agrammatical. Ex.

- (8) (a) J'ai pensé à vous. J'y pense sans cesse. [Lavedan, cité par Sandfeld 1970 : 137]
(b) Il était bien temps d'y songer, à la terre de ses ancêtres ! [Bénoziglio, f]
- (9) (a) Fallait y penser, à des rillettes de haricots blancs ! [w]
(b) J'y songe, à un ostéopathe. Mais j'ai peur de tomber sur un charlatan. Comment savoir si c'est un bon ? [w]
- (10) (a) Il est très populaire, songez-y. [Vautel, cité par Sandfeld 1970 : 136]
(b) J'y pense pas, j'y pense pas, j'y pense pas, que ça gratte ! [w]

/y/ dénote alors des objets déterminés (8), indéterminés (9) ou des faits (10) ; ces objets sont animés ou inanimés. Les V^{co}g n'imposent donc aucune restriction sur les traits essentiels du complément. Si /y/ est le seul pronom clitique possible, c'est aussi un fait de rection casuelle. Conclusion : le paradigme des cas du français comprend deux cas en /à/ distincts, caractérisés respectivement par la proportionnalité des compléments lexicaux [à SN] avec /lui, leur/ (cas **datif**) vs /y/ (cas **locatif**).

• Au locatif, les ProSN semblent assumer deux valeurs. On les emploie au lieu de /y/ lorsqu'on veut signifier explicitement que le référent est [+ déterminé]. Cf.

- (11) (a) Cette femme, j'y pense tout le temps, elle m'obsède. [w]
(b) J'aime vraiment cette femme. Je pense à *elle* très souvent. [w]

Mais on trouve aussi des ProSN référant à des objets indéterminés. Ils semblent employés de préférence à /y/ parce qu'il s'agit de désigner un référent [+ animé]. Et si le référent est une personne de discours, seul un ProSN est possible :

- (12) (a) *Un ami, on pense à lui* quand il n'est pas là. [w]
(b) Penser à *quelqu'un* signifie ne pouvoir lui dire qu'on est en train de *penser à lui*. [w]
(c) Je pense à *toi*. [w] / *Je te pense

Au locatif, les ProSN semblent donc être utilisés pour suppléer l'absence de pronoms clitiques autres que /y/. Celui-ci étant non marqué quant aux valeurs d'animation et de détermination, ce sont les ProSN qui sont chargés de signifier en cas de besoin les traits positifs [+ animé] et [+ déterminé]. Par suite, ces ProSN supplétifs ne marquent plus le degré de pertinence informationnelle : ils peuvent être [+ focus] ou non :

- (13) (a) J'ai pensé à *lui*, pas à *toi*. [w]
(b) [Théo] Il me plaît bien. C'est souvent que je pense à *lui*. [w]

• Une autre caractéristique du locatif réside dans les prépositions qui le marquent. Après certains *verba cogitandi*, /à/ alterne avec d'autres Pré^{loc} = {*en, dans, sur*} :

- (14) (a) Si j'ai bien compris, vous croyez *en* la Bible, *mais pas à* la totalité de ce qui y est écrit ? [w]
(b) Je crois *dans* la réalité de mes rêves. [w]

- (15) Il faut penser *dans* les autres il faut pas être égoïste i(l) faut penser *dans* ceux qui souffrent [oral, CFA80]

Ce qui caractérise le locatif, c'est donc la proportionnalité /y/ ÷ Prép^{loc}. La forme [à] recouvre apparemment deux prépositions homonymes : l'une, /à¹/, est la marque exclusive du datif ; l'autre, /à²/, fait partie des Prép^{loc} marquant le locatif.

Parmi les verbes régissant le locatif, les uns admettent diverses Prép^{loc} et des ProSN^{loc} = {là, là-bas, ailleurs, ici, où}. Les autres sélectionnent une seule Prép^{loc}, et n'admettent pas les ProSN^{loc}. Cf.

- (16) Je vais à Paris/ *en* Italie / *dans* le Midi/ *sur* la Côte / là-bas /ailleurs / où ?
 (17) la folle rumeur qui *incite à la haine* / **en/ dans/ sur* la haine / *là-bas /*ailleurs /*où ?
 je compte *sur* eux / *à/ *dans/en* eux/ *là / *ailleurs...

2.3. Résumé.

- Le système casuel du français comprend deux cas caractérisés par les proportionnalités suivantes :

Verbes	Pronoms clitiques	Prépositions	ProSN définis	Cas
A	<i>lui, leur</i>	/à/	[+focus]	Datif
B	<i>y</i>	Prép ^{loc}	[±focus]	Locatif

- Sont de type A les *verba dicendi*, les verbes de don, les lexies causatives [faire Vinf]. Sont de type B les *verba cogitandi*, les V de mouvement ou de transfert spatial.
- Il y a en outre de nombreux verbes polysémiques, qui ont deux constructions de sens différents, l'une dative et l'autre locative. Leur clitique complément fonctionne alors comme une marque de cas signalant un emploi de type A ou B (*Il y arrive* / *Ça lui arrive*)
- Il existe aussi quelques verbes trivalents qui construisent simultanément les deux cas :
 - (18) La plupart de nos joueurs ont pour principal objectif de réussir leur vie et l'anglais peut *leur y aider*. [w]
 - (19) (a) [Mon ex...] Je n'aurais jamais dû *lui répondre à son message* de saint valentin. [w]
 (b) Elle vous a envoyé un mail le 13/06/2020 à 14:01 ; je vous invite à le lire et à *lui y répondre* [w].
 - (20) Les enfants ont besoin du soutien financier de leurs parents, et la loi *leur y donne droit*. [w]

2.4. Type C.

2.4.1. Il y a cependant une 3^{ème} classe de verbes qui admettent concurremment /y/ et /lui, leur/ ou les pronoms de discours, sans différence apparente quant au procès signifié². Ces verbes ne semblent pas avoir deux constructions différentes, mais admettre les deux clitics comme instanciations de la même fonction argumentale. Contrairement aux verbes A et B, ils exercent en outre des restrictions sélectives sur les traits essentiels de leur argument.

2.4.2. /y/ dénote alors des référents déterminés ou indéterminés, tandis que /lui, leur/ apparaît réservé à des objets déjà déterminés :

- (21) (a) Et puis ne vous comparez pas à *la Bovary*. Vous n'y ressemblez guère ! [Flaubert, f]
 (b) *Ça y* [*lui] ressemble à une étuve ici. [w]
 (c) Puisque mon neveu veut faire le grand garçon, je vais l'aider à *y* [*lui] ressembler (à un grand garçon). [w]
 (d) Attendre n'est pas mourir, mais *ça y* [*lui] ressemble. [w]

² Ce sont entre autres les V de ressemblance (*ressembler, équivaloir, correspondre, associer, comparer...*) [Willems 1981 : 90], et divers autres comme *survivre, obéir, être fidèle...*

- (22) (a) Comme tu *lui* ressembles, à ton père ! [w]
 (b) Je n'irai pas sur *la Lune*, mais je sais un endroit qui *lui* ressemble... [Fallet, f]

On a donc affaire à une opposition privative : /*lui*/ : [+déterminé] vs /*y*/ : [∅]

2.4.3. Les deux pronoms s'emploient pour désigner des animés aussi bien que des inanimés (21-22). Mais un comptage des occurrences régies par le verbe *ressembler* (Frantext contemporain) fait apparaître la répartition suivante de leurs référents :

	+ dét	- dét	
+ an	341	0	92%
- an	31	0	8%
	100%	0%	

	+ dét	- dét	
+ an	6	2	5%
- an	61	43	95%
	60%	40%	

Il y a une nette tendance à préférer /*lui*, *leur*/ pour désigner les animés, et /*y*/ pour les inanimés.

2.4.4. Pour les régimes des verbes C, il semble donc y avoir interférence de deux règles :

- R1 : *lui*, *leur* : [+ déterminé] / *y* : [∅]
 R2 : *lui*, *leur* : [+ animé] / *y* : [- animé]

Ces règles sont susceptibles d'entrer en conflit, et les occurrences de pronoms prennent des valeurs différentes selon qu'elles satisfont à toutes les deux, à une seule ou à aucune :

R1	R2	
+	+	optimal
+	-	sub-optimal
-	+	inusité
-	-	inusité

Les occurrences où seule R1 est satisfaite sont attestées (22b), mais relativement rares (6% des occurrences totales). Les occurrences où seule R2 serait satisfaite (= /*lui*/ référant à des animés indéterminés) sont inattestées. Cela suggère que les deux règles sont soumises à un ordre de priorité, la violation de R1 engendrant toujours une agrammaticalité, tandis que celle de R2 est 'passable', dès lors que R1 est satisfaite. Cette hiérarchisation vaut du moins pour le genre de discours représenté dans Frantext (écrit savant)³.

3. CONCLUSIONS.

3.1. Conception de la grammaire

Les pronoms /*lui*, *leur*/ vs /*y*/ sont plurifonctionnels, et présentent trois rendements oppositifs différents, autrement dit obéissent alternativement à trois règles :

- R1 : *lui*, *leur* : [+ déterminé] / *y* : [∅] (traits conjoncturels)
 R2 : *lui*, *leur* : [+ animé] / *y* : [- animé] (traits essentiels)
 R3 : *lui*, *leur* : [+ datif] / *y* : [+ locatif] (traits casuels)

R1 et R2 semblent soumises à un ordre de priorité constant R1>R2 (v. *supra*). Le niveau de priorité accordé à R3 en revanche, semble conditionné par le type lexical du verbe recteur. Dans la dépendance des verbes A ou B, c'est le marquage du cas qui prévaut sur celui de traits essentiels, que les clitiques

³ R2 est évidemment le produit intériorisé des prescriptions normatives ambiantes. Cf. les notions de 'grammaire première' vs 'grammaire seconde' chez Blanche-Benveniste. [1990]. Contrairement à ce que ces termes pourraient laisser croire, je considère toutefois que les deux règles R1 et R2 ne relèvent pas de deux grammaires différentes, mais font partie de la même grammaire (non consistante). Seul diffère le poids relatif qui est attribué à chacune d'elles dans les divers genres de discours.

ne prennent pas en charge ($R3 > R1 > R2$) ; tandis qu'auprès des verbes C, c'est la priorité inverse qui s'applique ($R1 > R2 > R3$)¹.

Ce qu'il y a de plus intéressant à en retenir, c'est que la morpho-syntaxe d'une langue n'est pas un ensemble parfaitement consistant de règles, assimilable à un système formel ou à une algèbre homogène. Elle comprend ici et là des structures sous-spécifiées, dont notamment des oppositions de signifiants qui revêtent *ad libitum* plusieurs valeurs signifiées différentes.

3.2. L'effet « ± individualisant »

Une particularité des oppositions privatives est que leur cas non marqué, lorsqu'il est employé au lieu du terme marqué là où celui-ci serait pertinent, prend une valeur marquée inverse. C'est le cas pour R1 : dans

(23) Tu y survivras, à ta belle-mère.

le clitique /y/ prend implicitement la valeur [-déterminé, -animé]. Comme l'a bien vu Blanche-Benveniste [*& al.* 1987 : 47], il y a donc contradiction entre le pronom et le SN coréférentiel (= figure rhétorique de syllepse). L'interprète de (23) résout ordinairement cette contradiction en inférant que l'objet-de-discours dénoté est une entité à géométrie variable, qui est à la fois un individu et une réalité vague dont celui-ci est le centre (□ <ta belle-mère, et tout ce qui va avec (c'est tout un)>). De tels objets-de-discours aux contours élastiques témoignent du fait que la dénotation par le langage se fait souvent « en gros » : il n'est pas nécessaire pour être pertinent d'assigner à tout référent une extension claire et distincte. Voir la notion d'*objet indiscret* [Berrendonner 1994, 2013].

¹ Question pour l'instant sans réponse : à quelle particularité sémantique les verbes C doivent-ils le fait que le marquage du cas est ainsi sacrifié sur leurs pronoms compléments, au profit de traits essentiels ?

**La question des unités d'analyse en syntaxe :
Autour de la notion de séquence verbale/chaîne verbale
Étude sur corpus**

Mireille Bilger, CRESEM Université de Perpignan
Paul Cappeau, Université de Poitiers

Dans plusieurs articles et ouvrages, Claire Blanche-Benveniste a régulièrement discuté des unités sur lesquelles devait se fonder une analyse syntaxique apte à intégrer des données orales (entre autres, Blanche-Benveniste 1993, 2002). Elle rappelle, (cf. 2008 : 213), en faisant référence à Simone (2004) et Lazard (2001) que « les linguistes ont toujours estimé que la délimitation des unités était une de leurs tâches fondamentales ». On ne sera pas surpris, lorsqu'on connaît son parcours, qu'elle ait cherché à mieux appréhender la façon dont s'organisent le verbe et les éléments qu'il construit : cela a conduit à la notion de *construction verbale*, fondamentale dans l'outillage syntaxique qu'elle a élaboré. Elle a aussi cherché, à travers la notion de *chaîne verbale* (2002) ou d'entassement, à rendre compte de phénomènes d'agrégation peu ou mal pris en compte à ses yeux.

Notre propos sera certes de réfléchir sur la délimitation et la constitution de ces séquences/chaines verbales et de leur intérêt dans le cadre d'étiquetages formels mais aussi et surtout de voir comment elles se développent/distribuent dans les corpus.

Après avoir présenté quelques découpages en lien avec le verbe (en 1), nous reviendrons brièvement sur la description/constitution des entassements verbaux et sur divers problèmes que l'on rencontre, quant à l'intégration ou pas dans la *séquence/chaîne verbale* de certains éléments selon les verbes impliqués (en 2). Nous finirons par l'étude des données issues de différents corpus et par la description détaillée d'une séquence verbale dans laquelle apparaît le verbe « proposer de » (3).

1. Autour du verbe

1.1. Les entassements

Au moins deux raisons peuvent être avancées pour expliquer la réflexion conduite par Blanche-Benveniste autour de la notion de chaîne verbale : le comptage des verbes et l'étiquetage. D'une part, dans le prolongement des travaux de Halliday (1985) qui voyait l'opposition entre *nom* et *verbe* comme l'une des caractéristiques permettant d'opposer l'écrit et l'oral, Blanche-Benveniste (2002) a discuté de la difficulté que compter des catégories aussi larges pouvait présenter. Elle cite l'exemple repris en (1).

(1) *on va être amené à les faire travailler (cité comme exemple 9 dans Blanche-Benveniste 2002)*

Du point de vue morphologique, on dénombre cinq verbes (*aller, être, amener, faire et travailler*). Du point de vue syntaxique, plusieurs options se présentent : compter quatre unités (en traitant *être amené* comme une séquence au passif), trois (si l'on considère que *va être amené* forme un ensemble autour du modal *va*) et même un seul, si *va être amené* est analysé comme une modalisation du bloc avec causatif *faire travailler*. Cette fluctuation montre l'importance d'être au clair sur les éléments à comptabiliser. Dès 2002, Claire Blanche-Benveniste en s'appuyant notamment sur les travaux de Chu (1989) introduit la notion de *chaîne verbale* pour rendre compte d'exemples comme (2) et (3), notion qu'elle reprend dans son ouvrage de (2010, p. 111) :

(2) *il a l'air d'avoir commencé à se mettre à pleuvoir*

(3) *il a dû vouloir en faire de même*

L'étiquetage des corpus, même s'il n'est pas directement évoqué par Blanche-Benveniste, confronte aussi les linguistes à la question des unités à saisir. Dans cette perspective, Bilger et Campione (2002) identifient des séquences (sujet, verbe, etc.) qui permettent de regrouper des éléments solidaires. Ainsi l'exemple (4) est constitué d'une *séquence verbale*² :

(4) *il devrait pouvoir les acheter (cité p. 121)*

Ces découpages en grande masse ont pour intérêt principal d'isoler ces suites verbales sans entrer immédiatement dans l'analyse interne. Dans le cadre des analyses automatiques, cet aspect est loin d'être négligeable, tant on sait à quel point l'entrée immédiate dans les détails peut être compliqué et propice à des analyses divergentes (Beguelin & alii 2020).

1.2. Des découpages bien différents

Dans la terminologie traditionnelle, divers découpages qui incorporent un verbe sont proposés : *verbe* ou *prédicat*, *groupe verbal* et *syntagme verbal*. Comme le rappelle Deulofeu 2015, ces unités soulèvent des questions. Ainsi on peut se demander pourquoi isoler le seul regroupement verbe + COD/COI dans le groupe verbal... Aussi, nombreux sont les linguistes qui se sont interrogés sur l'intérêt de disposer « *d'une unité verbale qui représente la tête lexicale du verbe et des éléments adjacents comme modaux, causatifs et verbes supports, sans pour autant y inclure les compléments du verbe, (Abeillé 2002)* ». Plusieurs dénominations ont été proposées dans ce sens, entre autres, les informaticiens ont souvent choisi le terme de *verbal chunk*³, (Abney (1991)).

Même en s'en tenant à cette dernière conception d'une nouvelle unité d'analyse des formes verbales, différents choix sont envisageables, en lien étroit avec les faits que l'on souhaite dégager. On reviendra dans la partie suivante sur certains problèmes de délimitation de la séquence verbale.

2- Caractéristiques de la séquence/chaîne verbale⁴

2.1. Délimitation des séquences

La séquence ou chaîne verbale correspond à la partie verbale de la construction dans son extension linéaire. Dans un énoncé, cette séquence existe aux côtés d'autres types de séquences fonctionnelles, comme les séquences Sujet, Objet, etc., c'est-à-dire qui se comportent comme des « blocs » du point de vue de la syntaxe⁵.

Bilger & Campione (2002) avaient proposé la définition suivante : c'est la partie verbale de la production contenant tous les éléments « immédiats » dépendant du verbe qui ne sont pas susceptibles de développer une syntaxe interne propre. Elle est constituée de la forme verbale tensée, des pronoms clitiques (sujet et compléments), des particules négatives, des participes, des infinitifs, de certains adverbes. L'extension de cette séquence peut être plus ou moins grande (voir 2.3). On peut indiquer deux bornes à la séquence verbale : à gauche, c'est le sujet clitique (lorsqu'il est présent) et à droite la forme verbale (tensée, participiale ou infinitive) située la plus à droite et en dernier : [il traverse], [j'en ai parlé], [on ne peut pas rester] ici, son fils [va essayer de mieux les couper]

Par la suite, nous nous intéresserons plus particulièrement aux séquences verbales qui comportent au minimum un verbe à l'infinitif.

2.2. Les places contraintes dans les séquences.

² Cf. également les applications qui en ont été proposées dans Dister & al. 2008 et Simon & al. 2011

³ Les termes tels que *chaîne verbale* (Blanche-Benveniste 2002), *séquence verbale* (Bilger & Campione 2002), ou encore *amas verbal* (Gerdes & Kahane (2006) renvoient également à ces unités constituées de plusieurs *verbes*.

⁴ On ne fera pas, par la suite, de distinction entre ces deux désignations.

⁵ On retrouve la dualité que possède la notion de *verbe* à la fois partie du discours (comme *nom*, *adjectif*, etc.) et « fonction » (comme *sujet*, *objet*...).

On peut reprendre en grande partie le cadre de l'approche pronominale (Blanche-Benveniste 19xx) qui a servi à représenter les paradigmes syntaxiques dans une construction verbale. En effet, plusieurs éléments qui figurent dans une séquence verbale sont soumis à des contraintes catégorielles (la micro-grammaire des clitiques) qui rendent prévisibles leur apparition dans la chaîne linéaire.

Voici à titre d'illustration l'organisation des séquences avec un ou deux verbes.

- avec un seul verbe tensé

	clit Po	négation 1	clit O	Aux	négation 2	V
	je	ne	le lui	ai	pas	dit
(Julie)			s'	est		blessée

- avec 2 verbes, le second étant sous forme d'infinitif

	clit Po	nég. 1	clit O	Aux	nég 2	modal	Nég.	Clit O	Aux	V
	je	n'		ai	pas	voulu		les		voir
(Pierre)				a		su		le lui		dire
	il	n'		aurait	pas	pu	ne pas	le		faire

L'architecture devient rapidement complexe lorsque l'on doit tenir compte des diverses possibilités (négation, clitiques compléments, auxiliaires) qui peuvent être réalisés avec chacun des verbes. Les corpus montrent toutefois que, dans l'usage, les séquences verbes présentent rarement la réalisation maximale ou un cumul très important des diverses composantes. Bien évidemment, les séquences à plus de deux verbes, bien moins fréquentes, tendent à être d'une longueur plus conséquente (x)

(x) Madame Ewing, je crois que [vous avez raison de vous plaindre de ne pas avoir été consultée] sur cette question. (CERF, INST)

2.3. L'extension des séquences

La « longueur » des séquences dépend donc de nombreux facteurs ; elle peut varier en fonction de contraintes formelles et sémantiques.

a) Les modalités (négation, restriction) peuvent, bien entendu, concerner n'importe quel verbe de la séquence. Le premier verbe conjugué (x), le second verbe à l'infinitif (y) ou les deux (z) :

(x) [Je *ne* peux *que* m'en féliciter] (CERF, INST)

(y) Monsieur le Président, [je m'excuse de *ne pas* être arrivée] à l'heure (CERF, INST)

(z) un processus de connivence que [nous *ne* pouvons *pas ne pas* déplorer] (CERF, INST)

b) Comme les modalités, certains adverbes peuvent aussi porter sur le premier verbe (x) ou le second (y). Et l'on peut aussi envisager un cumul (z)

(x) Il pourrait *donc* ne plus pouvoir laisser passer en masse les réfugiés (Inter. FFig)

(y)

(z) Le clonage associé à la transgénèse pourrait *aussi* servir *rapidement* à produire des copies d'animaux (CERF, SCIE)

c) L'agent de l'infinitif peut être réalisé sous forme clitique (comme indiqué dans les codages en 2.1.) ou sous forme lexicale. Dans ce cas, la Séq V tend à enfler fortement (en nombre de mots du moins) (x).

(x) un logiciel qui permet à *des ingénieurs situés à des milliers de kilomètres de distance* de se retrouver (CERF, SCIE)

d) Des listes sont envisageables sur certains éléments (pour peu que leur catégorie de réalisation le permette). Ce sera le cas des éléments lexicaux : agent de l'infinitif (x), adverbe (y) ou infinitif lui-même (z).

(x) ce logiciel permet à *des ingénieurs et à des informaticiens* de suivre le satellite

(y) ils permettront *aujourd'hui ou demain* d'obtenir une réponse

(z) il ne devrait pas pouvoir les avoir souvent fait *peindre ou modifier*

e) Le cas des incises est plus discutable. On imagine facilement l'intérêt d'inclure ces incises (x) dans la Séq V : pour l'étiquetage en particulier, on ne doit/devrait se soucier que des bornes gauche et droite. Mais du point de vue syntaxique, il semble plus discutable d'inclure ces incises : leur place ne peut pas être prévue ou codifiée à l'aide de proformes et, contrairement aux adverbes qui sont en nombre limité et dont les paradigmes peuvent être délimités, elles sont peu prévisibles (tant dans leur forme que dans leur emplacement).

(x) Je puis, *si je veux*, vous mener chez Mme Saqui. (Balzac)

Ces divers phénomènes peuvent se cumuler, comme dans les exemples suivants, ce qui rallonge sensiblement la séquence verbale :

(x) (...) et qui par conséquent permettront *demain, j'en suis sûr, aux pays africains* de s'intégrer (CERF, POLI)

(y) Mais c'est précisément la division internationale du travail qui permettra *demain à l'humanité, débarrassée une fois pour toutes du capitalisme*, de produire suffisamment en qualité comme en quantité pour assurer à chacun selon ses besoins. (Politique)

2.4. *Quelques remarques sur les verbes impliqués*

Ces chaînes verbales peuvent présenter différents degrés de « complexité », liés au nombre des éléments qui les constituent (comme nous venons de le voir) mais aussi en ce qui concerne les relations entre les termes.

1- Ces deux bornes permettent de délimiter des chaînes verbales dans lesquelles apparaissent les verbes dits modaux⁶ (et quasi-modaux) où le sujet du verbe conjugué et l'agent de l'infinitif partagent la même référence⁷, comme dans la série suivante : *il part, il doit partir, il risque de devoir partir, il risque de vouloir devoir partir*.

Nous ne reviendrons pas ici sur les questions de définition des verbes modaux qui ont été abordées ailleurs (Le Querler 1996 ; Dendale et al. 2001 ; Chu 2008). Nous n'aborderons pas non plus le classement sémantique de ces verbes en fonction des valeurs qui y sont associés, notamment aspectuelles, phasiques ou pour l'expression de la modalité (voir François 2012, Labeau & Dister 2016)). On peut renvoyer également à Blanche-Benveniste & *alii* (1990 : 92) concernant la distribution et les différentes combinaisons possibles entre les différentes séries de ces verbes modaux.

⁶ Voir Chu (2008) sur les modaux. De façon rapide, les modaux peuvent, pour la plupart, être utilisés avec le verbe pleuvoir (*il va pleuvoir; il doit pleuvoir; il commence à pleuvoir...*). Voir Blanche-Benveniste et al. (1987) pour un inventaire.

⁷ Les modaux n'ont, par eux-mêmes, aucune valence.

2- Les séquences verbales peuvent également être constituées de verbes dits « à contrôle » (Blanche-Benveniste 2010 :112) (cf. aussi « verbes à pivot) où l’infinitif fait partie de la valence du verbe conjugué et où le sujet du verbe tensé et l’agent de l’infinitif sont différents comme : *empêcher de, aider à, permettre de*, pour lesquels l’agent peut ne pas être réalisé :

- 1- d’un mini-système d’exploitation *qui permet de manipuler* des informations
- 2- Un des boutons figurant sur le capot *permet de naviguer* dans les menus

ou l’être:

- 3- un logiciel *qui permet à des ingénieurs situés à des milliers de kilomètres de distance de se retrouver*
- 4- Une année d’enquête *ne nous a pas permis de trouver* des cas de décès
cette technique *nous permettrait de démultiplier* les performances de nos animaux
- 5 -il constitue donc très vite un un capital qui va lui permettre un capital affectif émotif émotionnel sensuel *qui va lui permettre de d’élaborer* une théorie

Cela concerne aussi certains emplois des verbes dits causatifs *faire/laisser* et des verbes dits de *perception* mais pour lesquels la complexité est d’un autre type, vu que c’est la relation entre l’agent et l’infinitif qui constitue la valence du verbe conjugué, ce qui crée un regroupement « complexe », l’ensemble étant l’équivalent à « ça » (assez proche des attributs de l’objet : « je le crois coupable »).

En ce sens, pour le découpage en séquence, on aura tout intérêt à proposer un découpage identique afin de signaler l’identité des structures, et ce quelles que soient la forme (clitique ou nominale) et la place que prend l’agent de l’infinitif :

- 1- on voyait les gens parler/ on les voyait parler/ on voyait parler les gens
- 2- *Christine le regarda monter l’escalier (corpus)*
- 3- je fais travailler les enfants/ je les fais travailler
- 4- je laisse les enfants travailler/je les laisse travailler/ je laisse travailler les enfants

De fait, si on prend le verbe « sentir » : « je la sens défaillir/je sens la fille défaillir/ je sens défaillir la fille », on a affaire au même sémantisme « particulier » du verbe « sentir ».

Si du point de vue syntaxique et de l’analyse interne, les découpages en séquences posent peu de problèmes, reste celui posé par l’étiquetage tel que proposé avec nos deux bornes bien définies (le sujet clitique et la forme verbale tensée, participiale ou infinitive).

Comme nous venons de le voir ci-dessus, l’agent de l’infinitif peut se réaliser après ce dernier sous forme lexicale (donc théoriquement exclu de l’étiquetage) :

- (*je vois/fais/laisse partir*) les enfants

Étiquetage qui serait alors identique pour une séquence du type :

- (*je vois/fais/laisse construire*) le pont

dans lequel cette fois le lexique « le pont » est la valence objet du verbe.

Ce même problème pourrait également se poser dans le cas des verbes comme « croire » qui ont deux constructions liées à des interprétations différentes:

- (*Ils ont pu croire*) (*l’enfant coupable*) / *ils ont pu le croire*/
- (*ils ont pu le croire*) *coupable*

Enfin, nombreux sont les verbes en français qui sont des locutions (*prendre conscience*, etc.), comme dans :

- *et qui [nous permettent de prendre conscience] sinon de la limite (DIVS)*

exclure de la séquence le lexique qui participe de fait au sens de ces verbes semble donc délicat et nécessite réflexions.

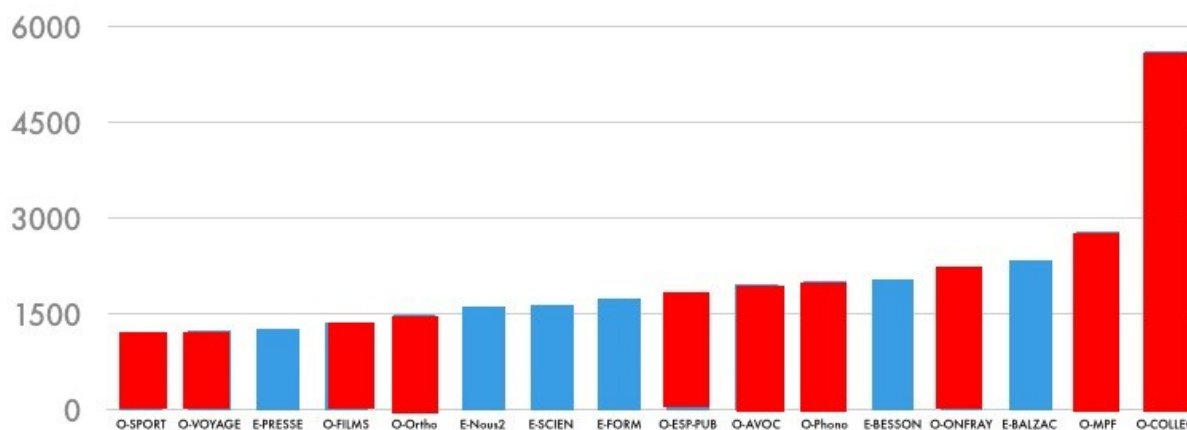
Ceci posé, il semble que l’étiquetage proposé, en dépit de ces limites, pourrait déjà permettre de dégager l’ensemble de ces séquences verbales, quitte à revenir sur certaines d’entre elles

(celles qui font intervenir certains verbes que l'on peut déjà prévoir et pré-lister) pour une analyse plus fine.

3- Les données des corpus

3.1. Présentation des corpus utilisés

Nous avons recensé les 30 premières séquences verbales constituées d'au moins deux verbes dans quinze petits corpus : 10 oraux et 5 écrits. Cela nous a permis de prendre en compte une certaine variété des productions (de oral formel, comme une conférence de type universitaire à de l'écrit plus immédiat, tel qu'on le rencontre sur un forum de cinéma). Pour atteindre les 30 occurrences de ce type de Sq verbales, il a fallu dépouiller un nombre de mots très variable : de moins de 1500 mots dans le commentaire oral d'un match à près de 6000 mots dans l'interview d'un collégien (Graphique 1). Cette interview semble assez à part puisque dans tous les autres corpus il faut entre 1500 et 3000 mots pour obtenir 30 Sq verbales. Aucune tendance nette ne se dégage quant à la coupure entre écrit (barres bleues) et oral (barres rouges).



Graphique 1. Nombre de mots pour trouver les 30 premières séquences verbales

Pour des études centrées sur des Sq qui incorporent un verbe lexical particulier (comme *permettre*), nous avons eu recours au Corpus Evolutif de Référence du Français⁸ (CERF) constitué de 10 tranches de 1 million de mots (9 d'écrits et 1 d'oral).

3.2. Les enseignements du petit corpus

Le travail à partir du petit corpus (de 450 exemples tout de même) permet d'observer certaines tendances et d'identifier des séquences dont certaines seront, par la suite, analysées dans le corpus CERF pour en vérifier la portée.

La très grande majorité des Sq comporte seulement deux verbes, (un tensé suivi d'un à l'infinitif). Les séquences constituées d'un verbe tensé (quelle que soit sa forme aspectuelle) suivie de deux verbes à l'infinitif, sont peu nombreuses, cela ne correspond qu'à 3% des occurrences et elles ne sont pas plus présentes à l'oral qu'à l'écrit. En voici quelques exemples :

- 1- la majorité des spectateurs *ne peut s'empêcher d'aller* avant tout voir (et donc lire...) (Écrit, Forum)
- 2- policier [...] qui *devait venir parler* de la bonne conscience (oral, Films)
- 3- on *n'aurait pas pu lui faire croire* que cela arriverait (roman d'amour)
- 4- Elle *s'efforça de faire taire* la douleur sourde (idem)
- 5- tu *as décidé de m'empêcher de dormir* (idem)
- 6- Le clonage associé à la transgénèse *pourrait aussi servir rapidement à produire* des copies d'animaux (Sciences-Avenir)

⁸ Constitué par Jean Véronis.

7- -il constitue donc très vite un un capital qui va lui permettre un capital affectif émotif émotionnel sensuel *qui va lui permettre de d'élaborer* une théorie (Oral- Onfray)

8-on *peut se faire aider* après manger (Oral-Anaïs)

9- Marine Le Pen *allait faire sauter* euh toutes les murailles (oral- Esprit public)

3.2.1. Le verbe en position V1

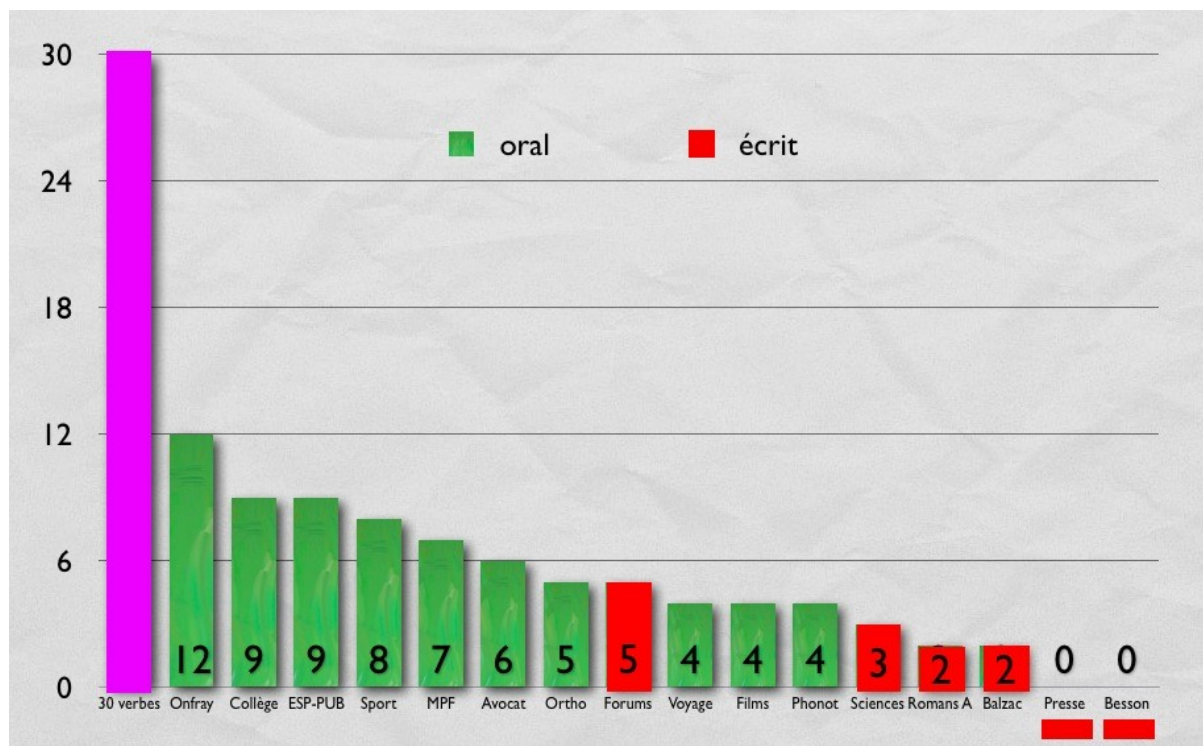
Comme cela se vérifie souvent dans les distributions de l'oral, une concentration peut être observée autour d'un nombre réduit de formes. Ainsi, près de 48% des séquences verbales font apparaître comme premier verbe tense seulement deux verbes (*pouvoir* (26%) et *aller* (22%). Dans les corpus écrits, la répartition entre les premiers verbes est nettement moins homogène, certes les verbes *pouvoir* (14%) et *aller* (6%) continuent à être les plus fréquents, mais ils ne totalisent que 20 % des séquences.

Le graphique 1 représente la répartition très inégale du verbe *aller* (en tant que V1). Globalement ce modal est présent ou très présent à l'oral. Il occupe plus d'un tiers des V1 dans la conférence de Michel Onfray, et contrairement à ce que l'on pourrait imaginer il s'agit moins, pour le conférencier d'annoncer ce qu'il va dire / faire (x), que de l'utiliser dans le récit qu'il délivre à son auditoire (y) :

(x) on va aborder la question de Thoreau sous l'angle.. (O-Onfray)

(y) c'est quelqu'un qui n- qui va passer sa vie à marcher (O-Onfray)

Dans le corpus écrit, c'est dans les forums - soit un écrit qui semble partager quelques traits avec l'oral – qu'il est le plus présent. En revanche, il est absent dans l'article de presse ou le roman contemporain (sur les 30 premières occurrences récupérées, s'entend).

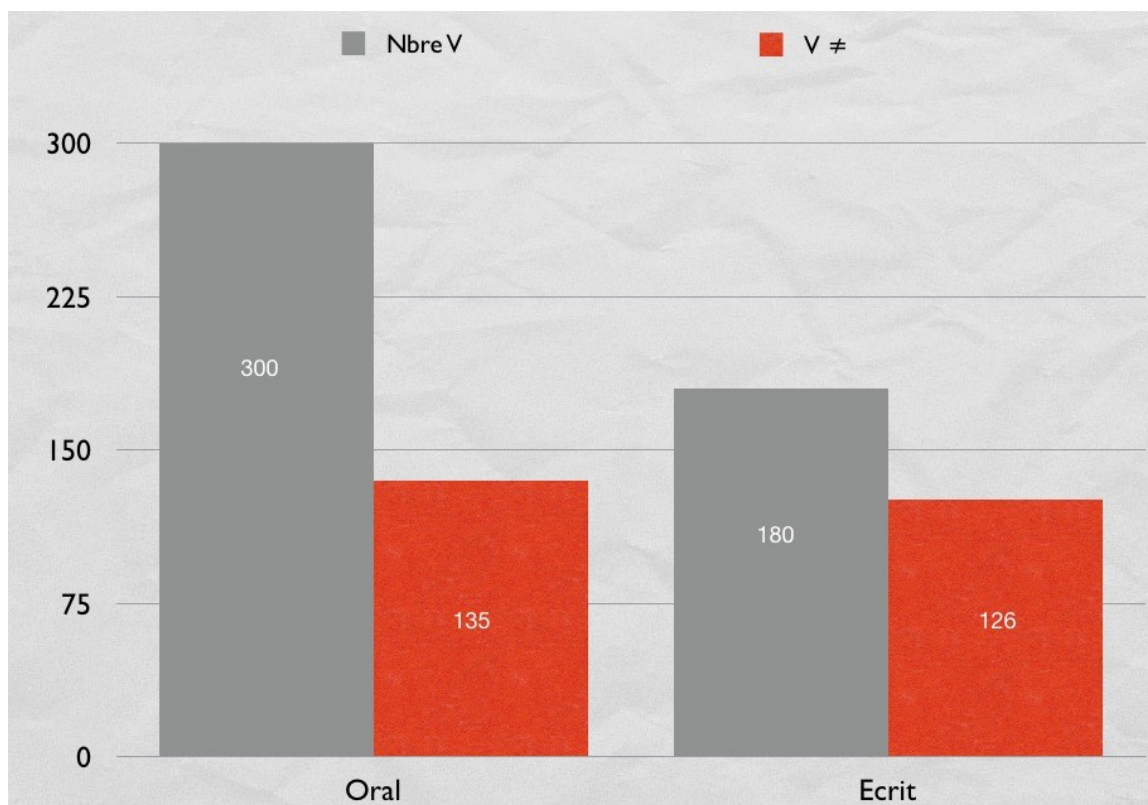


Graphique 2. Nombre d'occurrences du V1 *aller* dans les 30 séquences de chaque corpus

3.2.2. L'infinitif (souvent en position V2)

Comme signalé plus haut, ce sont les séquences à deux verbes (un verbe conjugué V1 + un infinitif V2) qui sont largement dominantes et aucune différence frappante n'apparaît entre

l'écrit et l'oral. Toutefois, une distinction entre médium concerne la variété lexicale (Graphique 3).



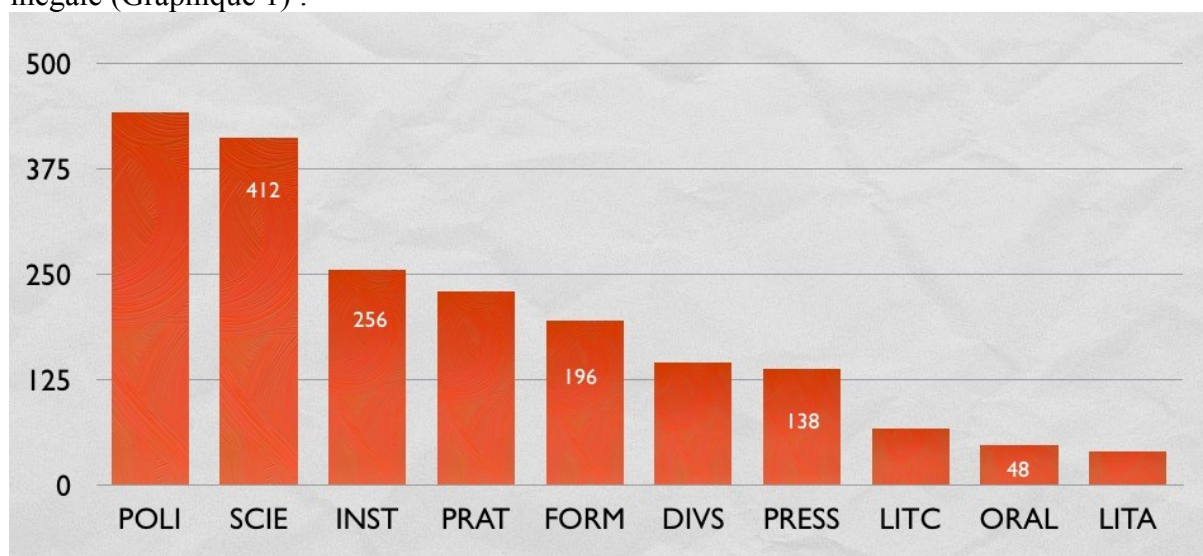
Graphique 3. Variété lexicale différenciée du V2 infinitif (oral vs écrit)

Dans les productions orales, sur 300 séquences verbales, on relève 135 verbes différents à l'infinitif - les plus fréquents étant *faire* (30) et *dire* (22).

A l'écrit, les verbes à l'infinitif sont nettement plus variés, sur 180 séquences, on relève 126 verbes différents (les plus fréquents étant *être* (11), *faire* 10, *dire* 9).

3.3. Etude sur gros corpus : les séquences avec *permettre* en V1

Ce verbe *permettre* a été retenu car il est présent dans tous les corpus, de façon toutefois assez inégale (Graphique 1) :



Graphique 4. Les occurrences de *permet- de* dans les corpus

Les corpus peuvent être répartis en trois ensembles :

- Ensemble 1 : Politique et Sciences qui comportent le plus d'exemples. Peut-être est-ce en lien avec le type de discours « démonstratif » présent dans ces deux corpus.

- Ensemble 2 : institutionnel, pratique, forums, divers et presse avec un nombre d'occurrences moyen
- Ensemble 3 ; le corpus oral et les deux corpus littéraires dans lesquels cette construction est assez faiblement présente.

L'analyse a porté sur les séquences verbales qui contiennent la forme *permet-* et dans lesquelles ce verbe construit un complément prépositionnel introduit par *de* (cela afin de disposer de séquences verbales longues). En voici des exemples (les Sq verbales sont en italiques) :

- (1) *Ils permettent de rétablir* une alimentation normale en eau potable. (PRESSE)
- (2) le contexte, qui est le médiateur *permettant de rendre* effective la différence (SCIE)
- (3) *on peut se permettre de faire* des fautes de grammaire (Oral)

Cinq corpus de 1 million de mots ont été retenus, en tenant compte des ensembles précédemment signalés : productions scientifiques (SCIE), textes institutionnels (INST), Divers (DIVS), presse (PRESSE) et Oral. On espère, par ce choix, avancer sur le lien entre séquences et « genres ».

Pour décrire les séquences verbales de façon plus fine, celles-ci ont été décomposées en deux parties⁹ :

- une bloc 1 : du clitique sujet à la préposition (mais sans inclure celle-ci)
- une bloc 2 : de la préposition à la dernière forme verbale soit infinitif, soit auxiliaire + ppé

Cette décomposition permet de mieux observer et différencier les entassements qui précèdent / suivent un verbe lexical. On pourrait aussi l'exploiter pour voir comment faire ressortir l'articulation entre ces deux sous-parties (mais cela n'a pas été exploité ici).

Le nombre d'éléments qui composent la bloc 1 présente une forte variation. On trouve ainsi des blocs à un (x), deux (y), trois (z), quatre (w) éléments voire plus (a)¹⁰.

- (x) Ces travaux [(permettaient) d'affirmer] que ce rayonnement était... (CERF, SCIE)
- (y) la recherche de voies [(lui permettant) de sortir] de la morosité (CERF, PRESSE)
- (z) [(ils vous permettent) de danser] toute la nuit sans vous arrêter (CERF, DIVS)
- (w) la communauté internationale [(pouvait désormais se permettre) de laisser] à leur sort les populations (CERF, INST)
- (a) [(ça pourra peut-être me permettre) de remettre] un C.V. euh sur pied (CERF, ORAL)

Le tableau 1 semble indiquer que l'organisation de la bloc 1 est sensible aux types de productions :

Nbre éléments	ORAL	SCIE	INST	DIVS	PRESSE
1	18 %	56 %	36 %	33 %	38 %
2	14 %	34 %	33 %	32 %	28 %
3	37 %	6 %	19 %	21 %	22 %
4 ou plus	31 %	3 %	12 %	14 %	12 %

Tableau 1. Répartition des blocs 1 par corpus

⁹ L'un des avantages de cette subdivision est de pouvoir mieux observer comment s'effectue l'enchaînement entre la préposition *de* et le Vinf i.e. avec ou sans insertion d'éléments (cf. *des thèmes permettant de mieux comprendre la politisation*)

¹⁰ La séquence V est indiquée entre crochets et la sous-séquence 1 sera placée entre parenthèses dans les quatre exemples qui suivent.

Ce tableau fait ressortir le caractère singulier du corpus oral à deux titres :

- le bloc à un élément est dominant dans tous les corpus écrits, ce n'est pas le cas dans le corpus oral ;
- le bloc « étoffé » à 4 éléments ou plus est peu présent dans les productions écrites, alors qu'il est bien attesté à l'oral.

Il s'agit là de deux pistes à creuser pour tenter d'expliquer cette différence. L'une tient à la présence de sujets clitiques qui est bien plus fréquente à l'oral qu'à l'écrit mais cette donnée ne peut, à elle seule, expliquer toutes les différences.

En outre le corpus SCIE se distingue des autres corpus écrits, plus convergents. C'est notamment le corpus où la part des blocs à un élément est si importante et celles à 4 éléments ou plus aussi faible.

Concernant le bloc 2, on trouve sa répartition dans le tableau 2 :

Nbre éléments	ORAL	SCIE	INST	DIVS	PRESSE
2	68 %	83 %	81 %	72 %	81 %
3	24 %	14 %	13 %	21 %	18 %
4	7 %	2 %	5 %	6 %	1 %
5+	< 1 %	< 1 %	< 1 %	< 1 %	0 %

Tableau 2. Répartition des blocs 2 par corpus

Ce bloc est composé au minimum de deux éléments puisque l'on compte la préposition ainsi que le verbe à l'infinitif (x). On trouve aussi des blocs à trois éléments (y), quatre (z) ou plus (a) – mais cette dernière configuration est en fait très rare).

(x) adaptées à la réalité de chacun + et qui [leur permettaient justement (de de communiquer)] parfaitement à travers ce patois (CERF, ORAL)

(y) des "résistances" qui [permettraient aux entreprises (de mieux cibler) les désirs des consommateurs. (CERF, PRESSE)

(z) sur le principe d'une "conscience extérieure" [permettant à chacun (de ne pas développer)] sa propre conscience (CERF, DIVS)

(a) un patronage c'est euh ce qui [va me permettre (de: de pouvoir euh mh + de de pouvoir en fait créer)] (CERF, ORAL)

Pour cette présentation on s'en tiendra à la description des blocs 1 à un élément et à quatre éléments ou plus ainsi que le bloc 2 à trois éléments.

a) Le bloc 1 à un élément

Lorsque le bloc ne comporte qu'un seul élément, cela peut être :

- une forme non tensée du verbe : l'infinitif (*permettre*) ou le participe présent (*permettant*) comme dans (les séquences sont en italiques, les blocs entre crochets) :

(4) il n'y a pas de programme d'ordinateur [*permettant*] de dire avec quelque succès si, dans une image, (SCIE)

- un verbe tensé sans sujet ni complément clitiques, ni verbe modal. C'est le cas dans :

(5) Cette nouvelle génération de lasers [*permettra*] de réaliser des expériences dans des domaines qui (SCIE)

Voici la distribution que le CERF permet de dégager :

ORAL – 17 %	SCIE – 56 %	INST – 36 %	DIVS – 33 %	PRESSE – 38 %
-------------	-------------	-------------	-------------	---------------

(13 sur 71)	(278 sur 495)	(121 sur 333)	(55 sur 182)	(70 sur 186)
Non tensé : 33 % <i>permettre</i> (3) <i>permettant</i> (1)	Non tensé : 40 % <i>permettre</i> (9) <i>permettant</i> (103)	Non tensé : 40 % <i>permettre</i> (9) <i>permettant</i> (39)	Non tensé : 31 % <i>permettre</i> (1) <i>permettant</i> (16)	Non tensé : 26 % <i>permettre</i> (4) <i>permettant</i> (14)
tensé : 67 % souvent sujet <i>qui</i> (7 ex)	tensé : 60 %	tensé : 60 %	tensé : 68 %	tensé : 74 %

Tableau 3. Distribution des blocs 1 comportant 1 élément

Pour ce qui concerne les formes non tensées, sans surprise à l'oral le participe présent est rare (un exemple). La surprise vient plutôt des deux corpus DIVS et SCIE ou cette forme verbale représente plus de 90 % des part de cette rubrique (les formes non tensées). Dans certains cas, on peut penser qu'on est proche d'énoncés formulaires ou de routines :

(6) arrêtons-nous un instant sur quelques définitions [*permettant*] de cadrer notre réflexion. (SCIE)

Dans les corpus écrits c'est généralement la réalisation lexicale du sujet qui explique pourquoi celui-ci n'est pas inclus dans le bloc. A l'oral l'explication tient plutôt à la réalisation d'un sujet sous la forme *qui* :

(7) la nouvelle structure de santé [*permettra*] d'offrir chaque mois aux populations un minimum (PRESSE)

(8) on commençait par le bas et ce qui [*permettait*] de cueillir le plus haut possible (ORAL)

On le voit, sur ce court exemple, une analyse plus ciblée sur certaines composantes de la séquence verbale (en l'occurrence le bloc 1) s'avère discriminante et fait effectivement ressortir des propriétés sensiblement distinctes selon les types de production.

b) Le bloc 1 à quatre éléments ou plus

Les blocs « longs » correspondent à deux schémas distincts selon que la forme verbale est tensée ou non.

Lorsque c'est l'infinitif *permettre* qui est réalisé, le bloc long est développé à gauche du verbe par la présence de divers clitiques (sujet, négation, complément) et d'un modal comme dans :

(x) [(on peut pas se permettre) de fermer] (CERF-ORAL)

(y) je suis consciente que [(je ne peux pas me permettre) de déroger] une journée (CERF-DIVS)

A noter que seul le corpus INST contient des blocs longs avec *permettant* comme dans :

(x) Ainsi des dispositions ont-elles été adoptées dans le cadre de l'article 34 de la loi de finances rectificative pour 1998 pour exonérer les acquisitions et les rétrocessions des SAFER de toute perception au profit du Trésor et pour introduire une nouvelle modalité de rétrocession [(permettant désormais à une SAFER titulaire d'une promesse de vente) de se substituer] un attributaire choisi par elle, qui ne supportera plus que les frais d'un seul acte notarié. (CERF, INST)

(y) en mesure de mener un petit élevage ovin [(leur permettant d'une part) de consommer et d'utiliser] les productions (CERF, INST)

Si c'est une forme tensée du verbe *permettre* qui est utilisée, le développement du bloc s'effectue plutôt par un ajout entre le verbe et la préposition *de* :

(x) ce sont des petites comptines sonores qui [(permettent donc une fois qu'il y a deux sons) de les associer] (ORAL)

(y) ... et la "masculinisation" de la progéniture [(permettent donc aux gobe-mouches à collier femelles) de limiter] (SCIE)

Les relevés montrent que c'est la forme non tensée /modalisée qui est largement privilégiée.

ORAL – 27 % (13 sur 48)	SCIE – 2 % (6 sur 412)	INST – 3 % (7 sur 254)	DIVS – 12 % (17 sur 145)	PRESSE – 5 % (7 sur 138)
non tensé 11	non tensé (4)	non tensé (6)	non tensé (15)	non tensé (5)
tensé 2	tensé (3)	tensé (1)	tensé (2)	tensé (2)

Tableau 4. Distribution des blocs 1 comportant 4 éléments ou plus

L'un des points qui se dégage de ces distributions concerne la présence plus forte de blocs longs à l'oral.

c) Le bloc 2 à trois éléments

On a retenu cette configuration qui présente une distribution contrastée entre d'une part l'ensemble oral et divers et d'autre part sciences et institutions. La presse se situant dans un entre-deux.

ORAL (24 %)	DIVS (21 %)	SCIE (14 %)	INST (13 %)	PRESSE (18 %)
clt Ct (14) 88 %	clit Ct (27) 75 %	Clit C (16) 70 %	clit. complément (23) 59 %	Clit Ct (11) 79 %
2 Vinf (2) 12 %	2 Vinf (4) 11 %	2 Vinf (4) 17 %	2 Vinf (2) 5 %	2 Vinf (2) 14 %
	<i>mieux</i> (2) <i>bien</i> (1) 8 %	<i>mieux</i> (2) <i>seulement</i> (1) 13 %	<i>mieux</i> (10) <i>tout</i> (1) <i>bien</i> (1) 31 %	<i>mieux</i> (1) 7 %
	Loc V (1) V acc (1) 6 %		Loc V (2) 5 %	

Tableau 5. La composition des blocs 2 à trois éléments

La composition est très contrastée selon les genres, même si le procédé le plus répandu reste la présence d'un clitique Objet comme en (x). Toutefois des variations importantes se manifestent en lien avec les genres. A l'oral la présence du clitique frôle les 90 %, soit nettement plus que dans les autres productions. La présence de deux verbes à l'infinitif (y) arrive généralement en seconde position sauf dans les textes institutionnels qui privilégient l'insertion d'un adverbe (très massivement *mieux* (z)). Enfin on trouve aussi quelques rares locutions verbales (a).

(x) [ça me permettra (de te définir)] quelques termes (CERF, ORAL)

(y) quelques grands scientifiques [leur permettent (de faire progresser)] l'audience de leurs théories (CERF, SCIE)

(z) de façon à [lui permettre (de mieux suivre)] la situation (CERF, INST)

(a) qui [nous permettent (de prendre conscience)] (CERF, DIVS)

Pour compléter cette observation des tendances qui émergent dans le gros corpus, on peut signaler que les séquences le plus souvent adverbiales¹¹ insérées dans les blocs 1 et 2 (Tableau X) présentent un double intérêt :

- d’une part, on ne retrouve aucun élément commun aux deux blocs
- d’autre part, certains adverbes apparaissent comme transgenres (ils se retrouvent dans plusieurs corpus) alors que d’autres sont plus spécifiques. Certes, il convient de rester prudent puisque le corpus reste de taille modeste pour observer des phénomènes qui relèvent du lexique, mais on dispose peut-être là de pistes en vue de renforcer le lien entre séquences verbales et genres.

Insertions	ORAL	SCIE	INST	DIVS	PRESSE
Dans Bloc 1	ainsi donc peut-être justement ben	ainsi donc peut-être aussi cependant également sans doute ici probablement même ensuite désormais notamment alors bien bien souvent bientôt certainement dans certains cas déjà donc bien essentiellement maintenant pas encore pas toujours	ainsi donc demain aussi cependant justement certainement non seulement également sans doute désormais notamment en effet d’ailleurs d’une part entre autres parfois progressivement	donc demain aussi probablement même enfin en effet le mieux	donc certainement non seulement également sans doute ici ensuite enfin du même coup vraiment

¹¹ On n’a pas retenu les séquences lexicales qui correspondent à une place de rection comme dans : pour [permettre *aux praticiens* d’enrichir] leur activité professionnelle. (CERF, INST).

Dans Bloc 2		mieux	mieux bien	mieux bien	mieux
	en fait je sais pas (avec reprise prép.)		tout réellement de nouveau	largement	seulement

Tableau 6. Les insertions (souvent adverbiales) dans les deux blocs

Ce relevé permet de s'interroger peut-être sur comment s'exprime « la modalisation » dans les séquences verbales : dans le bloc 1, certes celle-ci peut être portée par le verbe modal mais aussi par des adverbes tels que *donc, peut-être* ... Dans le bloc 2, celle-ci semble portée par d'autres adverbes nettement moins variés (*mieux, en fait, etc.*).

4. Conclusion

Nous avons voulu ici revenir sur une unité qui a beaucoup intéressé CBB, la séquence verbale, en essayant à la fois de combiner les acquis et la description d'une séquence particulière *permettre de*. L'étude de cette séquence particulière nous a permis de montrer l'intérêt qu'il y aurait à décomposer la séquence verbale en deux blocs pour en parfaire la description et mieux appréhender les liens entre séquence et genres.

Concernant l'étiquetage, si les deux bornes proposées semblent utiles pour une première analyse « automatique », certains verbes que l'on peut déjà prévoir (faire/laisser, les verbes de perceptions) ou encore certaines réalisations, par exemple, le quantifieur couplé à *en* comme dans *j'en ai trouvé quatre*, les locutions verbales) nécessiteront des ajustements par rapport à la délimitation générale de la séquence verbale.

Éléments de bibliographie

Abeillé, A., *Une grammaire électronique du français*, Paris, CNRS éditions, 2002.

Abney, S., « Parsing by chunks », in BERWICK, Robert, ABNEY, Steven, TENNY, Carole (eds.), *Principle-based Parsing*, Dordrecht, Kluwer Academic Publishers, 1991, p. 257-278.

Béguelin & alii, 2020

Benzitoun C. & alii, 2010, « Tu veux couper là, faut dire pourquoi. Propositions pour une segmentation syntaxique du français parlé », in F. Neveu & alii (éds.), *Congrès Mondial de Linguistique Française – CMLF 2010*, pp. 2075-2090.

Bilger, M. & Campione, E., 2002. « Propositions pour un étiquetage en « séquences fonctionnelles » ». *RSFP*. 17 : 117-136

https://repository.ortolang.fr/api/content/recherches-francais-parle/v1/Site_rsfp/rsfp-17.htm

Blanche-Benveniste, Cl., 1993. « Les unités : langue écrite, langue orale », *Proceedings of the Workshop on Orality versus Literacy : Concepts, Methods and Data*, Strasbourg. European Science Foundation, pp. 139-194.

- 2002. « Compter les noms, compter les verbes », in *Mélanges publiés en hommage à Gunnel Engwall*, I. Bartning, J. Falk, L. Fant, M. Forsgren, R.M. Jacobsson & J. Nystedt (éds.), Acta Universitatis Stockholmiensis, Suède : Almqvist et Wiksell international, 65-81,

- 2008. « Les unités de langue écrite et de langue parlée ». *Cahiers de l'Université de Perpignan*. 37 : 192-216.

- 2010. *Le français ; Usages de la langue parlée*. Leuven-Paris. Peeters.

Chu, X. 1989. « Examen de la structure de l'enchaînement des verbes modaux », *RSFP*, n° 9, 75-86. https://repository.ortolang.fr/api/content/recherches-francais-parle/v1/Site_rsfp/rsfp-17.htm

Chu X., 2008, *Les verbes modaux du français*, Col. L'essentiel du français, Ophrys, Paris.

- Deulofeu, J., 2015. « Les chaînes verbales : analyse externe et analyse interne ». *Repères-DoRiF*. 5. https://www.dorif.it/ezine/ezine_articles.php?art_id=194
- Dister & alii., 2008. « Approches syntaxiques en français parlé: vers la structuration en unités minimales du discours. Proceedings of the 27th Conference on Lexis and Grammar, L'Aquila, 10-13 September 2008, 27-34.
- Halliday M.A.K., 1985. *Spoken and written Language*. Oxford. Oxford University Press.
- Gerdes, Kim, Kahane, Sylvain, « L'amas verbal au cœur d'une modélisation topologique de l'ordre des mots », *Linguisticae Investigationes*, n. 29:1, 2006, p. 75-89,
- Labeau, E. & Dister, A., 2016, « La contribution des corpus oraux à la description de phénomènes de grammaticalisation. Que nous apprend le CFPB (Corpus de français parlé à Bruxelles) sur les périphrases en aller + infinitif ? », *Corpus* [En ligne], 15 | <http://journals.openedition.org/corpus/3071>
- Lazard, G., 2001, « De l'objectivité en linguistique », *Bulletin de la Société de Linguistique de Paris*, 106 : 9-22.
- Simon & Degand, 2011. « L'analyse en unités discursives de base : pourquoi et comment ? » *Langue française*, n°170, Armand Colin, pages 45 à 59
- Simone, R., 2004, « The object, the method and the ghosts. Remarks on a terra incognita » *The Linguistic Review*, 3 : 235-256.

L'adjectif adverbial : délimitation de l'objet, données, héritages épistémologiques

Gilles Corminboeuf, Université de Fribourg (CH)

1. L'adjectif adverbial : illustrations et définition

Notre exposé concerne le second volet du colloque (*La linguistique descriptive au XXIème siècle*) et consistera en un questionnement théorique et méthodologique à propos de la description du fonctionnement de l'« adjectif adverbial » (Noailly 1994) en français. Voici un aperçu des données concernées, à l'oral (1) et à l'écrit (2) :

- (1) (a) est-ce que on n'irait pas au cinéma **direct** après (oral, Ofrom)
(b) ça descendait très **raide** pis y avait un grand dévers (oral, Ofrom)
(c) tu foutais le feu ça flambait **impeccable** (oral, Ofrom)
(d) c'est pas une bannière que je vais mettre **haut** et **fort** (oral, Ofrom)
(e) putain ils parlent **bizarre** quoi (oral, Ofrom)
(f) bon au moins ça va filer **droit** (oral, Ofrom)
(g) je fais une soirée normale + enfin je regarde mon film **tranquille** (oral, Ofrom)
(h) elle était sur son lit d'hôpital elle a m- même pas encore de corset ni rien elle avait juste des anti-douleurs et elle rigolait **malade** (oral, pfc)
(i) je me suis mangé **méchant** (oral, à la volée, 15.12.2020 ; = 'je suis tombé violemment')
(j) mon papa il est assez sauvage + donc de pas voir les gens ça le dérange **moyen** (oral, à la volée, 12.3.2021)
(k) ça ça m'hallucine **complet** (oral, à la volée, 5.5. 2021)
- (2) (a) Derborence, le mot chante **doux** ; il vous chante **doux** et un peu **triste** dans la tête. Il commence assez **dur** et **marqué**, puis hésite et retombe [...]. (Ramuz, *Derborence*)
(b) [...] ils ont tracé **oblique** pour retrouver la berge sud. (Damasio, *La Horde du Contrevent*)
(c) Au bout de quelques semaines, ça se met à puer **terrible**. (Cavanna, *Les Ritals*)
(d) Il s'habille **moche**, mais attention, c'est une démarche d'expert. (presse écrite)
(e) Maintenant il va vagabonder **léger** [...]. (presse écrite, avis mortuaire)
(f) Les vieux ce n'était pas original / quand ils s'essuyaient **machinal** / d'un revers de manche les lèvres (Ferrat, « La Montagne »)
(g) Il faut dire qu'elle avait toujours couché **utile** [...] (Despentes, *Apocalypse bébé*)
(h) Clémentine Autain, députée de Seine-Saint-Denis : affirmer **haut** et **fort** la centralité de LFI à gauche. (titre, presse écrite)
(i) Ça me détend **total**. (bd, Zep)
(j) Ils sont **direct** venus à la maison [...]. (Genoux, *La barrière des peaux*)
(k) On a **facile** fait 15 bornes au total ! (web, blog)

L'adjectif est dit « adverbial » lorsqu'il forme un syntagme incorporé dans la construction du verbe, oeuvrant par conséquent en dehors du syntagme nominal, et qu'il n'est pas attributif. Cela dit, des questions de délimitation se posent immédiatement, notamment dans le cas où l'adjectif adverbial est incident à un autre adjectif – le plus souvent en antéposition et avec l'accord : **raide mort/gelé**, **fraîches écloses**, **grandes ouvertes**, **courte vêtue**, **folle amoureuse**, **bons derniers**, **beau droit**, **haut perché**, **flambant neuf**, **ras tondu**, **monstre cool**, **flambant neuf**, **une tête levée haut**, etc. Pour notre exposé, nous laisserons cet emploi de côté.

Les travaux sur l'adjectif adverbial, dont certains sont relativement récents (Delsaut 2013, Hummel & Valera 2017, Coiffet 2018, Hummel 2018a,b), nous paraissent un bon observatoire des pratiques actuelles en linguistique descriptive. En nous inspirant directement des recherches de Hummel – les plus abouties actuellement sur l'adjectif adverbial –, nous examinerons les trois questions suivantes, largement imbriquées :

- [1.] L'objet d'étude est-il délimité de la même manière d'une étude à l'autre ?
[2.] Les données convoquées sont-elles du même ordre ?
[3.] Les hypothèses sous-jacentes ou les postulats explicitement assumés par les différents auteurs sont-ils accordables avec les faits linguistiques ?

Pour cela, nous nous limiterons à l'unique question de l'invariabilité présumée de l'adjectif adverbial.

2. Le « dogme » de l'invariabilité

Par défaut, l'adjectif adverbial n'est pas accordé, parce qu'il n'est pas contrôlé par une catégorie qui requiert l'accord : « Invariable parce que ne trouvant pas à quoi s'accorder » (Noailly (1994 : 104). Or, il trouve parfois à quoi s'accorder...

Hummel (2017b : 34) montre en effet que l'accord est répandu, en particulier dans les textes anciens, en français parlé et dans les genres dits 'informels' (de même dans les autres langues romanes), mais qu'il est peu observé. Cela tient d'une part à la documentation, les genres informels et l'oral étant moins bien documentés dans les corpus à disposition que l'écrit normé destiné à la publication. D'autre part, un gros effort normatif a été produit dans le discours grammatical pour éradiquer l'accord lorsque l'adjectif est utilisé avec une fonction adverbiale (*ibid.*, 38, et Hummel 2018b). Enfin, il est peu repéré également parce qu'« à la différence d'autres langues romanes, le français possède des terminaisons qui sont explicites à l'écrit mais souvent inaudibles à l'oral » (Hummel 2018a : 263). Ainsi, pour l'unique forme /blø/ de l'oral, l'écrit possède quatre graphies distinctes (*bleu, bleue, bleus, bleues*) :

- (3) Il y avait des petits feux de bois qui fumaient **bleu** à toutes les cheminées carrées [...] (Ramuz, *La grande peur dans la montagne*, 1926)

Dans (3), *bleu* ne prend pas l'accord, mais l'écoute de cet extrait lu à haute voix ne permettrait pas de savoir s'il y a ou non accord.

Dans un exemple comme le suivant, seul l'écrit permet de repérer l'accord, les marques de flexion étant bien davantage marquées à l'écrit :

- (4) Après une deuxième mi-temps rigoureuse et appliquée et face à un adversaire qui a vendu **chère** sa peau, l'ASM s'imposait sur le fil 18 à 17. (web, lamontagne.fr)

A l'oral, l'accord au féminin de (4) serait inaudible. Voyons les exemples suivants :

- (5) (a) elles ont été testées /pɔzitiv/
(b) elles ont été testées /pɔzitif/
(c) ils ont été testés /pɔzitif/

A l'oral et avec un sujet au féminin, l'analyse est univoque : la variante (4a) est accordée et la variante (4b) n'est pas accordée. En revanche, la forme (/pɔzitif/) dans (4c) ne permet pas de savoir si l'adjectif est accordé (au pluriel).

Pour Hummel (2018a : 271) les terminaisons débiles de l'oral ont sans doute favorisé la variante normative. Les formes ambiguës de l'oral (p.ex. /blø/) sont d'office reversées dans les formes non accordées, renforçant le précepte de l'invariabilité de l'adjectif.

Hummel (*ibid.*, et 2018b) insiste sur le fait que l'invariabilité n'est pas une hypothèse ou un observable, mais un dogme. Un dogme auquel souscrivent la grande majorité des auteurs et qui, selon lui, a pour conséquence dommageable de restreindre « l'angle de vue » (Hummel 2018a : 264).

L'adjectif adverbial est accordé dans environ seize pour cent des attestations du *Dictionnaire historique de l'adjectif-adverbe* (Hummel 2018a : 273), qui compte 17'000 occurrences (la banque de données est en accès libre ici : <https://gams.uni-graz.at/context:aaif>). Voici quelques exemples où l'adjectif adverbial est accordé, dans la littérature publiée et dans les types de discours les plus divers :

- (6) (a) Et il lut les mots suivants qui étincelaient et tournaient **rapides**. (Sue, *Atar-Gull*, 1831 ; cité par Hummel 2018a)
(b) Cette grêle d'insectes tomba **drue** et **bruyante**. (Daudet, *Lettres de mon moulin*, 1869 ; cité par Goosse-Grevisse)
(c) La lampe brûlait très **haute**. (Zola, *Une page d'amour*, 1878 ; cité par Gazdik & Hummel)
(d) Chaque tartine coupée trop **épaisse** lui attirait des paroles dures. (Zola, *La Terre* 1887 ; cité par Goosse-Grevisse)

- (e) De laborieuses extravagances qui sonnent si **affectées et littéraires** que c'en devient insupportable. (Vercors, *Chevaux du temps*, 1977 ; cité par Goosse-Grevisse)
- (f) si elles étaient fringuées **pareilles** je me flinguerais (oral, transcription par un étudiant d'un énoncé saisi à la volée, Université de Fribourg 2010)
- (g) Ici on ne se moque pas de ceux qui s'habillent peu **chers**, mais de ceux qui s'habillent **chers**. (web, forum.psychologies.com)
- (h) Je connais exactement la peinture dont j'ai besoin suivant les marques ainsi que les marques qui chaussent **étroits** ou **larges**. (web, cité par Hummel 2018a)
- (i) [...] une chanson que seul lui pouvait entendre tant il la chantait **basse**. (web, cité par Hummel 2018a)

Avant de lister plusieurs exemples avec l'accord, Goosse-Grevisse (2016 : 1289) précise que « des épithètes détachées du nom et attachées au verbe sont traités comme des adverbes, jusqu'à l'invariabilité, mais celle-ci n'est pas obligatoire d'ordinaire ».

De fait, les hésitations sur l'accord sont communes :

- (7) (a) Les créateurs locaux jouent **groupés** (titre d'article, ladepeche.fr/2020/12/21)
- (b) Les agriculteurs jouent **groupé** (titre d'article, ladepeche.fr/2014/10/14)

Dans (7), le même prédicat (*jouer groupé*), avec un sujet au masculin pluriel dans les deux cas, issu du même périodique (*La Dépêche*) est accordé dans (7a) et non accordé dans (7b).

Hésitation également pour *testé positif / positive*, même si la version accordée est largement majoritaire :

- (8) (a) Ces personnes infectées malgré la vaccination représenteraient 0,12% de l'ensemble des personnes testées **positives** depuis fin janvier. (presse, *24heures*, 3.7.2021)
- (b) Plusieurs personnes testées **positif** dans l'entourage de Trump. (web, europe1.fr, 27.7.2020)

L'adjectif *court* dans (*se couper court*) s'accorde (9a, 9b) ou pas (9b, 9c), qu'il soit contigu au verbe (9a, 9c) ou non (9b, 9d) :

- (9) (a) Contrairement à Mithridate qui, pour suivre son époux au combat, se coupe **courts** les cheveux [...] (texte scientifique, books.openedition.org)
- (b) Monica Bellucci se coupe les cheveux **courts** à la garçonne [...] (web, elle.fr)
- (c) Elle a récemment coupé **court** ses cheveux et cela lui va bien. (web, lefigaro.fr/sortir-paris/2018/11/06)
- (d) Elle était persuadée que, dans la nuit, l'autre lui avait coupé les cheveux **court**. (Teulé, Googlebooks)

Hésitation encore chez cette étudiante qui, considérant l'accord fautif, fait peut-être le lien avec *remplir plein* :

- (10) L1 : Du thé ?
- L2 : Oui, mais pas **pleine** la tasse. (oral, exemple recueilli par une étudiante de l'Université de Fribourg qui voit une faute d'accord sur *pleine*, 2010)

La variation sur l'accord est souvent mentionnée par les linguistes, sans toutefois que des conséquences en soient tirées. Chez la plupart des auteurs, l'accord serait le révélateur que la construction échappe au champ d'application de l'adjectif adverbial. En général, c'est le concept d'« élargissement attributif » (ou « attribut accidentel ») qui est convoqué comme explication. Ainsi dans *il a vendu chère sa peau*, *vendre* ferait l'objet d'un élargissement attributif, élargissement qui n'a pas cours si *cher* n'est pas accordé (*il a vendu cher sa peau*). Pour le même exemple à l'oral, il serait par conséquent impossible de savoir s'il y a ou non élargissement attributif.

Levant habilement le voile sur plusieurs idées reçues qui ont la caution d'une longue tradition grammaticale¹, Delsaut (2013) parle toutefois d'« adjectif invarié » (déjà dans le titre de son article),

¹ Delsaut (2013) évoque le poids de certains postulats qui ont la caution de la tradition, en premier lieu les impasses de la polycatégorisation, la tentation endémique de ramener l'adjectif adverbial à des formes adverbiales (*La Grammaire méthodique du français* analyse ces formes au chapitre 13 « L'adverbe ») et la solidité du postulat selon lequel l'adjectif est inféodé au nom et ne peut, en conséquence, se rapporter au verbe.

précisant qu'il « ne s'accorde pas » (*ibid.*, 25). Assumant le dogme de l'invariabilité, il est en quelque sorte lui-même « victime » de cette tradition.

3. Les explications à la présence d'un accord

3.1. Une première hypothèse est celle d'une incidence au sujet. Plusieurs auteurs mentionnent ce cas de figure.

Wilmet (2010 : 347) cite les deux variantes suivantes (le point d'interrogation est de l'auteur), avec leur glose : « Marie bronze **idiot** / ? **idiote** = 'à une façon idiote de bronzer' mais aussi 'prouve son idiotie en bronzant' ». La seconde glose, qui correspond à la variante accordée, trahit une incidence au sujet. Dans *Les retraités ne veulent pas vieillir idiots*, « l'adjectif habituellement invariable s'accorde au sujet » (Noailly 1994 : 106), le caractère [+humain] de celui-ci n'y étant pas étranger. L'auteure précise en note : « Il y a aussi que le verbe *vieillir* permet un attribut accidentel du sujet : *Lise vieillira heureuse*. C'est un élément supplémentaire favorisant l'accord ».

Wilmet (*ibid.*, 578) mentionne de même le belgicisme *tourner fou* (= 'tourner à vide') qui « ne devrait pas plus accorder l'adjectif que p.ex. *Marie travaille dur*. On entend pourtant *la vis tourne folle* = 'est folle + à force de tourner' ».

Goosse-Grevisse (2016 : 1291) souligne que « *Droit dans se tenir droit* est traité comme un adjectif attribut ou comme un adverbe. [...] *Elle se tenait très droite et sage sur sa monture* (Germain, *Le livre des nuits*, 1984) [...] *Elle se tenait très droit* (Balzac, *Le curé de Tours*, 1832) »².

Gazdik (2016 : 75-76) considère que *dormir tranquille* est ambigu : ou *tranquille* modifie le verbe (lecture 'manière') ou l'adjectif fonctionne comme un « prédicat secondaire orienté vers le sujet » (caractérisant « l'état de la personne pendant qu'elle dort »). Dans cette seconde lecture, l'accord est possible. Nombreux sont les exemples d'accord au pluriel du type *ils dorment tranquilles* / *elles partent tranquilles* / *elles se tiennent tranquilles*. Pour *Blaise part tranquille*, Noailly parle également d'« attribut accidentel » ; *tranquille*, placé « après un verbe autre que copulatif » se trouve « qualifier le sujet » (1999 : 115) – sans toutefois endosser les propriétés habituelles des attributs (pronominalisation p.ex.).

Coiffet (2018 : 61) observe que *construire solidaire* dans :

- (11) Redon : les collégiens construisent **solidaire** en Roumanie pendant l'été16. (*Ouest-France*, cité par Coiffet 2018)

ne signifie pas forcément que l'agent est solidaire (cf. <les collégiens sont solidaires en construisant ceci>), mais si c'est cette lecture qui est privilégiée, l'accord est possible, comme dans les exemples (12) :

- (12) (a) Les commerçants proposent des micro-services et/ou des produits solidaires, les habitants consomment **solidaires**, les personnes sans domicile peuvent devenir ambassadeurs bénévoles. (web, rtes.fr, 10.12.2018 ; à propos d'un projet associatif)
(b) Les Français achètent **solidaires** (titre d'article, 1jourlactu.com, 30.11.2004)

Pas de doute que dans (12), l'interprétation <les habitants / les Français sont solidaires> est possible. Si pour un attribut périphérique (*elle est partie heureuse*) l'incidence au sujet peut rejaillir sur le verbe ('comment est-elle partie?'), Goes (1999 : 223-224) observe parallèlement des hésitations entre attribut et adjectif adverbial, surtout lorsque le sens de l'adjectif « a tendance à rejaillir sur le sujet ». Les exemples (13) – noter l'accord – sont de l'auteur :

- (13) (a) Le vent s'étant apaisé et la pluie tombant moins **serrée**, le cochon se remet en marche. (Aymé)
(b) La neige tombe **abondante**.
(c) L'oreille était implantée très **basse**. (TF1)

²Soulignant les incohérences du discours normatif, Hummel (2018b) note que « l'accord de l'adverbe devient 'illogique' dans *Ils vont droits au but* (alors qu'il serait 'logique' dans *Elle marche droite* ('debout')). [...] le rejet du tour *ils se font forts de* consacre la valeur adverbiale de *ils se font fort de*, à la différence de *elle se présente forte*, ou *fort* actualise sa valeur adjectivale ».

(d) Elles s'embrassent très **cordiales**. (A. Sand)

Goes (2008 : 35) observe que dans *La neige tombe dru / drue*, l'hésitation sur l'accord est commune ; une petite enquête de l'auteur sur Google montre que *dru* est, dans cette construction, accordé une fois sur deux.

Les hésitations ne sont donc pas rares entre une analyse attributive (propriété du sujet) et une analyse en adjectif adverbial (lecture 'manière') – hésitations alimentées par l'accord. Les exemples (14) le montrent encore une fois :

- (14) (a) [...] rendez-vous dans Paris, elle arrive **ponctuelle** cheveux dans la figure yeux brillants [...] (de Kerangal, *Naissance d'un pont* ; cité par Jollin 2020)
(b) Elle a répondu **présent / présente**.

L'accord est le plus souvent conçu comme une marque discriminante, qui permet d'assigner l'exemple concerné au domaine attributif. Mais, parallèlement, on pourrait également se demander, en reconduisant un raisonnement du même ordre, si un attribut non accordé n'est pas l'indice d'un emploi d'adjectif adverbial :

- (15) L1 : elles étaient bonnes les pizzas de Vito ?
L2 : ben elles étaient **moyen** (oral, à la volée, 5.2.2021 ; à propos de pizzas à l'emporter ; Vito est le prénom du patron de la pizzeria)

Avec l'accord au féminin, *moyennes* serait attribut³.

On sait en effet que tous les attributs (aussi bien du sujet que de l'objet) ne sont pas accordés :

- (16) (a) Allait-il manger tout **chaud** sa vengeance sur la tête de son agresseur ? (*LM* 18/05/1996, p. 19 ; Tobback 2004 : 216 ; cité par Goes 2008)
(b) Le Japon achète plus de films français que n'importe quel autre pays au monde (46 en 1993, 31 en 1994), il les achète **cher** (parfois plus d'un million de dollars [...]). (*ibid.*)

De manière convergente, Hummel (2018a : 272-273) souligne que Damourette & Pichon, Blinkenberg, Bauche et Goosse-Grevisse ont repéré des exemples de ce type (Ex. de Blinkenberg : *Ne soyons pas si fier ! / Un fog à rendre jaloux toutes les villes du Royaume-Uni*, etc.). Hummel (*ibid.* 272) conclut que « l'accord, qui est canonique pour la prédication seconde, n'est pas systématique non plus dans l'usage réel »⁴.

3.2. L'accord avec l'objet est commun également (voir *supra* 6d, 6g, 8a, 9a-b) :

- (17) (a) Elles sont enchantées qu'elle les laisse nigauder **peinardes**. (Despentes, *Apocalypse bébé*, 2010)
(b) [...] ils se défendaient en cherchant à vendre **chère** une vie pleine d'amertumes et d'opprobres [...]. (Boisrond-Tonnerre, 2014, Googlelivres)

Abeillé & Mouret (2010) signalent les cas suivants où « l'adjectif s'accorde avec l'objet » :

- (18) (a) Il a classé Marie **première**.
(b) Il donne la jument 3 **gagnante**.
(c) Il code **secrets** tous ces documents.

Dans (19a) vs (19b) ci-dessous, on verra probablement un adjectif adverbial dans (a) et une prédication seconde dans (b), sur la base de l'accord :

³L'énoncé *les pizzas étaient moyennes* pourrait avoir un autre sens (« de taille moyenne »), mais pas dans ce contexte, à la suite de la question de L1.

⁴Voir aussi la prescription de Vaugelas (1647, cité par Hummel 2018b) : « *Ils demeurèrent court*, et non pas, *ils demeurèrent courts* ; [...] *court*, s'emploie[] là adverbialement ».

- (19) (a) [...] les maisons se touchent toutes, parce qu'on a bâti **serré**. (Ramuz, *Posés les uns à côté des autres*, 1943)
(b) Je mets les tranches de ventrèche entre les navets boule d'or, les grosses carottes, les poireaux, les panais, les boules de céleri, rangés très **serrés** sur des plaques. (liberation.fr/futurs/2010/03/06)

Mais c'est faire reposer l'analyse uniquement sur la présence / absence d'un accord. A propos de (20),

- (20) On les a amarrées **serrées**, les barques. (cité par Muller 2000)

Les trois interprétations pour le seul exemple (20) proposées par Muller suggèrent que cela n'est pas aussi simple :

- (1) « on peut y voir, soit l'emploi quasi adverbial (d'autant plus plausible que l'adverbe correspondant n'existe pas : **serrément*⁵) » ;
- (2) soit « la supplétion d'un opérateur du type de *rendre* par le verbe *amarrer* », c'est-à-dire une forme d'élargissement attributif ;
- (3) soit encore « l'aboutissement d'une relation de concomitance entre deux actions (ou une action active et une passive) : *Quand on a amarré les barques, elles ont été serrées* » (un résultatif ?).

La lecture (i) correspond à l'adjectif adverbial et la lecture (ii) à l'élargissement attributif.

Pour Goes (2008), une différence sémantique peut être perçue, par exemple dans la paire d'exemples empruntée à nouveau à Muller (2000 : 31) :

- (21) (a) On les taille très **courtes**, les jupes, cette année. (cité par Muller 2000)
(b) On les taille très **court**, les jupes, cette année. (*ibid.*)

Goes (2008 : 31) dit de l'exemple (21a) que « l'adjectif est à la fois dans une relation attributive avec l'objet (incidence), et s'interprète comme l'aboutissement de l'action verbale (portée) ». Dans (21b) en revanche, l'adjectif selon lui, « ne qualifie que l'action verbale ». L'accord est ainsi à nouveau l'indice d'une incidence à l'objet et d'un rapport prédicatif de type attributif. De fait, Muller (*ibid.*, 31) considère que « l'emploi adverbial aura tendance à être différencié de l'emploi prédicatif par l'invariabilité de l'adjectif ».

Constatant toutefois « la nuance [...] très ténue entre les élargissements attributifs » et les exemples suivants : [...] *la neige tombe dru*. / *On achètera petit mais bon* », Goes (2008 : 25) souligne les

- (1) interférences entre l'élargissement attributif (EA)-objet et l'adjectif adverb(i)al » (*ibid.*, 30). Selon l'auteur, « les hésitations des locuteurs quant à l'accord de l'adjectif tendent à confirmer ce rapprochement : ils ont tendance à laisser l'adjectif invariable lorsque ce dernier semble qualifier plus naturellement l'action verbale » (*ibid.*). Dans la même perspective intégrative qu'Hummel, Goes (*ibid.*, 31) fait état d'une « triple alternance : EA-objet, adjectif invarié, forme en *-ment* ». Or, poursuit-il, « ce phénomène n'est possible que si l'adjectif en fonction d'EA-objet peut qualifier à la fois l'objet (incidence) et l'action exprimée par le verbe (portée) » (*ibid.*)

Certaines homonymies de construction peuvent compliquer l'assignation de l'accord :

- (22) Laissez ceux qui s'aiment **tranquille(s)**. (exemple fabriqué)

Deux lectures sont possibles pour (22), (i) l'une assimilable à un élargissement attributif (*laissez tranquilles*), (ii) l'autre à l'adjectif adverbial (*s'aimer tranquille*), glosables respectivement comme suit :

- (1) Laissez **tranquilles** ceux qui s'aiment.
- (2) Laissez-les, ceux qui s'aiment **tranquille(s)**.

Goes (2008 : 31) postule encore un point d'intersection entre EA-objet, adjectif adverbial et apposition, qui s'inscriraient tous trois dans un « continuum sémantique » :

- (23) (a) **Tranquille**, il reposait, vous l'avez déchaîné. (Balzac, *Cromwell*, 1820)
(b) Là, on est venues pour discuter, **tranquilles**. (Despentès, *Apocalypse bébé*, 2010)

⁵ Frantext livre toutefois 8 occurrences de *serrément*, certes anciennes (de 1561 à 1604, dont 6 occurrences d'Olivier de Serres).

(c) Je vous laisse cibler, **tranquilles**. (Despentès, *Apocalypse bébé*, 2010)

Verra-t-on dans (23a) une apposition adjectivale, ou un adjectif adverbial extraposé (sans incidence au sujet) ? Et dans (23b)-(c), l'accord oriente-t-il fatalement vers une lecture appositive ?

4. Conclusions

On observe des hésitations sur l'accord lorsque l'adjectif peut qualifier le sujet et l'objet, en même temps que l'action exprimée par le verbe. Ces hésitations sont beaucoup plus communes que ne le laissent penser les grammaires et les travaux qui souscrivent à l'invariabilité de l'adjectif adverbial. Hummel (2018a : 274) conclut que « l'analyse de l'adjectif adverbial comme unité invariable ne saurait se justifier. L'invariabilité est une réalité en tant que dogme normatif qui se confirme en règle générale dans la langue standard ».

Quand il y a reconstruction possible d'une incidence au sujet ou à l'objet – donc éventualité d'un accord –, on parle d'élargissement attributif ou d'attribut accidentel. Or ce n'est jamais un élargissement du domaine de l'adjectif adverbial qui est postulé – ce qui est une façon de sauvegarder le dogme de l'invariabilité. Hummel (*passim*) comme Goes (1999 : 224) concluent à un continuum entre attribut et adjectif adverbial (v. également Lauwers 2013 qui parle de pôle 'attribut' et de pôle 'complément de manière').

Selon Gazdik (2016 : 83), suivant sur ce point les observations de Ledgeway (2011) à propos des dialectes du sud de l'Italie, l'accord apparaît avec le sujet d'un verbe inaccusatif (sujet donc à rôle de thème, non agentif), ou avec l'objet du verbe transitif.

La présence de l'accord conduit à trois options descriptives :

- (1) Censurer les exemples concernés comme fautifs.
- (2) Faire reposer la distinction entre attribut (du sujet ou de l'objet) et adjectif adverbial sur l'accord. Autrement dit, considérer que l'accord conduit à une réanalyse de l'adjectif en ajout prédicatif – ce qui est probablement parfois le cas chez les sujets parlants. Le coût est un élargissement du domaine attributif, pour maintenir le dogme de l'invariabilité.
- (3) Ou, considérant que l'adjectif adverbial peut être accordé et que l'attribut peut ne pas l'être, postuler un continuum entre lecture attribut et lecture adverbiale de l'adjectif – lectures qu'il serait artificiel de distinguer sur la base de l'accord, dans la mesure où l'accord est conçu comme une propriété définitoire de l'adjectif. L'accord serait commun aux deux techniques, mais par défaut et non systématiquement, l'adjectif adverbial est non accordé (faute de support identifiable) alors que l'attribut prend l'accord.

On peut se demander s'il convient de traiter en bloc les différents usages de l'adjectif, ou s'il vaut mieux souligner la spécificité de chacune de ces constructions considérées individuellement. Hummel prône l'option intégrative, alors que la plupart des chercheurs isolent l'adjectif adverbial – et encore l'emploi où il est contiguement postposé au verbe – sur la base de l'accord. Si la séparation des deux techniques a l'avantage de mieux délimiter l'objet d'étude, elle inhibe toute ambition de saisir de manière globale le réseau constructionnel concerné et elle se fait au prix d'une approche normative qui se conforme au discours grammatical dominant et qui est empreinte de contradictions (cf. l'adjectif dont une propriété définitoire est d'être variable en genre et en nombre est invarié/invariable... Voir le titre de l'article de Gazdik (2016) : « Sur l'accord des adjectifs invariables du français »)

Le renvoi dos à dos des phénomènes de prédication seconde et d'adjectif adverbial – sur la base de l'invariabilité (une décision *ad hoc* et normative) – revient à privilégier le système linguistique tel qu'il se présente au regard externe du savant, par rapport au sentiment linguistique des sujets parlants (Béguelin 1990). La séparation des deux techniques, sans doute commode pour le linguiste dans sa pratique descriptive, apparaît comme un principe d'une part anti-saussurien, dans la mesure où il fait peu de cas de la conscience linguistique des sujets parlants, et d'autre part a-constructionnaliste, dans la mesure où il va à l'encontre du principe d'existence d'un réseau réticulaire de constructions associées.

La délimitation de l'objet d'étude sur la base de l'accord et de la position de l'adjectif (contigu ou non au verbe), la qualité de la base empirique à disposition (corpus diversifié ou non), ainsi que le poids de la tradition grammaticale (Delsaut 2013) apparaissent intimement imbriqués, et ont en définitive un impact considérable sur les résultats de la recherche dans le domaine de l'adjectif adverbial.

Bibliographie indicative

- Abeillé, A. & D. Godard (2004), « Les adjectifs invariables comme compléments légers en français », *L'adjectif en français et à travers les langues*. François, J. (éd.), Presses Universitaires, Caen, 209-224.
- Abeillé A. & F. Mouret (2010), « Les compléments adjectivaux des verbes transitifs en français », *Les Tables. La grammaire par le menu*, T. Nakamura (éd.), Presses universitaires de Louvain, 1-10.
- Béguelin M.-J. (1990). « Conscience du sujet parlant et savoir du linguiste ». Liver R., I. Werlen & P. Wunderli (éds), *Sprachtheorie und Theorie der Sprachwissenschaft. Festschrift für Rudolf Engler*. Tübingen : Gunter Narr Verlag, 208-220.
- Coiffet B. (2018), *Etude des emplois de l'adjectif invarié en français*, Thèse de doctorat de l'Université de Toulouse. En ligne : <https://tel.archives-ouvertes.fr/tel-02434692/document>
- Delsaut M. (2013), « La tradition corsète-t-elle la question de l'adjectif invarié jusqu'à un point de non-retour ? », *Travaux de linguistique* 67, 25-60.
- Gazdik A. (2016), « Sur l'accord des adjectifs invariables du français », *Grazer Linguistische Studien* 86, 71-84.
- Goes J. (1999), *L'adjectif. Entre nom et verbe*. Duculot, Bruxelles.
- Goes J. (2008), « Les prédications secondes à prédicat adjectival », *Travaux de linguistique* 57, 23-41.
- Guimier C. & Oueslati L. (2006), « Le degré de figement des constructions 'verbe + adjectif invarié' », *Composition syntaxique et figement lexical*, J. François & S. Mejri (éds), Presses Universitaires de Caen, 17-37.
- Hummel M. (2018a), « La structure "verbe+adjectif". *Parler vrai, dire juste, faire simple* et compagnie », *Revue romane* 53-2, 261-296. En ligne : <https://benjamins.com/catalog/getpdf?uuid=54cd51b6-1da7-4f56-8d51-b61da73f5642&href=%2Ffro%2Ffro.53.2%2Ffro.16021.hum%2Ffro.16021.hum.pdf>
- Hummel M. (2018b), « Éléments d'une diachronie grammaticographique et normative de l'adverbe français accompagnés de trois méthodes pour mesurer l'effet du discours normatif sur l'usage », *Romanische Forschungen* 130, 3-35. En ligne : https://static.uni-graz.at/fileadmin/projekte/Adjective-Adverb/Hummel_Elements_D_une_Diachronie.pdf
- Hummel M. & S. Valera (eds) (2017), *Adjective Adverb Interfaces in Romance*. Benjamins, Amsterdam / Philadelphia. En ligne : <https://doi.org/10.1075/la.242>
- Hummel M. *Adjective-Adverb Interfaces in Romance. Open Access Database*. En ligne : <https://gams.uni-graz.at/context:aaif>
- Lauwers P. (2013), « Compléments de manière ou attributifs ? Les Sprép après *partir* », *Scolia* 27, 161-181.
- Ledgeway A. (2011), "Adverb agreement and split intransitivity: Evidence from Southern Italy". *Archivio glottologico italiano* 95, 31-66.
- Muller C. (2000), « Les constructions à adjectif attribut de l'objet, entre prédication seconde et complémentation verbale », *Langue française* 127, 21-35.
- Noailly M. (1994), « Adjectif adverbial et transitivité », *Cahiers de grammaire* 19, 103-114.
- Noailly M. (1999), *L'adjectif en français*, Paris, Ophrys.

Exploitation des descriptions du français parlé : l'exemple de « pendant »

Clara Cousinard, ATILF Université de Lorraine et CNRS

Les travaux de Claire Blanche-Benveniste, et notamment les études grammaticales sur le français parlé (Blanche-Benveniste *et al.* 1997, Blanche-Benveniste 2000), ont renouvelé les sciences du langage depuis une trentaine d'années, en permettant de formuler de nouvelles hypothèses sur le fonctionnement des langues. Nous avons réinvesti ces études dans le cadre de nos travaux, dont un des axes de recherche porte sur l'utilisation des descriptions du français parlé dans l'enseignement du Français Langue Etrangère (FLE), et plus précisément sur l'écart entre le français parlé présenté dans les documents de référence, et le français parlé authentique (Giroud, Surcouf 2016). Pour rendre compte de cet écart, nous nous sommes intéressée à la préposition temporelle « pendant ». Dans un premier temps, nous avons listé les utilisations de « pendant » tels que nous les trouvons dans des documents de référence. Nous avons ainsi analysé 8 manuels de FLE¹². Pour chacun des manuels, nous avons observé comment « pendant » était décrit. Plus précisément, nous avons cherché à voir si la leçon du manuel proposait une phase d'observation, c'est-à-dire un ensemble d'énoncés dans lesquels « pendant » est utilisé, amenant les apprenants à dégager les règles d'utilisation en contexte. Les apprenants sont ainsi amenés à faire de la grammaire inductive, qui nécessite un travail réflexif de leur part par l'observation d'un corpus d'énoncés, pour aboutir à la construction du savoir par eux-mêmes (Vincent *et al.* 2013). Notons que seuls deux manuels présentent des énoncés authentiques, donc bien que la phase d'observation soit présente dans 6 des 8 manuels, elle n'est constituée d'énoncés authentiques que dans 2 de ces 6 manuels. Ensuite, nous avons décrit les segments qui suivent « pendant » dans les énoncés présents dans les leçons des manuels, en dégagant deux tendances majeures. La première, celle des « pendant » suivis d'action ou de durée passée, donc nous avons coché d'une croix dans le tableau lorsque « pendant » était suivi d'un verbe au passé. La seconde, les « pendant » suivis de durées chiffrées. Nous entendons par « durées chiffrées » les segments contenant un chiffre, un nombre, ou une référence numérique, tels que « pendant cinq ans », « pendant plusieurs semaines », etc. Enfin, nous avons trié les manuels selon la présence ou l'absence du « pendant » de simultanéité, et nous avons fait la liste des définitions associées à « pendant ». Nous reprenons ces analyses dans les tableaux suivants :

¹² *Tendances A2* (2016) et *B1* (2018), et *Vite et bien A1/A2* (2018) publiés chez Clé International ; *Edito A2* (2016) et *B1* (2018), *L'atelier A2* (2019), et *Saison A2+* (2014) et *B1* (2018) chez Didier.

	Phase d'observation	« pendant » suivi de référence à action/durée passée	« pendant » suivi de référence à une durée chiffrée
<i>Tendances A2</i>	∅	X	X
<i>Tendances B1</i>	X (authentiques)	∅	∅
<i>Vite et bien A1/A2</i>	∅	X	X
<i>Edito A2</i>	X (1 seul énoncé par formule, factice)	X	X
<i>Edito B1</i>	X (1 seul énoncé par formule, factice)	X	∅
<i>L'atelier A2</i>	X (authentiques, mais ne contiennent pas « pendant »)	X	X
<i>Saison A2+</i>	X (1 seul énoncé par formule, factice)	X	X
<i>Saison B1</i>	X (texte factice + ne contient pas « pendant »)	X	X
	Présente : 6/8 → 75% (dont 2/8 authentiques → 25%) Absente : 2/8 → 25%	7/8 → 88%	6/8 → 75%

	« pendant » de simultanéité	Définition utilisation « pendant »
<i>Tendances A2</i>	∅	« exprimer la durée sans référence à un moment présent »
<i>Tendances B1</i>	X	« introduit une action qui se passe en même temps qu'une autre » + « relation de simultanéité »
<i>Vite et bien A1/A2</i>	∅	« pendant + verbe au passé composé (durée définie) »
<i>Edito A2</i>	∅	« pour indiquer une durée limitée » + « la durée de l'action est délimitée dans le temps »
<i>Edito B1</i>	X	« durée » + « simultanéité »
<i>L'atelier A2</i>	∅	« pour parler d'une action ponctuelle »
<i>Saison A2+</i>	∅	« la durée d'une action » + « durée totale d'une action »
<i>Saison B1</i>	X	« durée terminée » + « la durée d'une action » + « exprimer une simultanéité »
	3/8 → 38%	Associé à : - « durée » : 6/8 → 75% - « action » : 3/8 → 38% (≠ 100% parce que <i>Saison A2+</i> dans les deux cas)

Figures 1 et 2 : Récapitulatifs des analyses de 8 manuels de FLE

Deux constats majeurs peuvent être faits. Premièrement, « pendant » est présenté comme accompagné d'une durée chiffrée dans 75% des cas. Ensuite, « pendant » est très majoritairement présenté comme associé à une durée ou une action passée, exprimée par un verbe la plupart du temps au passé composé, dans 7 des 8 manuels, soit 88% des cas. Les exemples suivants illustrent ces deux remarques :

GRAMMAIRE

LES EXPRESSIONS DE TEMPS

DEPUIS + verbe au présent (*idée de continuité*)

- Depuis quand est-ce que vous travaillez en France?
- Je travaille en France depuis deux ans. (*et ça continue*)
- Depuis combien de temps est-ce qu'ils sont mariés ?
- Ils sont mariés depuis dix ans. (*et ils sont toujours mariés, ils n'ont pas divorcé*)

PENDANT + verbe au passé composé (*durée définie*)

- Pendant combien de temps est-ce qu'il a vécu en Italie ?
- Il a vécu en Italie pendant cinq ans. (*et c'est fini*)

Elle a été mariée pendant six ans. (*maintenant, elle est divorcée*)

Figure 3 : Extrait du manuel *Vite et bien A1/A2*, Clé International

LES INDICATEURS DE TEMPS		
Pour indiquer l'heure	à	<i>On se voit à onze heures.</i>
Pour indiquer une année, un mois	en	<i>C'est en 2014, en juillet</i>
Pour indiquer le point de départ d'une action qui dure encore	depuis	<i>J'habite à Montréal depuis dix ans.</i>
Pour situer un événement dans le passé	il y a	<i>Je suis arrivée ici il y a une heure.</i>
Pour indiquer une durée limitée	pendant	<i>Je t'ai attendu pendant dix minutes.</i>
Pour indiquer une durée prévue	pour	<i>Je pars en vacances pour deux semaines.</i>
Pour indiquer un point de départ	à partir de	<i>Je serai en vacances à partir de lundi prochain.</i>
Pour indiquer la limite temporelle d'une action	jusqu'à	<i>Le magasin est ouvert jusqu'à 19 h.</i>
Pour indiquer la succession d'actions	d'abord puis ensuite enfin	<i>D'abord, j'ai fait une licence d'anglais, puis j'ai fait un master 2 en linguistique, ensuite j'ai fait un stage de 6 mois, enfin j'ai trouvé un travail à l'école Louise Michel.</i>

Figure 4 : Extrait du manuel *Edito A2*, Didier

LES indicateurs de temps

👁 **Observez.**
J'ai travaillé dans cette entreprise **pendant** trois ans.
Il y a deux mois, j'ai obtenu mon diplôme.
J'étudie à l'université **depuis** 2 ans.
Quand j'étais à l'université...

🌀 **Réfléchissez.**
Nommez le temps de chaque phrase.
Indiquez si les phrases présentent une **action ponctuelle**, une **situation**, ou une **durée non terminée**.

✍ **Appliquez.**
Dans le document 4, relevez les verbes au passé composé et ceux à l'imparfait. Indiquez si ce sont des actions ou des situations.

Figure 5 : Extrait du manuel *Atelier A2*, Didier

Les indicateurs de temps

- **En** indique la **date** d'un événement avec un mois ou une année.
EXEMPLE : J'ai commencé **en septembre**.
- **Depuis** (+ date ou durée) indique une **action** ou une **situation qui continue** maintenant.
EXEMPLES : J'étudie le français **depuis novembre**.
Je travaille ici **depuis un an**.
- **Pendant** indique la **durée totale** d'une action.
EXEMPLE : J'ai travaillé **pendant deux ans** comme décoratrice.
- **Il y a** indique la **durée** entre une **action terminée** et maintenant.
EXEMPLE : Je me suis inscrit **il y a un an**.
- **Dans** indique la **durée** entre maintenant et une **action future**.
EXEMPLE : Je vais partir **dans une heure**.

→ Précis, p. 195

Figure 6 : Extrait du manuel *Saison A2+*, Didier

LES INDICATEURS DE TEMPS

► Unité 4 pp. 80 et 88, unité 5 pp. 101 et 109

► Exprimer le temps écoulé, une durée, une date

Exprimer quoi ?	Quoi ?	Exemples
une date	<i>en</i>	<i>En 2000, j'ai obtenu mon master.</i>
le temps écoulé	<i>il y a</i> <i>il y a... que</i> <i>cela / ça fait... que</i> <i>voilà... que</i>	<i>Il y a 10 ans, j'ai commencé à travailler.</i> <i>Il y a 10 ans que j'ai commencé à travailler.</i> <i>Ça fait 10 ans que j'ai commencé à travailler.</i>
le début d'une action	<i>depuis / depuis que [passé]</i> <i>dans [idée future]</i>	<i>Depuis qu'il est parti, je me sens revivre !</i> <i>Dans deux ans, je pars au Japon.</i>
la durée d'une action	<i>depuis / depuis que</i> <i>pendant</i>	<i>Je travaille dans cette entreprise depuis 10 ans.</i> <i>J'ai travaillé pendant 10 ans et ensuite, j'ai fait le tour du monde.</i>
la date de fin d'une action	<i>jusqu'à ce que + subj.</i>	<i>J'ai travaillé jusqu'à ce que j'en aie assez.</i>

Figure 7 : Extrait du manuel *Saison BI*, Didier

Pour compléter l'analyse des manuels, nous pouvons regarder comment « pendant » est présenté dans l'Encyclogram². Nous trouvons alors « pendant » dans la notice « L'aspect verbal » (Gosselin 2018), où la préposition est présentée comme introduisant un complément de durée, associée à l'exemple « pendant une heure ». Dans la suite de cette notice, nous trouvons « J'ai joué au piano pendant dix ans. ». La description de « pendant » laisse ici penser que cette préposition s'associe là encore à une durée passée, et chiffrée.

Nous avons voulu voir si les utilisations proposées dans les documents de références reflétaient ou non les utilisations effectives de « pendant », en comparant ces résultats à ceux de l'analyse d'un corpus de français parlé en interaction (André 2018). Pour cela, nous avons réalisé une analyse distributionnelle de « pendant » dans le corpus FLEURON³, outillé d'un concordancier (André 2016), et avons abouti à une liste de 34 occurrences⁴ (figure 6). Nous avons choisi le corpus FLEURON parce qu'il s'agit d'un dispositif d'apprentissage à destination d'apprenant de FLE, ce qui est au cœur de nos recherches.

² www.encyclogram.fr/

³ www.fleuron.atilf.fr

⁴ Le corpus FLEURON compte à ce jour un peu plus de 9h de données, et est régulièrement enrichi. Le nombre d'occurrences donné ici correspond au moment de sa consultation le 15/05/2020.

A gauche du mot	Mot recherché	A droite du mot
L2 : et donc	pendant	5 ans j'ai été responsable de l'espace langue L1 : hum hum
A : elle sera là pour vous aider hein	pendant	les quelques mois
E : mais j'ai déjà vécu en France	pendant	plusieurs années
A :	pendant	tout le séjour hein en France E : cité universitaire A : cité universitaire d'accord
E : d'accord A : alors à plus forte raison	pendant	vos séjour en France
A : ah oui mais il faut y aller	pendant	les permanences hein
E2 : c'est marqué que c'est uniquement	pendant	les périodes de cours
L1 :	pendant	le temps où
A : pour pouvoir travailler	pendant	les vacances viennent nous voir
A1 : mais vous devez savoir que	pendant	l'été il y a eu des changements
A1 : et je suis bien certaine que vous n'aurez pas de grosses difficultés pen- vo-	pendant	vos séjour ici
et ça vous permet en fait d'avoir tous les musées de Nancy et alentours gratuitement	pendant	un an
A : voilà c'est juste un peu petit peu plus cher mais après ça vous donne les mêmes droits sur tous les musées de Nancy et alentours musée du fer musée lorrain etc. gratuitement	pendant	un an E : d'accord
renseignez vous au niveau des de l'agenda culturel au CAN donc un samedi sur deux à 11h15 on peut vous faire une visite	pendant	une demi-heure
E1 :	pendant	les cours pendant les ex- les examens et
E1 : pendant les cours	pendant	les ex- les examens et
A : alors vous pouvez emprunter les livres	pendant	un mois et vous pouvez renouveler deux fois l'emprunt
1 : ok 2 : oui et	pendant	les vacances les hôpitaux sont toujours étaient presque tous fermés
A : et si vous venez pour déposer c'est	pendant	nos horaires d'ouverture qui sont du lundi au vendredi
A : d'accord sinon vous revenez nous voir	pendant	les horaires d'accord E : OK merci OK merci
E2 : euh j'ai commencé à trou- à chercher un directeur de thèse	pendant	mon master
E2 : voir les choses que j'avais déjà faites euh	pendant	mon master pendant la licence
E2 : voir les choses que j'avais déjà faites euh pendant mon master	pendant	la licence
avant quand tu étais au Brésil et	pendant	et après
euh qui pour eux euh ils peuvent pas trouver un logement et je connais des doctorants qui ont habité en dans un hôtel très cher	pendant	trois mois
chaque année nous on a	pendant	toute l'année environ 1850 doctorants
qui est capable avec ce travail avec ce qu'il a appris	pendant	ce travail de de faire beaucoup de choses
en suggérant que le docteur eh bien il a été amené	pendant	sa formation à faire d'autres choses
ben j'ai découvert Fleuron euh	pendant	euh pendant le Master
ben j'ai découvert Fleuron euh pendant euh	pendant	le Master
mais en fait euh j'ai ap- j'ai appris le français en Chine	pendant	quatre ans quand j'étais à l'université euh en Chine
et puis après ben j'ai enseigné euh j'ai enseigné le français	pendant	mon stage
donc ben j'espère que vous allez vous plaire	pendant	le six mois le semestre ou l'année que vous allez passer ici
vous venez quand vous voulez	pendant	les horaires d'ouverture

Figure 8 : Résultats de la recherche de « pendant » dans le corpus FLEURON via son concordancier

Le concordancier de FLEURON est un concordancier multimodal, ce qui permet à l'utilisateur, lorsqu'il clique sur une occurrence, d'accéder directement à la ressource correspondante. Plus précisément, la ressource s'ouvre automatiquement sur le moment où le terme recherché est énoncé.

Après l'analyse de chaque occurrence, nous avons pu conclure que « pendant » n'est associé à une durée chiffrée que dans 29% des cas (voir exemples suivants) contre 75% dans les documents de références. Les exemples 1 à 3 présentent des occurrences de « pendant » suivies d'une durée chiffrée, provenant du corpus FLEURON. L'exemple 1 est extrait d'une présentation d'un personnel d'une université à des étudiants étrangers :

L:	vous venez quand vous voulez	pendant	les horaires d'ouverture	:
		hum hum		

Exemple 1 : extrait de la ressource « Pôle européen » (FLEURON)

L'exemple 2 provient d'un dialogue entre une bibliothécaire et un étudiant qui demande des renseignements sur l'inscription et le fonctionnement de la bibliothèque municipale :

B:	et vous avez la possibilité d'emprunter trois livres	pendant	trois semaines
----	--	---------	----------------

Exemple 2 : extrait de la ressource « Bibliothèque municipale » (FLEURON)

L'exemple 3 est extrait d'un dialogue entre une étudiante et un employé du musée des beaux-arts de Nancy, pendant lequel l'employé explique à l'étudiante ce qu'il est possible de faire avec la « carte jeune » :

et ça vous permet en fait d'avoir tous les musées de Nancy et alentours gratuitement	pendant	un an
---	---------	-------

Exemple 3 : extrait de la ressource « Musée des beaux-arts et carte jeune » (FLEURON)

Le reste des occurrences de « pendant » accompagnent des durées nominalisées, et représentent 24 des 34 occurrences, soit 71% des cas. Nous illustrons cela par trois exemples également présents dans le corpus FLEURON. Le premier concerne une étudiante qui souhaite voyager en train à travers la France, et qui se rend à un guichet SNCF pour échanger avec une employée au sujet d'une carte jeune pour les 12-28 ans :

E: d'accord A: alors à plus forte raison	pendant	votre séjour en France
--	---------	------------------------

Exemple 4 : extrait de la ressource « Carte SNCF – 12-28 ans » (FLEURON)

L'exemple 5 est issu d'une présentation donnée par un médecin à des étudiants, concernant les différents services que propose la santé universitaire. Le médecin explique que les étudiants prennent parfois rendez-vous avec le service de santé pour obtenir un certificat médical et pouvoir travailler pendant les vacances :

A: pour pouvoir travailler	pendant	les vacances viennent nous voir
----------------------------	---------	---------------------------------

Exemple 5 : extrait de la ressource « Santé universitaire – partie 2 » (FLEURON)

L'exemple 6 provient d'un échange entre une étudiante et un agent d'accueil d'un centre de ressources en langues, autour du fonctionnement de ce dernier :

Exemple 6 : extrait de la ressource « Fonctionnement d'un centre de ressources en langue » (FLEURON)

Les verbes associés à « pendant » sont à 41% au passé, contre 88% dans les documents de référence, 47% au présent, et 12% au futur. Notons que ces deux derniers temps, le présent simple et le futur simple, ne sont pas représentés dans les manuels de FLE, alors que leur présence est significative dans les corpus oraux.

En conclusion de cette recherche, nous pouvons aisément affirmer que les usages de « pendant » dans les manuels de FLE ne reflètent pas ou peu les usages effectifs en français parlé. Ces résultats ne font que renforcer l'idée que les représentations de la langue française, notamment dans les documents de référence, reposent sur les données de la langue écrite ou de langue construite. Les résultats des recherches sur le français parlé trouvent donc leur utilité en didactique des langues, de façon à enseigner une langue authentique, qui existe réellement. L'exemple de « pendant » n'est pas isolé, nous remarquons cet écart entre les utilisations authentiques et celles proposées dans les manuels pour d'autres phénomènes comme l'enseignement du futur ou la formation d'énoncés interrogatifs (Ravazzolo *et al.* 2015). D'une perspective didactique se pose alors la question de la langue enseignée aux apprenants : quelle variété les documents de référence tentent-ils d'enseigner, s'ils ne prennent pas en compte les descriptions du français parlé ? Nous pouvons par ailleurs nous demander comment intégrer les corpus en didactique, pour enseigner une langue plus authentique que celle présentée jusqu'à maintenant ? De la même façon, nous pouvons également nous demander comment prendre en compte les résultats des recherches en français parlé, en didactique ? Pour cela, une piste de réponse pourrait être l'exploitation des corpus à des fins didactiques, en les intégrant à des séquences pédagogiques en suivant les principes de la démarche d'apprentissage sur corpus, ou *data-driven learning* (Johns 1991). Les possibilités sont multiples, Breyer (2006 : 162) allant jusqu'à dire que les utilisations pédagogiques des corpus sont « limited only by the imagination of the user ».

Bibliographie

André, V. (2016). FLEURON : Français Langue Etrangère Universitaire, Ressources et Outils Numériques. Origine, démarches et perspectives. Mélanges CRAPEL, 37, 24.

André, V. (2018). Nouvelles actions didactiques : faire de la sociolinguistique de corpus pour enseigner et apprendre à interagir en Français Langue Etrangère. Action didactique, 1, 71-88.

Blanche-Benveniste, C. (2000). Approches de la langue parlée en français. Paris : Ophrys.

Blanche-Benveniste, C., Bilger, M., Rouget, C., Van Den Eynde, K. (1997). Le français parlé : études grammaticales. Paris : CNRS.

Breyer, Y. (2006). My Concordancer: Tailor-made software for language learners and teachers. In S. Braun, K. Kohn & J. Mukherjee (eds), *Corpus Technology and Language Pedagogy: New resources, new tools, new methods*. Frankfurt: Peter Lang, p. 157-176.

Giroud, A., Surcouf, C. (2016). De « Pierre, combien de membres avec-vous ? » à « Nous nous appelons Marc et Christian » : réflexions autour de l'authenticité dans les documents oraux des manuels de FLE pour débutants. *SHS Web of Conferences*, 27, 07017.

Gosselin, L. (2018). « L'aspect verbal », in Encyclopédie grammaticale du français, en ligne : encyclogram.fr

Johns, T. (1991). Should you be persuaded: Two examples of data-driven learning. In T. Johns & P. King (dir.), *Classroom Concordancing. English Language Research Journal*, 4, 1-16.

Ravazzolo, E., Traverso, V., Jouin, E. (2015). Interactions, dialogues, conversations : l'oral en français langue étrangère. Vanves : Hachette.

Vincent, F., Dezutter, O., Lefrançois, P. (2013). Enseigner la grammaire selon une approche inductive ou déductive ? *Québec français*, 170, 93-94.

À propos de la typologie de la relativisation

Denis Creissels, Université Lumière (Lyon 2)

1 Introduction

.

Comme l'ont bien montré les différents auteurs qui ont abordé cette question, et notamment Claire Blanche-Benveniste, les données du français parlé obligent à poser la question des fondements même de la description de la relativisation, à commencer par la définition de ce qu'est (ou n'est pas) une proposition relative. Les questions que doivent se poser les linguistes qui s'intéressent à la relativisation en français parlé rejoignent ainsi très largement celles qui se posent aux typologues ayant pour objectif de rendre compte de la diversité qui existe à travers les langues dans le domaine des constructions où on peut reconnaître une proposition subordonnée qui exprime ou évoque une propriété et un terme nominal caractérisé par la propriété en question. Dans cette présentation, sans prétendre apporter des réponses définitives, j'ai pour objectif de soulever un certain nombre de questions qui me semblent cruciales dans cette perspective, à partir notamment d'un examen critique de l'approche de la typologie des propositions adnominales (noun-modifying clauses) développée par Yoshiko Matsumoto et Bernard Comrie.

2 Propositions relatives et propositions adnominales

.

Il convient tout d'abord de rappeler que les notions de proposition adnominale et de proposition relative ne sont pas équivalentes.

En ce qui concerne une notion générale de proposition relative, le problème est de dégager ce qu'ont en commun les propositions relatives telles qu'elles ont été traditionnellement décrites dans un certain nombre de langues, et de proposer à partir de là une définition faisant le plus possible abstraction de tout ce qui pourrait rendre la notion de relativisation inapplicable à une proportion plus ou moins importante des langues du monde. Dans cette perspective, la définition proposée par Andrews (2007: 206) constitue un point de départ utile. Selon cette définition

“A relative clause (RC) is a subordinate clause which delimits the reference of an NP by specifying the role of the referent of that NP in the situation described by the RC.”

Maintenant, si une proposition adnominale est définie comme une proposition qui modifie un nom et fait partie du groupe nominal projeté par le nom qu'elle modifie, il est évident qu'il existe (mais de manière très variable selon les langues) des propositions communément considérées comme relatives (et qui répondent à la définition citée ci-dessus) qui ne sont pas adnominales : relatives libres (1), relatives 'internes', qui incluent syntaxiquement le nom qu'elles modifient (2), relatives adjointes à gauche (3) et relatives adjointes à droite (4).

(1) anglais

The dog ate [what the cat left in its bowl].
le chien a.mangé ce.que le chat a.laissé dans so bol
n

‘Le chien a mangé ce que le chat a laissé dans son bol.’

(2) navajo (athabascan ; Platero 1974)

)

[Ashkii at'ééd yiyiltsá-néé] yálti'.
garçon fille Io:3SG.IPF.Is:3SG.voir-REL.PAS IPF.Is:3SG.parler
‘Le garçon qui a vu la fille est en train de parler.’ ou ‘La fille que le garçon
a vue est en train de parler.’ (ambigu)

(3) mandinka (mandé)

[Kèè míŋ nãa-tá bĩ], ñ máŋ wõo lôŋ.
homme. lequel venir-ACP aujourd'hui 1SG ACP.NEG DEM connaître
D

‘Je ne connais pas l'homme qui est venu aujourd'hui.’

lit. ‘Lequel homme est venu aujourd'hui, je ne connais pas celui-là.’

(4) hongrois (ouralien)

)

Azt a könyvet olvasom, [amelyet most vettem].
DEM.ACC D livre.ACC lire.Is:1SG.Io:3.D lequel.ACC maintenant acheter.Is:1SG.Io:3.D
‘Je lis le livre que je viens d'acheter.’
lit. ‘Ce livre je lis, lequel je viens d'acheter.’

Inversement, il est communément admis qu'une proposition adnominale peut ne pas relever de la relativisation. Par exemple, il y a un assez large consensus (qui n'est toutefois pas total) pour considérer que dans *l'idée [que tu puisses me quitter]*, la proposition adnominale ne doit pas être considérée comme une relative.

Des difficultés surgissent toutefois lorsqu'on s'efforce de préciser la distinction entre propositions adnominales relevant de la relativisation et propositions adnominales demandant un autre type d'analyse. L'analyse d'une proposition adnominale comme relative ne va de soi que lorsque le nom-tête est repris par un pronom (pronom ordinaire ou pronom relatif) qui par sa morphologie et/ou sa position informe sur le rôle du référent du nom-tête dans l'événement décrit par la proposition adnominale.

3 Propositions adnominales à rôle implicite

.

La question qui se pose est de savoir comment situer par rapport à la notion de relativisation des propositions adnominales qui caractérisent le référent de leur tête par son rôle dans l'événement qu'elles décrivent tout en n'incluant aucune référence explicite à ce rôle (désignées dans ce qui suit comme PROPOSITIONS ADNOMINALES À RÔLE IMPLICITE pour éviter de se prononcer prématurément sur la façon dont il convient de les analyser). Par exemple, (5a) et (5b) ne diffèrent pas en ce qui concerne la relation sémantique entre le nom-tête et la proposition adnominale : dans les deux cas, on comprend que les pompiers ont été appelés pour éteindre le feu. La différence est qu'en (5a), *pour lequel* encode explicitement le rôle du nom-tête dans l'événement décrit comme *il a fallu appeler les pompiers*, alors que ce rôle reste entièrement implicite en (5b), où *que* est un pur marqueur de subordination qu'il n'y aurait aucune justification à analyser comme 'pronom relatif'.

(5) français (standard (a), non standard (b))

(5a) *un feu [pour lequel il a fallu appeler les pompiers]*

(5b) *un feu qu'[il a fallu appeler les pompiers]*

En français, il est difficile de discuter l'analyse de telles constructions en faisant abstraction du fait que leur acceptabilité varie selon les registres, et que la grammaire traditionnelle ne les reconnaît que de manière très limitée. Il importe donc de préciser immédiatement qu'il existe des langues où les propositions adnominales à rôle implicite sont très productives quel que soit le registre considéré, et où leur usage ne constitue en aucun cas un marqueur sociolinguistique.

4 La typologie des propositions adnominales selon Matsumoto & al. (2017)

.

En linguistique générale, la question de la typologie des propositions adnominales a été essentiellement discutée à partir de données du japonais, langue où les propositions adnominales à rôle implicite sont particulièrement productives, et se prêtent tous registres confondus à l'expression d'une variété de relations allant au-delà de tout ce qui a pu être observé dans d'autres langues pour des constructions semblables.

Matsumoto & al. (2017) proposent trois catégories fonctionnelles de propositions adnominales (noun-modifying clause constructions, en abrégé NMCC). Ce classement fonctionnel ne préjuge en rien de la répartition des propositions adnominales en types formels dans les langues particulières :

- (a) propositions adnominales ARGUMENTALES (argument NMCC), dont la propriété caractéristique est que « le nom-tête est coréférent avec un argument de la proposition adnominale » ;
- (b) propositions adnominales ADJONCTIVES (adjunct NMCC), dont la propriété caractéristique est que « le nom-tête est coréférent avec un adjectif de la proposition adnominale » ;

- (c) propositions adnominales ÉTENDUES (extended NMCC), qui impliquent « un autre type de relation entre le nom-tête et la proposition adnominale ».

Les deux premières catégories sont illustrées en (6) par des exemples japonais, langue dont la particularité est d'avoir une même construction disponible pour la totalité de ces types fonctionnels. Comme le montre cet exemple, les catégories (a) et (b) correspondent à des constructions dont l'analyse en termes de relativisation fait l'objet d'un consensus, même en l'absence de tout élément morphologique explicitant le rôle du référent du nom-tête dans l'événement décrit par la proposition adnominale. En effet, il n'y a pas de difficulté particulière à identifier ce rôle à un rôle non exprimé: sujet/agent en (6a), objet/patient en (6b), circonstanciel de lieu en (6c).

(6) japonais

(6a) [*hon o katta*] *gakusei* (proposition adnominale argumentale)
)

livre ACC acheter.PAS étudiant
'l'étudiant qui a acheté un livre'

(6b) [*gakusei ga katta*] *hon* (proposition adnominale argumentale)
étudiant SUJ acheter.PAS livre
'le livre que l'étudiant a acheté'

(6c) [*gakusei ga hon o katta*] *mise* (proposition adnominale
adjonctive)
étudiant SUJ livr ACC acheter.PAS boutique
e
'la boutique où l'étudiant a acheté un livre'

Matsumoto & al. (2017) reconnaissent ensuite trois sous-catégories de propositions adnominales 'étendues', qui sont illustrées en (7) :

- (c1) propositions adnominales étendues CADRATIVES (frame NMCC), dans lesquelles « le nom-tête instancie un élément du cadre de la proposition non couvert par les propositions argumentales ou adjonctives » (7a) ;
- (c2) propositions adnominales COMPLÉTIVES (content noun NMCC), dans lesquelles « la proposition adnominale exprime le contenu du nom-tête » (7b) ;
- (c3) propositions adnominales RELATIONNELLES/PERCEPTIVES (relational/perceptional noun NMCC), dans lesquelles « un nom-tête exprimant relation ou perception est caractérisé en relation avec l'événement ou état décrit par la proposition adnominale » (7c).

(7) japonais (Matsumoto & al. 2017)

(7a) [*atama no yokunaru*] *hon* (proposition adnominale cadrative)
tête SUJ s'améliorer livr

e

lit. ‘un livre que la tête s’améliore’ > ‘un livre qui rend plus intelligent’

(7b) [*hon o katta*] *sirase* (proposition adnominale complétive)

livre ACC acheter.PAS nouvelle

lit. ‘la nouvelle qu’on a acheté le livre’ > ‘la nouvelle de l’achat du livre’

(7c) [*hon o katta*] *oturi* (proposition adnominale relationnelle)

livre ACC acheter.PAS monnai

e

lit. ‘la monnaie qu’on a acheté le livre’ > ‘la monnaie de l’achat du livre’

Crucialement, cette classification, qui se veut pourtant fonctionnelle, ne comporte pas une catégorie qui réunirait toutes les propositions adnominales ayant en commun de caractériser le référent du nom-tête par son rôle dans la situation décrite, et regroupe au contraire une partie des propositions adnominales qui caractérisent le référent du nom tête en termes de participation à une situation (les ‘cadratives’) avec les propositions adnominales exprimant d’autres types de relations sémantiques avec le nom-tête.

Dans ce qui suit, la discussion portera donc essentiellement sur la question des propositions adnominales que Matsumoto & al. désignent comme ‘cadratives’, qui ont en commun avec les propositions adnominales ‘argumentales’ et ‘adjonctives’ la propriété de caractériser le référent du nom-tête par sa participation à la situation qu’elles décrivent, mais sont censées s’en distinguer par l’absence de relation de coréférence entre le nom-tête et un terme de la proposition adnominale.

5 Propositions adnominales ‘adjonctives’ et ‘cadratives’

.

De manière générale, la classification des propositions adnominales proposée Matsumoto & al. (2017) pose un certain nombre de problèmes, dus notamment à une certaine imprécision dans la formulation des définitions, qu’il n’est pas possible de passer en revue dans le cadre de cette présentation. Dans ce qui suit, je me limiterai à ce qui me paraît être le point le plus faible de cette classification, à savoir le fait qu’elle implique l’existence d’une limite tranchée entre deux types de propositions adnominales qui ont en commun de caractériser le référent du nom-tête par référence à son rôle dans la situation que décrit ou évoque la proposition adnominale : d’une part les propositions adnominales ‘argumentales’ ou ‘adjonctives’, qui sont présentées comme se prêtant à une analyse relevant de la notion de relativisation telle que couramment conçue, selon laquelle la proposition adnominale inclut une fonction syntaxique exprimant le rôle du référent du nom tête dans la situation décrite, et d’autre part les propositions adnominales ‘cadratives’, qui sont présentées comme ne se prêtant pas à une telle analyse.

Le problème se pose pour toutes les langues où l’usage de propositions adnominales à rôle implicite ne se limite pas à la relativisation de rôles argumentaux.

Comme cela a été mentionné ci-dessus, le japonais a la particularité d’exprimer un éventail particulièrement large de relations entre une situation et ses éléments constitutifs au moyen de propositions adnominales à rôle implicite. Sans avoir forcément la même productivité qu’en japonais, ce phénomène est largement attesté dans les langues du monde, comme l’illustrent les exemples (8) à (11).

- (8 mungbam (bantoïde ; Lovegren & Voll 2017)
)

ì-dʒī ì-nī [bũ gbà nà kǎ-tǎ kǎ]
 cl5-bruit cl5-REL cl2 couper.INACP être.INACP cl12-tree cl12.D
 ‘le bruit qu’ils font en coupant l’arbre’
 litt. ‘le bruit qu’[ils coupent l’arbre]’

- (9) hinuq (nakh-daghestanais ; Comrie & al. 2017: 134)

[de ɣwero t’otto goła yi]
 1SG.ERG vache traire.CVBI être.PTCP lait
 ‘le lait que j’ai eu en trayant la vache’
 litt. le lait que [j’ai trait la vache]’

- (10) chinois mandarin (Zhang 2008 : 1004)

- (10a *[Lulu tan gangqin] de shengyin*
)

Lulu jouer piano JCT son
 ‘le son du piano joué par Lulu’
 litt. ‘le son que [Lulu joue du piano]’

- (10b *Wo wendao-le [[mama chao cai] de weidao].*
)

1SG sentir-ACP maman frire légume JCT odeur
 ‘J’ai senti l’odeur de légumes que maman faisait frire.’
 litt. ‘J’ai senti [l’odeur que [maman faisait frire des légumes]].’

- (10c *Wo jisuanchu-le [[qiche zhengzai xingshi de sudu].*
)

1SG calculer-ACP voiture PROG courir JCT vitesse

e
 ‘J’ai calculé la vitesse à laquelle la voiture roulait.’
 litt. ‘J’ai calculé [la vitesse que [la voiture roulait]].’

(11) jóola fóoñi (atlantique)

(11a) *hamaa y-an [e-purumi sindeey]*
y

projet.D clE-REL Is:clA- foyer.D
quitter.ACP

‘le projet qui l’a conduit à quitter son foyer’
litt. ‘le projet qu’[il a quitté le foyer]’

(11b) *bɛ-tɛsmab b-umbeem b-an [ɪ-jaakalm]*

)

bouche.D clB-POSS.1SG clB-REL Is:1SG-se.tracasser.ACP.REL

‘ma bouche avec laquelle j’ai un problème’
litt. ‘ma bouche que [j’ai un problème]’

(11c) *bɛrɔkab b-an ɛmanj’ɔɔn [nɪcɛɛ lookuŋ lɛt e-riiŋ]*

)

travail.D clB-REL tu.sais.que⁰ parfois semaine FUT.NEG Is:clA-arriver

dɛɛ sindeey]

LOC foyer.D

‘un travail qui l’oblige parfois à rester absent de la maison une semaine entière’
litt. ‘un travail que [parfois il ne rentre pas à la maison pendant une semaine]’

De telles constructions abondent dans les registres informels du français, et il est évident que la grammaire traditionnelle, qui les ignore purement et simplement, ne fournit pas les outils d’analyse appropriés. Des constructions identiques ou semblables à celles présentées en (12) sont mentionnées dans tous les ouvrages classiques sur le français parlé, cf. notamment Deulofeu (1981, 1988), Gadet (1992, 1997), Blanche-Benveniste (2000), Ball (2000).

(12) français

(12a) *le jour que [Jean est venu]*

(12b) *un feu qu’[il a fallu appeler les pompiers]*

)

(12c) *un chien que [je suis jamais seul]*

(12d) *les choses qu’[il m’a parlé]*

)

(12e) *ma petite maison que [je tenais tant]*

(12f) *le café qu’[on s’était donné rendez-vous]*

⁰ En Jóola Fóoñi, *ɛmanj’ɔɔnɛ*, litt. ‘tu sais (que)’ a été grammaticalisé comme marque des propositions relatives exprimant une propriété non-épisodique du référent du nom-tête.

(12g *le couteau qu'[on coupe le pain]*
)

(12h *la façon que [tu me parles]*
)

Le problème qui se pose est donc tout d'abord celui des critères selon lesquels on peut décider que de telles propositions adnominales doivent être classées comme adjonctives ou cadratives selon la typologie de Matsumoto & al.

La distinction entre propositions adnominales argumentales ou adjonctives d'une part, et propositions adnominales cadratives d'autre part, est en fait reprise de la tradition japonaise représentée notamment par Teramura (1975-1978), qui caractérise cette distinction en termes de RELATION INTERNE vs. RELATION EXTERNE. Le critère proposé est que la proposition adnominale puisse ou non être mise en correspondance avec une proposition indépendante dont elle diffère seulement par l'absence d'une expression spécifiant le rôle assumé par le référent du nom-tête dans la situation décrite par la proposition adnominale.

Or en reprenant les exemples cités en (12), il est facile de voir que tous se prêtent à cette manipulation, et méritent donc d'être reconnus comme propositions adnominales adjonctives selon la classification de Matsumoto et al.

(12') français

(12'a) *le jour que [Jean est venu ~~ce jour-là~~]*

(12'b) *un feu qu'[il a fallu appeler les pompiers ~~pour (éteindre) ce feu~~]*

(12'c) *un chien que [je suis jamais seul ~~avec / grâce à ce chien~~]*

(12'd) *les choses qu'[il m'a parlé ~~de ces choses~~]*

(12'e) *ma petite maison que [je tenais tant ~~à cette maison~~]*

(12'f) *le café qu'[on s'était donné rendez-vous ~~dans cette maison~~]*

(12'g) *le couteau qu'[on coupe le pain ~~avec ce couteau~~]*

(12'h) *la façon que [tu me parles ~~de cette façon~~]*

Ce n'est donc pas en français qu'on peut espérer trouver de bons exemples d'adnominales cadratives. Mais si on examine de près les exemples plus exotiques qui ont été cités dans la littérature comme illustrant la notion d'adnominale cadrative, notamment pour les langues d'Asie orientale où ces constructions sont particulièrement productives, on s'aperçoit qu'ils ne sont pas fondamentalement différents. L'équivalent français littéral de quelques-uns de ces exemples (signalé par le symbole '°' pour éviter toute confusion avec des exemples français attestés) est donné en (13). Il s'avère que dans tous les cas, exactement comme en (12), la proposition adnominale peut s'obtenir à partir d'une proposition indépendante en effaçant simplement une expression qui explicite le rôle sémantique par lequel la proposition adnominale caractérise le référent du nom-tête.

- (13a) °la vitesse que [la voiture roulait à ~~cette~~ vitesse]
- (13b) °un livre qu'[on devient plus intelligent ~~grâce à ce~~ livre]
- (13c) °une maladie que [beaucoup de gens sont morts ~~de cette~~ maladie]
- (13d) °l'odeur que [maman fait frire des légumes ~~avec comme résultat cette~~ odeur]
- (13e) °le bruit qu'[on coupe un arbre ~~avec comme résultat ce~~ bruit]
- (13f) °le lait que [j'ai trait la vache ~~pour (avoir) ce~~ lait]
- (13g) °l'argent que [j'ai fait une traduction ~~en échange de cet~~ argent]
- (13h) °des bonbons qu'[on ne grossit pas ~~à cause de ces~~ bonbons]

Il est certain que d'un cas à l'autre (et d'une langue à l'autre) l'expression qu'il faut ajouter à la proposition adnominale à rôle implicite pour obtenir une proposition indépendante spécifiant le rôle du nom-tête peut être plus ou moins complexe. Mais il y a un continuum entre des cas où une simple préposition (ou postposition) totalement grammaticalisée suffit à préciser le rôle du nom-tête et des cas nécessitant l'intervention d'une locution prépositionnelle (ou postpositionnelle) plus ou moins grammaticalisée, et il ne serait pas raisonnable de prétendre fixer dans ce continuum une frontière qui séparerait les propositions adnominales à rôle implicite en deux types nettement distincts.

6 Pour une conception large de la notion de relativisation

.

Il n'y a en réalité aucune distinction nette entre les propositions adnominales à rôle implicite couramment analysées comme relatives adjonctives et les propositions adnominales 'cadratives' de Matsumoto & al., dont la seule véritable particularité est d'être ignorées par les grammaires traditionnelles des langues les mieux connues de linguistes. A partir de là, la seule décision cohérente consiste à ne pas retenir cette distinction.

Il est utile à ce stade de rappeler la définition de la notion générale de proposition relative citée dans l'introduction, selon laquelle

“A relative clause (RC) is a subordinate clause which delimits the reference of an NP by specifying the role of the referent of that NP IN THE SITUATION DESCRIBED BY THE RC.”
Andrews (2007: 206)

En fait l'interprétation de cette définition repose crucialement sur la façon dont on comprend 'préciser le rôle du référent'.

Si 'préciser le rôle du référent' est compris au sens restreint d'encoder explicitement le rôle du référent, alors les propositions adnominales à rôle implicite ne peuvent pas être reconnues comme relatives, quel que soit le rôle qu'elles attribuent à leur tête. Or la suite de

l'article d'Andrews ne laisse aucun doute sur le fait que ce n'est pas ainsi qu'il faut interpréter cette définition, puisqu'il cite l'absence de toute expression du rôle relativisé comme une stratégie possible de relativisation.

Maintenant, si 'préciser le rôle du référent' est compris comme 'fournir d'une manière ou d'une autre des informations permettant de reconstituer le rôle du référent dans l'événement dénoté par la proposition relative', il n'y a aucune raison de ne pas considérer les constructions citées en (12) et (13) comme relevant de la relativisation. Simplement, en l'absence de tout codage explicite du rôle du référent du nom-tête, et si la langue n'impose pas de fortes restrictions sur les rôles susceptibles d'un tel traitement, le rôle relativisé doit être identifié parmi les rôles qui ne sont pas exprimés dans la proposition adnominale en tenant compte de la compatibilité entre les rôles qui restent vacants et le sens lexical du nom-tête.

Si on accepte cette conception large de la relativisation, on doit bien sûr distinguer deux types de relatives adnominales selon que la construction inclut ou non référence explicite au rôle relativisé. On peut d'ailleurs estimer que dans une typologie des propositions relatives, la distinction la plus fondamentale est entre présence ou absence de codage explicite du rôle relativisé. On pourrait parler de RELATIVISATION EXPLICITE vs. RELATIVISATION IMPLICITE. Ce qui semble par contre difficile à justifier, c'est de considérer (comme le proposent Matsumoto & al.) que certaines seulement des constructions citées en (12) et (13) seraient des relatives adjectives, les autres relevant d'une notion de 'propositions adnominales cadratives' qu'il est en fait impossible de délimiter de manière cohérente.

Toute définition étant arbitraire, on pourrait aussi imaginer de formuler la définition de la notion de relativisation de façon à la limiter à ce que j'ai proposé d'appeler relativisation explicite. Une telle restriction de la notion de relativisation aurait toutefois deux inconvénients sérieux : d'une part elle exclurait de la relativisation un certain nombre de constructions que les linguistes ont pris l'habitude de considérer comme des cas de relativisation (à commencer par les relatives en *that* de l'anglais), et d'autre part, compte tenu des affinités évidentes entre les constructions relevant d'une définition étroite de la relativisation et les propositions nominales à rôle implicite, il serait indispensable de créer un nouveau terme qui les réunisse.

7 Quelques généralisations

.

Si on accepte une conception large de la notion de relativisation qui englobe la totalité des propositions adnominales à rôle implicite, on peut formuler les deux généralisations suivantes :

- (1) la relativisation implicite n'est possible que pour les relatives adnominales ; pour les autres types possibles (relatives libres, internes, adjectives à gauche et adjectives à droite), seule la relativisation explicite est attestée ;
- (2) dans les langues du monde, la relativisation implicite est plus courante lorsque le rôle relativisé correspond à un argument d'un terme de la proposition relative que lorsqu'il est encodé comme adjectif.

La généralisation (2) est cohérente avec le fait que dans les langues du monde, il est courant que les arguments nuls se prêtent de façon plus ou moins systématique à une interprétation anaphorique, alors que rien de semblable ne s'observe pour les adjoints.

Ensuite, les données disponibles semblent indiquer que les différents rôles sémantiques susceptibles d'être exprimés par des adjoints ne se prêtent pas avec la même facilité à la relativisation implicite. La relativisation implicite semble particulièrement fréquente pour les rôles de localisation temporelle (*le jour qu'il est venu*) ou de localisation spatiale (*le café qu'on s'était donné rendez-vous*). Inversement, la relativisation implicite du rôle d'effet collatéral (*°le bruit qu'on coupe un arbre*) semble beaucoup moins courante.

Beaucoup plus de données sur une plus grande variété de langues seraient nécessaires avant de poursuivre dans cette voie, mais il serait certainement intéressant de reprendre la hiérarchie d'accessibilité à la relativisation en traitant séparément la relativisation implicite et la relativisation explicite et en distinguant les différents rôles sémantiques susceptibles d'être exprimés par des adjoints.

8 Conclusion

.

Dans cette présentation, je me suis efforcé d'apporter un éclairage typologique sur les propositions adnominales à rôle implicite. Ce type de construction, largement attesté en français parlé, peut être jugé problématique en ce qui concerne sa relation au concept de relativisation.

Après avoir fait observer que pas mal de langues à travers le monde exploitent la possibilité d'avoir recours à ce type de construction pour attribuer au nom-tête un large éventail de rôles dans la situation que décrit la proposition adnominale, j'ai essayé de montrer que, contrairement à la proposition de Matsumoto & al., il n'y a pas lieu de séparer les propositions adnominales à rôle implicite en deux catégories dont une seule relèverait de la relativisation.

Enfin, j'ai fait remarquer que les définitions de la relativisation couramment proposées dans la littérature typologique permettent en fait de reconnaître comme relatives la totalité des propositions adnominales à rôle implicite. J'en ai conclu que la position la plus cohérente et la plus pratique consiste à admettre une notion large de relativisation, tout en proposant de considérer que la distinction la plus fondamentale dans la typologie de la relativisation est la distinction entre relativisation explicite (avec codage explicite du rôle relativisé) et relativisation implicite (sans codage explicite du rôle relativisé).

Abréviations

ACC : accusatif, ACP : accompli, cl : classe nominale, CVBI : verbe imparfait, D : défini, DEM : démonstratif, ERG : ergatif, Io : indice d'objet, FUT : futur, INACP : inaccompli, IPF : imparfait, Is : indice de sujet, JCT : joncteur, LOC : locatif, NEG : négatif, PAS : passé, POSS : possessif, PROG : progressif, PTCP : participe, REL : marqueur de relativisation, SG : singulier, SUJ : sujet

Références

- Andrews. 2007. Relative clauses. In Timothy Shopen (ed.), *Language typology and syntactic description. Volume II: Complex constructions (second edition)*. Cambridge University Press. 206-236.
- Ball, Rodney. 2000. *Colloquial French Grammar: A Practical Guide*, Oxford: Blackwell.
- Blanche-Benveniste, Claire. 2000. *Approches de la langue parlée en français*, Paris: Ophrys.
- Comrie, Bernard, Diana Forker & Zaira Khalilova. 2017. General noun-modifying clause constructions in Hinuq and Bezhta, with a note on other Daghestanian languages. In Yoshiko Matsumoto, Bernard Comrie & Peter Sells (eds.) *Noun-modifying clause constructions in languages of Eurasia*. John Benjamins. 121-146.
- Deulofeu, José. 1981. Perspective linguistique et perspective sociolinguistique dans l'étude des relatives en français. *Recherches sur le français parlé* 3.
- Deulofeu, José. 1988. Syntaxe de *que* en français parlé et le problème de la subordination. *Recherches sur le français parlé* 8.
- Gadet, Françoise. 1992. *Le français populaire*. Paris: PUF.
- Gadet, Françoise. 1997. *Le français ordinaire*, 2ème édition. Paris: Armand Colin.

- Lovegren, Jesse & Rebecca Voll. 2017. Relative clause construction in two Yembe-Kimbi languages. In Gratien Gualbert Atindogbé & Rebecca Grollemund (eds.), *Relative clauses in Cameroonian languages*. De Gruyter Mouton. 169-208.
- Matsumoto, Yoshiko, Bernard Comrie & Peter Sells. 2017b. Noun-modifying clause constructions in languages of Eurasia: Rethinking theoretical and geographical boundaries. In Yoshiko Matsumoto, Bernard Comrie & Peter Sells (eds.) *Noun-modifying clause constructions in languages of Eurasia*. John Benjamins. 3-21.
- Teramura, Hideo. 1975-1978. Rentai shushoku no shintakusu to imi [Syntax and semantics of nominal modifiers]. In *Nihongo Nihonbunka* 4-7. 71-119, 29-78, 1-35, 1-24.
- Zhang, Niina. 2008. Gapless relative clauses as clausal licensers of relational nouns. *Language and Linguistics* 9(4). 1003-1026.

L’alternance actif/passif en français : un modèle statistique sur corpus écrits et oraux

Yanis da Cunha, Laboratoire de Linguistique Formelle, Université de Paris
Anne Abeillé, Laboratoire de Linguistique Formelle, Université de Paris

1. Introduction

La fréquence de la construction passive varie selon les genres de texte, notamment entre écrit et oral (Roland et al. (2007) pour l’anglais, Poiret et Liu (2020) pour le français). Les facteurs qui conditionnent l’utilisation de cette construction varient également à travers les langues (Bresnan et al., 2001). En anglais, ces facteurs ont été étudiés quantitativement à l’oral (Weiner & Labov, 1983 ; Estival, 1985 ; Dingare, 2001) et à l’écrit (Henley et al., 1995 ; Hundt et al., 2018). L’effet de facteurs variés a ainsi été notée : animéité, définitude, personne et longueur des arguments, classe sémantique du verbe, amorçage.

En français de telles études quantitatives n’ont pas encore été menées. Des études existent pour le français écrit littéraire (Attal, 1985) et journalistique (McLaughlin, 2010), et pour le français oral (Blanche-Benveniste, 1989 ; Hamma et al., 2017 ; Druetta, 2020), mais celles-ci ne comparent pas les constructions actives et passives de façon systématique, ni n’emploient de méthode quantitative.

Le présent travail vise donc à étudier l’utilisation de la construction passive en français à l’aide d’outils statistiques, en adoptant une approche probabiliste, multifactorielle et comparative entre écrit et oral.

2. Méthodologie

2.1 Sélection des corpus

Nous utilisons quatre corpus, tous annotés en dépendances :

Le French TreeBank (FTB) qui est un corpus d’écrit journalistique composé d’articles du journal Le Monde de 1990 à 1993 (Abeillé et al., 2019)

Le Corpus de Français Parlé Parisien (CFPP), le Corpus de Référence de Français Parlé (CRFP), et le C-Oral-Rom, qui sont des corpus oraux principalement conversationnels couvrant conjointement la période 1994-2012 (Benzitoun et al., 2016)

2.2 Échantillonnage

À l’aide des annotations syntaxiques des corpus, nous extrayons un total de 900 phrases des corpus, 500 du FTB et 400 des corpus oraux (désormais, CO). L’échantillon extrait contient autant d’actifs que de passifs et a été filtré pour ne contenir que des verbes transitifs passivables arguments nominaux ou pronominaux, et fléchis à un temps composé (afin d’éviter l’ambiguïté entre passif adjectival et verbal aux temps non composés).

Parmi les passifs, nous avons extrait autant de passifs courts (1), sans complément d’agent, que de passifs longs (2), avec complément d’agent en par (et non de). En correspondance à l’actif,

nous avons extrait autant d'actifs à sujet clitique (3) que d'actifs à sujet Syntagme Nominal (sujet SN) (4).

- (1) a. Saint-Antoine c'est là où ma mère a été opérée (CFPP)
b. 42000 permis poids lourds ont été délivrés en 1991. (C. de Chenay, 08/07/1992, FTB)
- (2) a. J'avais été enthousiasmé par le premier d'ouverture (REN, fmedsp02, C-Oral-Rom)
b. Un agriculteur a été grièvement blessé à la main par une grenade lacrymogène (02/07/1992, FTB)
- (3) a. Donc il m'a expliqué il m'a donné des petites ficelles (L1, PRI-PSO-2, CRFP)
b. Elle a pris des initiatives nombreuses et intervient dans les domaines les plus divers (A. Lebaube, 02/06/1993, FTB)
- (4) a. Et ma mère m'a raconté quelque chose d'extraordinaire (L1, PRI-POI-2, CRFP)
b. Le gouvernement allemand a adopté le 1 juillet un projet de budget pour 1993 (E. Le Boucher, 03/07/1992, FTB)

L'échantillon étudié se compose donc comme suit :

Corpus	FTB	CO	Total
Verbes transitifs passivables	500	400	900
dont Actifs	250	200	450
dont actifs à sujet SN	125	100	225
dont actifs à sujet clitique	125	100	225
dont Passifs	250	200	450
dont passifs longs en par	125	100	225
dont passifs courts	125	100	225

Tableau 1 – Composition de l'échantillon étudié

2.3 Annotation

Nous décrivons le passif à l'aide des correspondances suivantes :

Argument 1 = Sujet actif / Complément d'agent passif

Argument 2 = Objet actif / Sujet passif

Nous annotons alors une dizaine de facteurs qui prennent en compte des informations des arguments, du verbe, et du contexte plus général de la phrase :

- 1) Personne : 1^{ère}, 2^{ème}, 3^{ème}
- 2) Nombre : singulier, pluriel
- 3) Genre : masculin, féminin
- 4) Animéité : animé (humain, institution, machine), inanimé (lieu, objet concret ou abstrait) (Zaenen et al., 2004)
- 5) Définitude : défini (déterminants et pronoms définis, démonstratifs, possessifs), indéfini (déterminants et pronoms indéfinis, interrogatifs, négatifs)
- 6) Catégorie syntaxique : syntagme nominal, nom propre, pronom
- 7) Longueur : nombre de mots (Thuilier, 2012)
- 8) Classe sémantique du verbe, à l'aide du dictionnaire Les Verbes Français (LVF) de Dubois et Dubois-Charlier (1997)

9) Lemme du verbe

10) Type syntaxique de la phrase : matrice, subordonnée complétive, relative, circonstancielle

11) Présence d'un passif dans les énoncés qui précèdent (une phrase à l'écrit, trois à l'oral) Nous définissons deux variables supplémentaires à l'aide des annotations précédentes :

12) La longueur relative, qui est la différence entre les logarithmes des longueurs des arguments (Thuilier, 2012)

$$\text{Longueur relative} = \log(\text{Longueur de l'argument 1}) - \log(\text{Longueur de l'argument 2})$$

13) La saillance, qui est une échelle de critères formels (personne, définitude, catégorie syn-taxique) permettant de hiérarchiser l'accessibilité discursive des référents (Lambrecht, 1994 ; Aissen, 2003)

Pronom 1ère/2ème pers. > Pronom 3ème > Nom propre > SN défini > SN indéfini

Dans la plupart des cas, l'argument 1 omis au passif court est présent dans l'interprétation et peut être inféré grâce au contexte (Attal, 1985 ; Druetta, 2020). Ce fut le cas pour 95% des passifs courts de notre corpus, et l'argument 1 a donc fait l'objet d'annotation dans ces cas.

2.4 Analyse statistique

Nous proposer d'analyser les données de corpus à l'aide de modèles statistiques. Nous utilisons ici des modèles de régression logistique à effets mixtes (Agresti, 2003), réalisés à l'aide du package lme4 sur R (Bates et al., 2014).

La régression logistique à effets mixtes

La régression logistique à effets mixtes est un modèle :

- *de régression*, ce qui permet de décrire la relation entre une variable à expliquer (ex. la probabilité d'utiliser le passif) et des variables explicatives (ex. le fait que l'argument 2 soit animé ou inanimé)
- *logistique*, car la variable à expliquer est binomiale (deux valeurs, soit actif soit passif)
- *à effets mixtes*, car le modèle peut prendre en compte plusieurs variables explicatives en même temps (effets fixes) mais aussi prendre en compte des sous-groupes de données, selon la modalité écrit/oral ou selon le locuteur par exemple (effets aléatoires)

Les statistiques descriptives (chiffres bruts, proportions, moyennes etc) possèdent plusieurs limites : les résultats se cantonnent à l'échantillon, les valeurs agrégées comme la moyenne peuvent masquer la distribution des données ou les relations entre variables etc. L'apport méthodologique de la modélisation statistique est de dépasser les statistiques descriptives en envisageant :

1. des statistiques *inférentielles* : un modèle permet de faire des inférences sur la population (ex. le corpus) à partir de l'échantillon. Un modèle estime ainsi la distribution de la population, en quantifiant la relation entre variable à expliquer et variables explicatives. Pour savoir si l'effet d'une variable explicative est également valable dans la population, et non dû au hasard de l'échantillonnage, le modèle fournit une mesure de significativité statistique (*p-value*).
2. des statistiques *prédictives* : une fois qu'un modèle a appris une relation entre variable à expliquer et variables explicatives, celui-ci est capable de généraliser à de nouveaux points de données. Ainsi, nos modèles seront capables de calculer la probabilité qu'une phrase soit au passif selon les variables explicatives qu'on annoté. Le modèle peut donc être utilisé dans une *tâche de classification*, en prédisant si une phrase est

plus probable au passif ou à l'actif. Mesurer la qualité de la prédiction d'un modèle permet ainsi de savoir si celui-ci généralise correctement à partir des données apprises et rend donc bien compte du phénomène étudié.

Les modèles calculent avec les paramètres suivants :

- La variable à expliquer est la voix (actif/passif). Nous posons que Passif =1 et Actif =0.
- Les variables explicatives correspondent aux facteurs annotés (les variables catégorielles sont codées de façon binaire et normalisées).
- Les variables aléatoires correspondent au lemme du verbe et, lorsque c'est pertinent, au locuteur et au corpus.

Nous utilisons deux mesures d'évaluation de la qualité de prédiction des modèles : l'exactitude (calculée par validation croisée répétée 100 fois) et l'aire sous la courbe ROC, l'AUC (Area Under Curve). Nous ne détaillons pas ici le fonctionnement de ces méthodes (voir Thuilier, 2012, p. 82-86 pour une présentation et références). Les modèles sont considérés comme prédictivement intéressants si leur évaluation par ces mesures dépasse 0.80 (Harrell, 2001).

3. Modélisation statistique

3.1 L'alternance actif/passif

Nous proposons 4 modèles pour l'alternance entre actif et passif :

1. un modèle général du passif long (oral et écrit) : modèle FTB+CO (tableau 2)
2. un modèle général du passif court (oral et écrit) : modèle Passif Court (tableau 3)
3. un modèle du passif long à l'écrit : modèle FTB (tableau 4)
4. un modèle du passif long à l'oral : modèle CO (tableau 5)

Les modèles fournissent les informations suivantes :

- Chaque variable reçoit un coefficient. Si celui-ci est positif, la variable augmente la probabilité d'obtenir un passif et s'il est négatif, c'est la probabilité d'obtenir un actif qui augmente.
Exemple : Dans le modèle FTB+CO, la variable « arg2 défini » a un coefficient positif ($c = 0.85$). Il est donc plus probable d'utiliser un passif lorsque l'argument 2 est défini plutôt qu'indéfini.
- La valeur absolue du coefficient indique la force de l'effet : plus la valeur est grande, plus la probabilité d'utiliser un actif ou un passif va varier. Les variables des modèles sont rangées par ordre décroissant des valeurs des coefficients (les coefficients les plus hauts ont un effet plus fort que ceux plus bas).
Exemple : Dans le modèle FTB+CO, la variable qui a l'effet le plus fort sur la probabilité d'utiliser le passif est la longueur ($c = 1.06$) et celle qui a l'effet le moins fort est la classe sémantique du verbe ($c = 0.36$).
- Les résultats des tests de significativité statistique correspondent à la z-value et à la p-value. Toutes les variables des modèles ont un effet significatif sur la probabilité de passif (c-à-d dont on peut supposer avec confiance que l'effet est valable pour la population). En effet, les p-values des variables sont toutes inférieures au seuil de 5% ($p < 0.05$).
- L'exactitude, en pourcentage, mesure la capacité du modèle à généraliser à de nouvelles données. On peut l'interpréter comme un taux moyen de prédiction correcte du modèle.
Exemple : Le modèle FTB+CO a une exactitude de 83%, et ne se trompe donc que dans 17% des cas pour prédire si une phrase est un actif ou un passif.

- L'AUC (Area Under Curve) mesure la capacité du modèle à prédire correctement les données sur lesquelles il a appris. Une AUC de 1 indique que le modèle est un classificateur parfait, et une AUC de 0.5 indique que le modèle est un classificateur aléatoire (aucune capacité prédictive).

	Coefficient	Erreur-type	z-value	p-value
Intercept	-7.16	5.06	-1.415	> 0.1
Longueur relative	1.06	0.16	6.79	< 0.001
arg2 pronom relatif	0.96	0.23	4.24	< 0.001
arg2 défini	0.85	0.18	4.64	< 0.001
arg1 défini	-0.65	0.17	-3.82	< 0.001
arg2 pluriel	0.42	0.15	2.75	< 0.01
Amorce passive	0.42	0.15	2.88	< 0.01
arg2 animé	0.40	0.15	2.68	< 0.01
Verbe de type frapper	0.36	0.13	2.68	< 0.01
Exactitude = 83% ; AUC = 0.99				

Tableau 2 – Modèle FTB+CO (n=675)

Le modèle FTB+CO est un modèle général du passif long, à l'écrit et à l'oral. Il montre un effet significatif de multiples facteurs : si l'argument 2 est un pronom relatif, s'il est défini, animé ou pluriel, la probabilité d'utiliser un passif augmente. Si l'argument 1 est défini, c'est la probabilité d'utiliser un actif qui augmente. On trouve également un effet de la longueur relative : plus la longueur relative augmente (c'est-à-dire plus l'argument 1 est long par rapport à l'argument 2), plus la probabilité d'utiliser un passif augmente. Enfin on trouve un effet d'amorçage passif (la présence d'un passif dans le discours favorise la réutilisation du passif) et un effet de classe sémantique (les verbes de type frapper comme « assassiner », « détruire », favorisent la construction passive).

	Coefficient	Erreur-type	z-value	p-value
Intercept	0.15	0.17	0.87	> 0.1
arg1 1 ^{ère} /2 ^{ème} pers.	-0.70	0.12	-5.90	< 0.001
Amorce passive	0.55	0.12	4.42	< 0.001
arg1 pluriel	0.53	0.14	3.93	< 0.001
arg2 animé	0.38	0.13	2.97	< 0.01
Longueur arg2	-0.35	0.16	-2.16	< 0.05
arg1 féminin	0.34	0.13	2.66	< 0.01
arg2 défini	0.29	0.13	2.21	< 0.05
Exactitude = 70% ; AUC = 0.80				

Tableau 3 – Modèle Passif Court (n=300)

Le modèle Passif Court a des facteurs communs avec le modèle FTB+CO : la présence d'une amorce passive et un argument 2 animé ou défini augmentent la probabilité d'utiliser le passif. On trouve à nouveau un effet de longueur : plus l'argument 2 est long, plus la probabilité d'utiliser l'actif augmente. Enfin, plusieurs effets de l'argument 1 sont à noter : les arguments 1 pluriels ou féminins favorisent la construction passive, tandis qu'un argument 1 de 1^{ère}/2^{ème} personne favorise l'actif.

	Coefficient	Erreur-type	z-value	p-value
Intercept	-4.25	3.20	-1.33	> 0.1
Longueur relative	0.83	0.16	5.14	< 0.001
arg2 défini	0.82	0.20	4.00	< 0.001
arg2 plus saillant que arg1	0.58	0.17	3.36	< 0.001
arg2 pluriel	0.56	0.19	2.89	< 0.01
arg2 animé	0.55	0.19	2.87	< 0.01
arg1 pluriel	0.55	0.17	3.20	< 0.01
Exactitude = 84% ; AUC = 0.99				

Tableau 4 – Modèle FTB (n=375)

Le modèle FTB montre la significativité de la définitude, de l'animéité, de la saillance et de la longueur (un argument 2 défini, animé, court ou plus saillant que l'argument 1 favorise le passif). On trouve également un effet significatif du nombre : un argument 1 ou un argument 2 au pluriel augmente la probabilité d'utiliser un passif.

	Coefficient	Erreur-type	z-value	p-value
Intercept	-5.56	3.21	-1.73	< 0.1
Longueur de arg1	1.62	0.42	3.87	< 0.001
arg2 pronominal	1.17	0.25	4.64	< 0.001
arg2 défini	0.86	0.30	2.85	< 0.01
arg1 défini	-0.81	0.29	-2.83	< 0.01
Verbe de de type frapper	0.76	0.26	2.95	< 0.01
arg2 féminin	-0.71	0.25	-2.78	< 0.01
Exactitude = 86% ; AUC = 0.99				

Tableau 5 – Modèle CO (n=300)

Le modèle CO, tout comme le modèle FTB, montre qu'un argument 2 défini favorise l'utilisation d'une construction passive. Un argument 2 pronominal ou féminin augmente également la probabilité d'utiliser le passif, de même qu'un argument 1 long. On trouve à nouveau un effet de la classe des verbes de type frapper. Enfin, l'actif est plus probable lorsque l'argument 1 est défini.

A travers ces modèles nous avons pu voir que les arguments définis, animés, courts, saillants, pronominaux ou de 1^{ère}/2^{ème} personne favorisent systématiquement la construction où ils apparaissent comme sujet. Ainsi, dans le modèle FTB+CO, un argument 2 défini est plus probable au passif, donc comme sujet ($c = 0.85$) tandis qu'un argument 1 défini est plus probable à l'actif, donc à nouveau comme sujet ($c = -0.65$). Inversement, les arguments indéfinis, inanimés, longs, peu saillants, nominaux, ou de 3^{ème} personne apparaissent donc plus souvent comme complément (objet à l'actif, complément d'agent au passif).

Par conséquent, la fonction sujet tend à être la fonction des arguments définis, animés, courts, saillants, pronominaux ou de 1^{ère}/2^{ème} personne, et la fonction complément celle des arguments aux propriétés opposées. Cette asymétrie dans les préférences de codage des arguments s'appelle l'alignement harmonique (Aissen 1999, 2003 ; Bresnan et al., 2001, Haspelmath, 2020). Le choix de construction entre actif et passif permet ainsi de satisfaire la contrainte d'alignement harmonique, en codant les arguments dans leur fonction privilégiée selon leurs propriétés d'animéité, de définitude, de longueur etc. L'alternance actif/passif permet ainsi un codage harmonique, efficace des arguments et nos modèles montrent que cette tendance est présente aussi bien à l'écrit qu'à l'oral.

Comme évoqué en section 2.4, les modèles peuvent être utilisés de façon prédictive. A partir d'une paire d'arguments annotés et de propriétés sémantiques et discursives générales (classe sémantique du verbe, présence de passifs dans le discours), les modèles peuvent assigner une probabilité de passif P aux phrases. En guise d'illustration, voici les valeurs de probabilité de passif attribuées à quelques phrases de corpus par le modèle FTB+CO.

- (5) a. $P = 92\%$: Les notaires affirment que [leurs clients]_{arg2} ont été déçus [par les mesures récentes de relance du logement]_{arg1} (04/06/1993, FTB)
b. $P = 92\%$: [Le raffermissement]_{arg2} a été provoqué [par des ventes de marks au profit du yen]_{arg1} (F. Renard, 06/01/1992, FTB)
c. $P = 95\%$: euh donc à la fin de la guerre il y avait un pont [qui]_{arg2} avait été refait [par les américains]_{arg1} (L1,PRI-ROU-2, CRFP)
d. $P = 99\%$: enfin il y a eu un moment où [ils]_{arg2} ont été chassés [par euh par un nationaliste haïtien]_{arg1} (DAV, ffamd110, C-Oral-Rom)
- (6) a. $P = 12\%$: [M. Deng Xiaoping]_{arg1} a relancé [la politique d'ouverture économique du pays]_{arg2} (F. Deron, 06/07/1992, FTB)
b. $P = 1\%$: [les vieux]_{arg1} avaient fait [un petit sapin]_{arg2} (Gabriel Pujade, CFPP)
- (7) $P = 3\%$: [Huit ans passés à la tête d'Air France au Royaume-Uni]_{arg2} ont été vécus [par lui]_{arg1} comme un "plus" déterminant. (01/1990)

Les exemples en (5) illustrent le cas de passifs attestés en corpus et jugés très probables par le modèle. Plusieurs facteurs contribuent en effet à utiliser un passif dans ces cas : l'argument 2 est court, défini, et animé dans le cas de (5a), relatif en (5c) tandis que l'argument 1 est indéfini en (5b) et (5d). Ces phrases illustrent donc des cas de codage harmonique des arguments.

Les exemples en (6) illustrent le cas de phrases correctement classifiées comme actives par le modèle : elles sont à l'actif dans le corpus et le modèle leur assigne également une faible probabilité de passif (le passif est probable à 12% en (6a), donc l'actif à 88%). Ces phrases à l'actif mettent à nouveau en jeu un codage harmonique des arguments : l'argument 1 est défini, animé, court, par rapport aux arguments 2 qui sont inanimés, plus longs ou indéfinis.

L'exemple (7) est un cas de passif attesté en corpus qui est jugé peu probable par le modèle, et qui devrait donc plutôt être à l'actif (3% pour le passif, donc 97% pour l'actif). Cet exemple ne respecte en effet pas l'alignement harmonique : l'argument 2 sujet est long, indéfini, inanimé, tandis que l'argument 1 complément est court, pronominal, inanimé. Cet exemple montre que l'alternance actif/passif peut être soumise à d'autres facteurs que ceux pris en compte par nos modèles (malgré l'apport prédictif déjà correct de ces derniers). Remarquons toutefois que l'exemple (7) reste marginal et semble présenter une certaine dégradation d'acceptabilité. En le classifiant incorrectement, le modèle met en évidence son caractère marginal.

3.2 Le passif à l'oral et à l'écrit

Nous avons montré des similitudes dans l'alternance actif/passif entre français oral et écrit

(12) travers divers modèles statistiques (effet d'alignement harmonique, de longueur, de classe sémantique, d'amorçage). En suivant la méthode proposée par Bresnan et Ford (2010), nous allons également utiliser des modèles statistiques pour mesurer les différences significatives entre français oral et français écrit.

Pour tester les différences entre écrit et oral, nous calculons des interactions statistiques. L'interaction statistique est la situation dans laquelle les effets de deux facteurs ne sont pas simplement additifs, mais que l'effet d'un des facteurs (ex. la définitude) dépend des valeurs d'un autre facteur (ex. la modalité oral/écrit). Les interactions sont notées Facteur1 Facteur2. Un effet qui n'est pas une interaction s'appelle un effet principal.

Du fait que les coefficients d'interaction sont souvent difficiles à interpréter, nous accompagnons nos modèles de graphiques représentant la fréquence du passif à l'oral et à l'écrit selon les facteurs suivants, que nous proposons d'étudier spécifiquement :

- La définitude, en figure 1
- La pronominalité, en figure 2
- La longueur de l'argument 1, en figure 3
- La longueur de l'argument 2, en figure 4

En figure 1 nous observons un effet principal de définitude : les arguments 2 définis sont plus fréquemment codés comme sujets de passifs aussi bien à l'oral qu'à l'écrit. On trouve également un effet d'interaction significatif ($p < 0.05$) : lorsque l'argument 2 est indéfini, il est plus rarement sujet de passif à l'oral (13%) qu'à l'écrit (29%). Autrement dit, l'oral affiche une préférence plus importante pour la construction active lorsque l'argument 2 est indéfini.

En figure 2, on trouve un effet principal de pronominalité : les arguments pronominaux sont plus souvent sujets de passifs que les arguments nominaux, à l'écrit et à l'oral. S'observe également un effet significatif d'interaction ($p < 0.001$) : si l'argument 2 est pronominal il sera plus souvent sujet de passif à l'oral (69%) qu'à l'écrit (58%) et s'il est nominal c'est à l'écrit qu'on trouve plus de codage comme sujet de passif (48%, contre 28% à l'oral).

Les figures 3 et 4 montrent des résultats convergeants quant aux effets de longueur. Nous observons que plus un argument devient long, moins il est utilisé comme sujet. Ainsi si l'argument 1 s'allonge, la fréquence du passif augmente, construction où il est complément. De même, si l'argument 2 s'allonge, la fréquence du passif diminue, de sorte qu'il ne soit pas codé comme sujet passif mais comme objet actif. Ces effets de longueur se trouvent à la fois à l'écrit et à l'oral. Les effets d'interaction entre longueur des arguments et modalité sont significatifs ($p < 0.05$ pour Longueur arg1 Modalité, $p < 0.001$ pour Longueur arg2 Modalité). Sur chaque figure on constate que la pente des données de l'oral est plus forte que celle de l'écrit. Autrement dit la dispréférence générale envers les sujets longs est plus forte à l'oral : lorsque l'argument 1 s'allonge, le passif est plus fréquent à l'oral qu'à l'écrit, construction où cet argument 1 sera complément et non sujet. Inversement lorsque l'argument 2 s'allonge, le passif est moins utilisé à l'oral, pour éviter son codage en sujet.

De tous les facteurs étudiés émergent donc deux conclusions. D'abord, l'écrit et l'oral affichent des tendances communes, notamment une préférence pour les sujets définis, pronominaux et courts, et pour les compléments indéfinis, nominaux et longs : c'est l'alignement harmonique. Ensuite, l'oral affiche systématiquement une préférence plus forte que l'écrit pour le codage harmonique, avec une proportion plus importante de compléments indéfinis, nominaux et longs, ou de sujets pronominaux et courts.

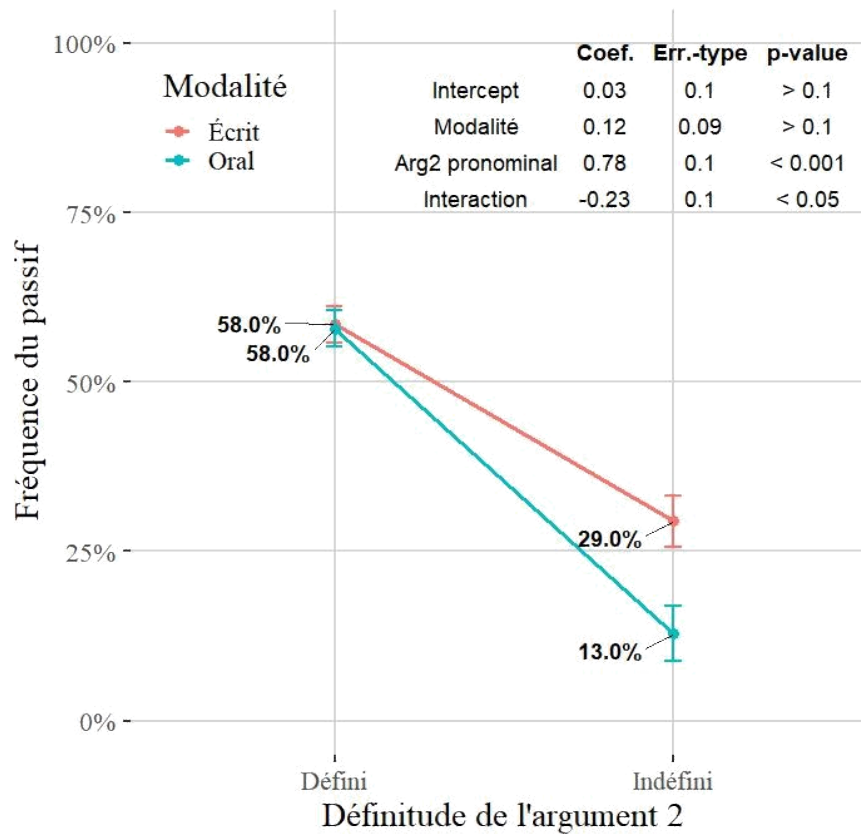


Figure 1 – Fréquence du passif selon la définitude de l'argument 2 à l'oral et à l'écrit. Le tableau présente une régression logistique qui calcule l'interaction arg2 défini * Modalité.

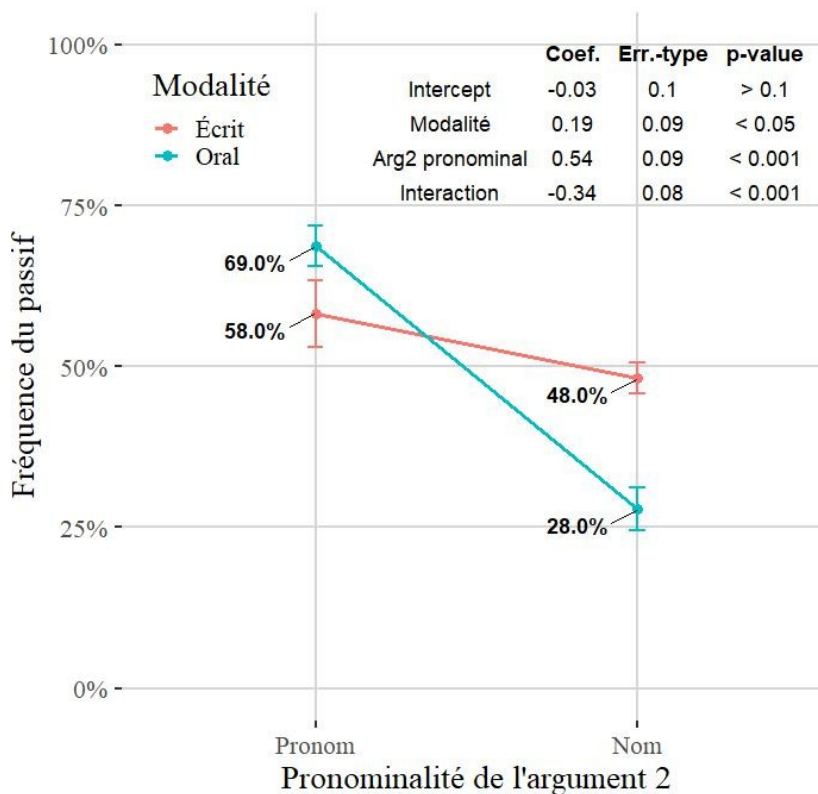


Figure 2 – Fréquence du passif selon la pronominalité de l'argument 2 à l'oral et à l'écrit. Le tableau présente une régression logistique qui calcule l'interaction arg2 pronominal * Modalité.

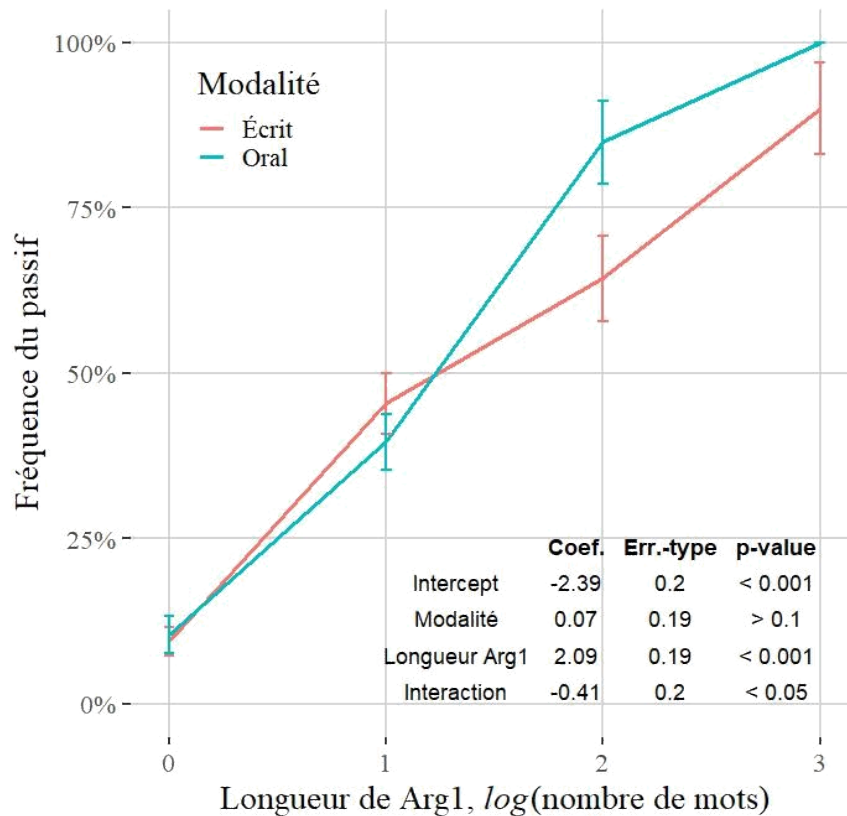


Figure 3 – Fréquence du passif selon la longueur de l'argument 1 à l'oral et à l'écrit. Le tableau présente une régression logistique qui calcule l'interaction Longueur arg1 * Modalité.

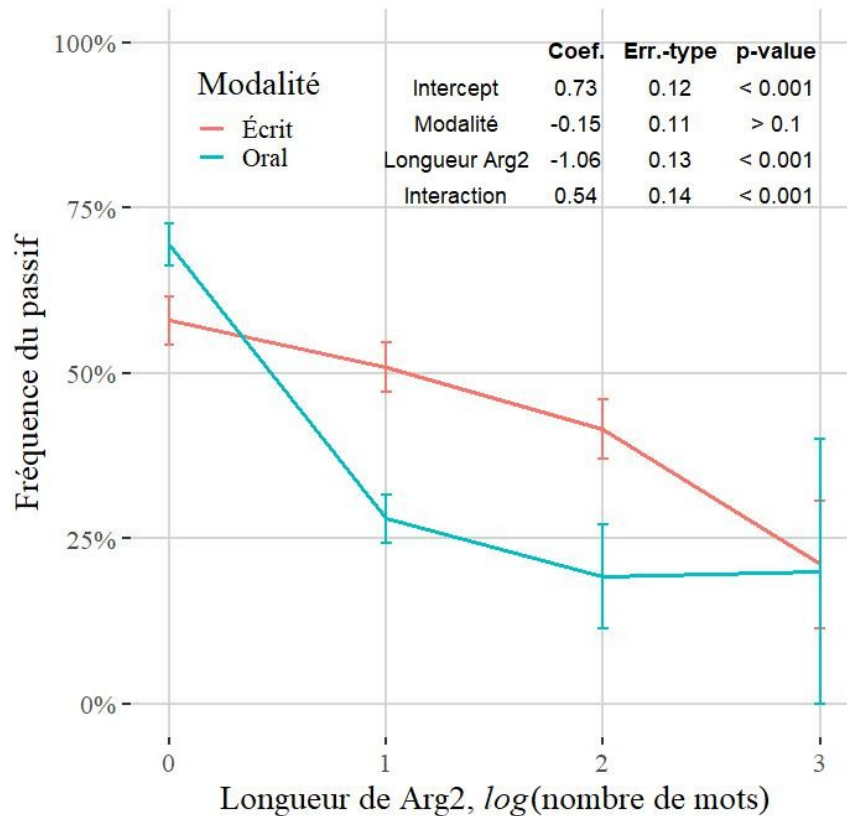


Figure 4 – Fréquence du passif selon la longueur de l'argument 2 à l'oral et à l'écrit. Le tableau présente une régression logistique qui calcule l'interaction Longueur arg2 * Modalité.

4 Discussion

Nous avons proposé une description de l'alternance actif/passif en français écrit et oral à l'aide de modèles statistiques qui apprennent sur des données de corpus. Ces modèles ont mis en évidence le caractère multifactoriel de cette alternance, en montrant la significativité de facteurs hétérogènes qui conduisent au choix de construction, aussi bien à l'oral qu'à l'écrit :

- Les arguments définis, pronominaux, courts, saillants, animés ou de 1^{ère}/2^{ème} personne sont préférablement codés dans la construction où ils sont sujets (argument 1 pour l'actif, argument 2 pour le passif). Il s'agit d'un effet d'alignement harmonique des arguments
- Les verbes de type frapper se construisent davantage au passif.
- La présence d'un passif dans les énoncés antérieurs favorise la réutilisation du passif.

Ces résultats sont cohérents avec les études menées en anglais, suggérant une certaine généralité des facteurs proposés, et ils permettent également de quantifier les résultats d'études précédentes en français.

A l'exception du modèle Passif Court, tous nos modèles possèdent une exactitude de plus de 80% et une AUC proche de 1. Les modèles proposés ont donc un intérêt prédictif dans l'alternance de construction : ils généralisent relativement bien à de nouvelles données. La capacité prédictive des modèles permet ainsi de rendre compte de la gradience dans l'alternance entre actif et passif, par exemple par la possibilité d'assigner des probabilités aux constructions.

Enfin, au delà des similarités entre oral et écrit dans l'utilisation du passif, nous avons pu montrer des différences significatives entre ces deux modalités grâce au calcul des interactions entre facteurs. Nous avons notamment pu voir que l'oral exprime une préférence plus forte que l'écrit pour l'alignement harmonique : le choix entre actif et passif opère de sorte à maximiser davantage les codages harmoniques à l'oral, comme les sujets pronominaux et courts, ou les compléments indéfinis, nominaux et longs. Autrement dit, l'écrit admet mieux les configurations non harmoniques (sujets longs, indéfinis ou nominaux).

La différence observée entre écrit et oral peut s'expliquer par les différences de conditions de production et de traitement des deux modalités. Alors que l'écrit est détaché de la situation d'interaction et que l'information linguistique y est accessible de façon statique, l'oral est dynamique et est soumis à une pression à communiquer directement, en canal bruité ou en débit rapide (Halliday, 1989 ; Mazur-Palandre, 2015 ; Poiret & Liu, 2020). L'alignement harmonique peut être vu comme une tendance au codage efficace des arguments (Haspelmath, 2020). L'oral, par sa préférence plus forte pour le codage harmonique, tend donc à proposer des productions plus efficaces communicativement (voir Poiret et Liu, 2020 pour des résultats similaires en corpus en français, et Mazur-Palandre, 2015 en expérience). L'écrit en revanche permet de plus facilement s'éloigner des configurations harmoniques, et présente donc des productions moins efficaces.

Pour terminer, cette étude montre l'intérêt d'analyser les données de corpus à l'aide de méthodes statistiques, notamment dans le cadre des alternances de construction. Ces méthodes permettent de tenir compte de facteurs multiples, de décrire des effets probabilistes, non catégoriels, et de mettre en évidence des différences entre variétés de langue.

Références

- Abeillé, A. & Clément, L. & Liégeois, L. (2019). Un corpus annoté pour le français : le French Treebank. *TAL*, 60, 19-43.
- Agresti, A. (2003). *An Introduction to Categorical Data Analysis*. Wiley.
- Aissen, J. (1999). Markedness and subject choice in optimality theory. *Natural Language & Linguistic Theory*, 17 (4), 673-711.

- Aissen, J. (2003). Differential object marking : iconicity vs. economy. *Natural Language and Linguistic Theory*, 21 (3), 435-483.
- Attal, P. (1985). Le passif : étude de corpus. *L'Information Grammaticale*, 27 (1), 10-13. Bates, D. & Mächler, M. & Bolker, B. & Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv :1406.5823.
- Benzitoun, C. & Debaisieux, J.-M. & Deulofeu, H.-J. (2016). Le projet ORFÉO : un corpus d'étude pour le français contemporain. *Corpus*, (15), 91-114.
- Blanche-Benveniste, C. (1989). Les régulations syntaxiques dans les productions de français parlé [Publisher : Persée - Portail des revues scientifiques en SHS]. *LINX*, 20 (1), 7-20.
- Bresnan, J. & Dingare, S. & Manning, C. D. (2001). Soft constraints mirror hard constraints : Voice and person in English and Lummi (M. Butt & T. Holloway King, eds.). In M. Butt & T. Holloway King (eds.), *Proceedings of the LFG01 Conference*, Hong Kong, Stanford : CSLI Publications.
- Bresnan, J. & Ford, M. (2010). Predicting syntax : Processing dative constructions in American and Australian varieties of English [Publisher : JSTOR]. *Language*, 168-213.
- Dingare, S. (2001). The effect of feature hierarchies on frequencies of passivization in English. Unpublished MA thesis, Stanford University. ROA-467-0901.
- Druetta, R. (2020). Le passif à l'oral. Phénoménologie et propriétés aspectuelles dans « OFROM » *Studia linguistica romanica*, (4), 150-174.
- Dubois, J. & Dubois-Charlier, F. (1997). *Les verbes français*. Paris, Larousse-Bordas.
- Estival, D. (1985). Syntactic priming of the passive in English. *Text-Interdisciplinary Journal for the Study of Discourse*, 5 (1), 7-22.
- Halliday, M. A. K. (1989). *Spoken and written language*. Oxford University Press, USA.
- Hamma, B. & Tardif, A. & Badin, F. (2017). Le passif à l'oral (P. Larrivée & F. Lefevre, eds.). *Français contemporain vernaculaire (FRACOV)*.
- Harrell, F. E. (2001). *Regression Modeling Strategies : With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. Springer International Publishing.
- Haspelmath, M. (2020). Role-reference associations and the explanation of argument coding splits. *Linguistics*, 59 (1), 123-174.
- Henley, N. & Millier, M. & Baezley, J. A. (1995). Syntax, Semantics and Sexual Violence : Agency and the Passive Voice, 14, 60-84.
- Hundt, M. & Röthlisberger, M. & Seoane, E. (2018). Predicting voice alternation across academic Englishes. *Corpus Linguistics and Linguistic Theory*, 1.
- Lambrecht, K. (1994). *Information Structure and Sentence Form : Topic, Focus, and the Mental Representations of Discourse Referents* (T. 71). Cambridge, Cambridge University Press.
- Mazur-Palandre, A. (2015). Overcoming Preferred Argument Structure in written French : development, modality, text type. *Written Language & Literacy*, 18 (1), 25-55.
- McLaughlin, M. (2010). L'influence de l'anglais sur la syntaxe du français : une étude de cas concernant la voix passive. *2ème Congrès Mondial de Linguistique Française*, 129.
- Poiret, R. & Liu, H. (2020). Some quantitative aspects of written and spoken french based on syntactically annotated corpora. *Journal of French Language Studies*, 30 (3), 355-380.
- Roland, D. & Dick, F. & Elman, J. L. (2007). Frequency of basic english grammatical structures : a corpus analysis. *Journal of Memory and Language*, 57 (3), 348-379.
- Thuilier, J. (2012). *Contraintes préférentielles et ordre des mots en français* ([thèse de doct.]) [Université de Paris]. Université de Paris.
- Weiner, E. J. & Labov, W. (1983). Constraints on the agentless passive. *Journal of Linguistics*, 19 (1), 29-58.
- Zaenen, A. & Carletta, J. & Garretson, G. & Bresnan, J. & Koontz-Garboden, A. & Nikitina, T. O'Connor, M. C. & Wasow, T. (2004). Animacy encoding in English : Why and how, In *Proceedings of the Workshop on Discourse Annotation*.

Apport des études sur corpus : le domaine nominal

Jeanne-Marie Debaisieux, LATTICE Université Sorbonne Nouvelle

L'objectif de cette présentation est de mettre en lien deux observations relevées dans l'ouvrage de Claire Blanche-Benveniste *Le français. Usages de la langue parlée* paru en 2010. La première est un constat sur la permanence d'un manque d'outils pour l'analyse de la grammaire du français parlé. La deuxième observation concerne le fonctionnement multiple de certains mots lexicaux–grammaticaux.

En conclusion de son ouvrage Claire Blanche-Benveniste formulait un vœu :

« Il serait bon d'organiser une grande collecte de français parlé, à la hauteur de ce qui a été fait pour d'autres langues. Elle devrait rassembler des productions variées, dans des situations très diversifiées, traitées selon toutes les techniques modernes, mais sans que ces techniques empêchent de s'intéresser au contenu des textes ». (2010, 223)

Plus de dix ans après, on pourrait penser que cet objectif est en bonne voie d'être accompli. En effet, comme le soulignent Corminboeuf G., Rothenbühler J. & Sauzet M. (2020), il existe aujourd'hui de nombreux corpus de français parlé accessibles en ligne. Je reprends ici la liste qu'ils dressent :

CEFC = Outils et ressources sur le français écrit et oral (ORFÉO) (éd.) Corpus d'étude pour le français contemporain.

CFPB = Anne Dister, Emmanuelle Labeau (éds.) Corpus de français parlé à Bruxelles.

CFPP2000 = Sonia Branca-Rosoff et al. (éds.) Corpus de français parlé parisien des années 2000.

CFPQ = Gaétane Dostie (éd.) Corpus de français parlé au Québec.

CIEL-F = Lorenza Mondada et al. (éds.) Corpus international écologique de la langue française.

CLAPI = Laboratoire ICAR (éd.) Corpus de langue parlée en interaction.

ESLO = Laboratoire ligérien de linguistique (éd.) Enquêtes sociolinguistiques à Orléans.
<http://eslo.huma-num.fr>.

FRAN = France Martineau et al. (éds.) Corpus du français en Amérique du Nord.

MPF = Françoise Gadet (éd.) Multicultural Paris French.

OFROM = Mathieu Avanzi, Marie-José Béguelin, Federica Diémoz (éds.) Corpus oral de français de Suisse romande.

PFC = Jacques Durand, Bernard Laks, Chantal Lyche (éds.) Phonologie français contemporain.

À ces corpus de langue parlée, s'ajoutent des millions de mots collectés souvent extraits du web pour la constitution de données écrites massives.

Tout corpus est en soi source de recherche, mais comme le développent Egbert J., Larsson T. & Biber D. (2020), la question qui se pose au chercheur devant la diversité des données concerne la pertinence du corpus en fonction de l'objet d'étude mais également ses limites.

C'est ainsi qu'on aurait envie d'appliquer au domaine français la remarque des auteurs « *Paradoxically, doing corpus linguistics is both easier and harder than it has ever been before* » (2)

En effet, pour évaluer la pertinence d'un corpus par rapport à l'objet de recherche visé le chercheur doit disposer d'un certain nombre d'informations, en premier lieu sur la nature et la composition du corpus. Ces informations concernent d'abord ce qu'il est coutume d'appeler les métadonnées du corpus : nature des sous-corpus, situations date et lieu des enregistrements, durée, qualité sonore et les métadonnées concernant les locuteurs : sur le sujet des échanges, sur le nombre de locuteurs et les relations qu'ils entretiennent ainsi que quelques informations minimales sur l'âge et leur situation socioculturelle au moment de l'enregistrement. Le chercheur doit également avoir un accès direct au contenu même des transcriptions afin de vérifier si la nature des textes correspond aux comportements langagiers qui l'intéressent ou au domaine d'étude qu'il souhaite explorer. C'est seulement si ces informations sont disponibles facilement que le chercheur pourra répondre à la nécessité de s'intéresser au contenu des textes, signalée par Claire Blanche-Benveniste dans la citation ci-dessus.

Mais d'autres informations de nature proprement linguistique sont nécessaires à la description. Dans une approche descriptive de la syntaxe, un des objectifs de la recherche sur corpus est de déterminer les contextes dans lesquels la structure observée apparaît. Il s'agit en d'autres termes d'analyser sa distribution à partir des informations que peut fournir ligne à ligne un concordancier. Ces informations, concernant notamment la taille du contexte et les propriétés de la structure recherchée, doivent pouvoir être contrôlées par le chercheur au moyen d'un retour direct au texte dans lequel apparaît l'occurrence à analyser ainsi qu'un accès simple au signal sonore. Le corpus utilisé dans la recherche doit permettre l'accès à l'ensemble de ces informations qui constituent la base de toute analyse qualitative sur corpus. Pour illustrer cette problématique, je partirai des descriptions de Claire Blanche-Benveniste développées sous la rubrique « *Multiplés fonctionnements des noms* ». (2010, 105) L'auteur y note :

« *Comme il n'y a pas coïncidence entre catégorie et fonction on rencontre dans les syntagmes nominaux des noms qui, à part le rôle majeur de constructeur de valence nominale assument différents autres fonctionnements dont quelques-uns seront mentionnés ici : rôle de modifieurs, de quantifieurs, de classifieurs, de proformes ou d'éléments de locutions* »

Ainsi *restaurant* est un modifieur dans *chèque restaurant* ; *tas* est un nom lexical dans *un tas de boue* mais fonctionne comme quantifieur dans *un tas de raisons*. Les noms comme *grains*, *pincée*, *filet* fonctionnent comme des classifieurs, les noms génériques : *truc chose*, *type*... sont utilisés comme des proformes et les noms comme *côté*, *question*, *genre*, ... se comportent comme des prépositions. Pour Claire Blanche-Benveniste « *Ces cinq types d'emplois montrent la grande plasticité syntaxique et sémantique de certains noms de grande fréquence*. (2020, 110)

L'analyse de ces sous-classes de Noms « grammaticaux » dont la syntaxe idiosyncrasique ne peut se révéler qu'au travers de distributions particulières, qui, pour l'auteur, sont « *très sensibles aux différences de situation et aux registres de parole* » (220, 106) pose les questions de recherche suivantes :

- Quelles sont les conditions dans lesquelles un Nom lexical peut adopter un comportement grammatical ?

- Combien de comportements grammaticaux différents peut-on observer pour les noms concernés ?

Ces questions ont été à l'origine d'études en sémantique dans le cadre des « Noms généraux » Kleiber G. (1987), (Benninger C. 2014), (Huyghe R., 2015), (Adler S. & Moline E., 2018), mais également d'études sur les faits de grammaticalisation (Rosier 2002) et d'études sociolinguistiques, (Trimaille C., Pereira C, Ziamari K., Gasquet-Cyrus M., 2020). Dans le cadre de l'approche pronominale, Claire Blanche-Benveniste (1985) offre un bel exemple d'analyse du nom « chose ». Des études existent sur des emplois grammaticaux d'autres catégories (Blanche-Benveniste, (2001) pour l'adjectif « plein », Lafontaine F. & Krejčí K. (2018), pour « juste ». Mais nous ne disposons pas d'un inventaire systématique à partir de données vérifiables des fonctionnements des « Noms grammaticaux ».

De telles études sont pourtant nécessaires pour cerner les limites entre la compétence grammaticale des locuteurs et leur usage du lexique, problématique centrale pour l'Approche pronominale. Cette compétence grammaticale que Claire Blanche-Benveniste (2010) définit comme « *une pluralité de compétences* » ne peut être décrite qu'au travers de corpus présentant des situations variées relevant de genres différents.

Pour illustrer cette problématique, j'ai choisi d'étudier les divers fonctionnements des noms *limite* et *début* qui relèvent respectivement dans leurs emplois lexicaux du domaine spatial et temporel. Pour ce faire, j'exploiterai les corpus de français hexagonal cités plus haut afin de déterminer s'ils fournissent des données pertinentes pour aboutir à une description satisfaisante.

Au-delà des Noms, c'est bien évidemment les fonctionnements grammaticaux de l'ensemble des catégories lexicales qu'il faut étudier à partir des corpus existants et de ceux qu'il nous faut encore construire afin de pouvoir répondre, entre autres, à une question centrale pour l'Approche Pronominale : distinguer le domaine de la grammaire de celui du lexique.

Références

- Adler S. & Moline E., (2018), « Les noms généraux : présentation », *Langue française* 198, no. 2, 5-18.
- Benninger C. (2014), « La question de la définition sémantique du nom atypique *chose* », *Travaux de linguistique* 69, 35-55.
- Blanche-Benveniste, C., 1985, « Le mot *chose* dans la syntaxe verbale » *Recherches sur le français parlé*, 7, Publications U. de Provence. 141-167
- Blanche-Benveniste, C. & Martin, P. (2010) : *Le français. Usages de la langue parlée*. Leuven/Paris : Peeters
- Blanche-Benveniste, C., (2001) « *Juste et plein, tout juste et tout plein* », *Langage et référence. Mélanges offerts à Kerstin Jonasson à l'occasion de ses soixante ans*, sous la dir. de Hans Kronning, Coco Norén, Bengt Novén, et al., Uppsala, Acta Universitatis Upsaliensis, 65-74.
- Corminboeuf G., Rothenbühler J. & Sauzet M. (éds) (2020), *Studia Linguistica Romanica* n° 4, « Français parlés et français 'tout court' », Publication électronique : <https://studialinguisticaromanica.org/index.php/slr/issue/view/4>
- Huyghe R. (2015), « Les typologies nominales : présentation », *Langue française* 185, 5-28.
- Kleiber G. (1987), « Mais à quoi sert donc le mot *chose* ? Une situation paradoxale », *Langue française* 73, 109-128.
- Lafontaine F. & Krejčí K. (2018), « Syntaxe et sémantique des tournures telles que c'est juste énorme ! » SHS Web of Conferences 46,

14003(<https://doi.org/10.1051/shsconf/20184614003> Congrès Mondial de Linguistique Française - CMLF 2018

Rosier, L. (2002), « Genre : le nuancier de sa grammaticalisation », *Travaux de linguistique* 44, no1 79-88.

Trimaille C., Pereira, C., Ziamari K., Gasquet-Cyrus M., (dir), 2020, *Sociolinguistique des pratiques langagières de jeunes. Faire genre, faire style, faire groupe autour de la méditerranée*, UGA Éditions Collection Langues gestes paroles.

Les verbes recteurs faibles dans l'expression de l'opinion en français et en espagnol

Juliette Delahaie, STL Université de Lille

Emmanuelle Canut, STL Université de Lille

Laurence Rouanne, Universidad Complutense de Madrid

Driss Sadoun, PostLab

Les parenthétiques (appelés « verbes recteurs faibles » par Blanche-Benveniste, 1989) sont présentés comme pouvant être utilisés dans deux types de constructions : en tête de construction, suivis d'une « que-phrase », ou en incise (médiane ou finale), après une séquence p (a. Je pense qu'il fait beau demain. b. Il fait beau, je pense, demain. c. Il fait beau demain, je pense). La proposition p apporte l'information principale, et le verbe parenthétique est souvent considéré comme ayant une fonction de modalisation et/ou d'atténuation. Ces constructions sont également considérées par certains chercheurs comme des « marqueurs discursifs propositionnels » (Andersen, 2007). Bolly (2010) donne même une définition des marqueurs parenthétiques qui diffère de celle de Blanche-Benveniste (1989) et d'Urmson (1952), où les marqueurs parenthétiques sont totalement apparentés à des marqueurs discursifs.

Notre objectif est d'étudier les constructions syntaxiques et les valeurs pragmatiques des verbes parenthétiques en français et en espagnol parlé, dans des conversations ordinaires portant sur l'expression d'une opinion sur un film/une série. Pour ce faire, nous adopterons le cadre analytique proposé par Blanche-Benveniste car il nous paraît le plus à même d'expliquer les constructions verbales liées à l'expression de l'opinion, comme l'a déjà montré González Ruiz (2015) pour l'espagnol.

Cette étude s'appuie sur les corpus constitués dans le projet DOC (Didactique Oral Corpus, Université de Lille, soutenu par le consortium Huma-Num). Ce projet vise à mettre à la disposition de la communauté des chercheurs des données comparables d'interactions orales dans différentes langues. Pour l'instant, sont accessibles (juin 2021) sur la plate-forme Ortolang des corpus oraux en français et en espagnol relevant de trois genres interactionnels différents et comportant plusieurs tâches discursives : des scènes de rencontre (42 enregistrements en français, 47500 mots ; 13 en espagnol, 5300 mots), des discussions autour d'un film/d'une série (32 enregistrements en français, 37500 mots ; 10 en espagnol, 5300 mots) et des interactions commerciales (22 enregistrements en français, 4300 mots ; en cours d'archivage pour l'espagnol).

Nous nous intéresserons dans cette communication aux verbes parenthétiques qui servent à l'expression de l'opinion, essentiellement dans les données corpus français et espagnol de discussions autour d'un film/série. Cette catégorisation fonctionnelle associée à une activité discursive permet de regrouper des formes verbales différentes, jusque-là non nécessairement associées à l'expression de l'opinion, et qui ont à première vue un fonctionnement syntaxique proche des verbes recteurs faibles.

On commencera par proposer une description syntaxique et sémantique des expressions du français *je pense/je trouve/je trouve pas*, mais aussi *je sais pas*, et *tu vois*. Par exemple :

1. mais justement je trouve en série ça aurait été super bien

2. et pareil Millie Bobbie Brown {Jeune actrice britannique } euh on a été habitué à mieux je trouve +[
 3. je sais pas +[pron=Sepa] c'est c'est bizarre
 4. tu vois en sortant je me suis dis ouais bon sans plus quoi
 5. je pense que pour eux c'est un très mauvais calcul
- (Extraits du corpus DOC)

À partir d'une analyse de fréquence quantitative et d'une description qualitative, d'ordre syntaxique et sémantico-pragmatique, nous montrerons que les formes en verbe faible + que p en français n'ont pas exactement la même valeur pragmatique et sémantique que les verbes sans que p. Ainsi, les formes *je sais pas* et *tu vois* sont très éloignées de leur équivalent en que p ou si p.

Nous proposerons pour finir une comparaison avec les verbes d'opinion en espagnol, tels que *creo* (*que*) et *me parece que* comme dans les exemples suivants tirés de DOC :

6. a mi me gustó porque vas en el personaje vas viendo qué tal está el amor entonces a mí me parece interesante
7. yo por eso creo que es buena porque cada personaje juega un papel importante y en cada episodio te hacen pensar que es un o otro el matador y al final no es así

Les verbes parenthétiques d'opinion de l'espagnol ont un fonctionnement syntaxique et des valeurs sémantico-pragmatiques très différentes de ceux du français. Tout d'abord, parce que la construction en que p en espagnol semble plus employée que la construction sans que p. Ensuite, parce que les verbes et locutions verbales sollicitées pour exprimer l'opinion n'ont pas les mêmes valeurs sémantico-pragmatiques : *creo que* et *me parece que* servent à modaliser et atténuer l'opinion du locuteur (Bonafont, 2018, pour *creo*), tandis que du côté francophone les locuteurs n'hésitent pas à affirmer leur prise de position personnelle en préférant employer *je trouve/je pense*, à la place de *je crois* (voir Gosselin, 2015 pour la distinction sémantique entre *je crois/je trouve/je pense*). De plus, si les locuteurs en français modalisent, avec *je sais pas* ou *tu vois*, c'est moins pour atténuer leur propos que pour introduire un point de vue conflictuel ou problématique pour l'interlocuteur. Se dessine ainsi à travers l'étude de ces verbes une conception de l'expression « polie » de l'opinion, au sens pragmatique du terme, radicalement différente en-deçà et au-delà des Pyrénées.

Bibliographie

- Andersen, Hanne Leth (2007), « Marqueurs discursifs propositionnels », *Langue française*, 154, 13-28.
- Bolly, Catherine (2010), « Pragmaticalisation du marqueur discursif tu vois. De la perception à l'évidence et de l'évidence au discours », in F. Neveu, V. Muni Toke, J. Durand et al. (éds.), *Congrès mondial de linguistique française*, 673-693. (<http://www.linguistiquefrancaise.org>)
- Bonafont, Amparo Soler (2018), « Fingimientos y atenuación en el uso de *creo* », *TILCE*, 34.3, 1104-1128.
- Blanche-benveniste, Claire (1989), « Constructions verbales < en incise > et rectification faible des verbes », *Recherches sur le français parlé*, 9, 53-74.
- González Ruiz, Ramón (2015), « Los verbos de opinión entre los verbos parentéticos y los verbos de rectificación débil : aspectos sintácticos y sémantico-pragmáticos », *CÍRCULO de Lingüística Aplicada a la Comunicación* 62, 148-173.
- Gosselin, Laurent (2015), « L'expression de l'opinion personnelle : « je crois/pense/trouve/considère/estime que P », *L'information grammaticale*, 144, 34-40.

Textes provisoires. Merci de ne pas citer.

Urmson, James Opie (1952), "Parenthetical Verbs", *Mind*, 61, 480-496.

Le masculin et le féminin dans la désignation des référents humains. Étude à partir de l'exploitation des corpus CFPP2000 et CFPB.

Anne Dister, Université Saint-Louis

Sonia Branca-Rosoff, CLESTHIA Université Sorbonne Nouvelle

L'asymétrie des usages du masculin et du féminin dans la représentation des référents humains fait partie des objets de polémique entre partisans et opposants de l'écriture dite inclusive. Pour les premiers, le genre masculin invisibiliserait les femmes, et ils préconisent de mentionner leur présence de manière explicite ; les seconds, quant à eux, considèrent que l'usage du masculin inclusif fait partie de l'économie de la langue et ils combattent l'ajout d'une couche idéographique compliquée à un système orthographique qu'il faudrait plutôt simplifier.

Déplaçant le problème vers l'observation des usages, nous souhaitons voir comment fonctionne l'oral non planifié, c'est-à-dire le médium que les sujets parlants traitent de façon quasi automatique, sans interférence de décisions militantes. Notre question est la suivante : que nous dit l'observation des corpus sur les usages du masculin et du féminin pour les noms référant à des humains ? Nous nous situons ainsi dans la perspective si bien ouverte par Claire Blanche-Benveniste de l'observation des données orales qui permet, par contraste, de mieux situer l'écrit. Nous soulignerons le clivage oral/écrit, notamment en morphologie, ce qui est un axe fort de l'approche développée par Claire Blanche-Benveniste, et nous nous demanderons comment fonctionne la catégorie du genre pour les noms correspondant aux humains.

Dans cette communication, nous analysons six entretiens issues du *Corpus de français parlé parisien* (CFPP2000, Branca *et al.* 2012) et du *Corpus de français parlé à Bruxelles* (CFPB, Dister et Labeau 2017), soit sept heures de parole. Le CFPB, dont l'entreprise a débuté huit ans après celle du CFPP, est conçu selon le même protocole d'enquête : l'entretien semi-dirigé de type labovien qui se base sur le même questionnaire, adapté à la situation bruxelloise. La grille d'entretien porte sur le rapport des habitants à leur quartier (arrivée, rapport au voisinage et aux commerces, changements récents...), sujets qui permettent à chacun de donner facilement son avis. Dans le sous-corpus, constitué d'entretiens réalisés par des intervieweuses avec une ou plusieurs locutrices, nous avons relevé systématiquement tous les noms associés à des référents humains. Notre objectif est de voir comment, dans de l'oral non planifié, se répartissent les genres grammaticaux qui désignent des humains, d'analyser quand les locutrices estiment devoir dire explicitement qu'il y a des femmes (ou seulement des hommes), et *in fine* de répondre à la question de la supposée invisibilisation de celles-ci par l'emploi du genre grammatical masculin.

A. Cinq catégories morphologiques

Nous avons distingué cinq grandes catégories de noms en ce qui concerne les marques de genre associées à des humains. Nous n'avons pas tenu compte des noms propres qui ont un comportement spécifique (Leroy 2004), à l'exception des gentilés et ethnonymes.

1) les noms marqués en genre pour lesquels on a une alternance du morphème dérivationnel

1. *tu vois et c'était genre une une Africaine qui sortait avec un Marocain* (B2-Sonia)
2. *et ensuite euh j'ai habité à Waterloo pendant quatre ans hum grâce à mon amoureux qui vivait là* (B3-Joëlle)
3. *pour moi la rive droite c'était les Halles auxquelles je me rendais la nuit avec une crémère + qui avait sa crèmerie à Nogent-sur-Marne* (P3-Denis)

2) les noms marqués en genre pour lesquels on a une alternance lexicale

4. *car mon arrière-grand-père Christophe ayant perdu son épouse a épousé une Wallonne* (B3- Maminou)
5. *et puis euh et puis mes mon p- mon père ainsi que sa soeur et ses deux frères sont venus là dans cette maison et mon père a passé toute sa vie euh professionnelle dans cette maison* (B3- Maminou)
6. *attends c'était qui encore des deux (?) c'était le type ou la femme qui avait l'alzheimer (?)* (B2-Sonia)
7. *après mes humanités j'ai eu l'occasion de rencontrer euh une jeune fille qui faisait de la sculpture* (B3-Maminou)
8. *la boulangère + l'autre boulangerie pas la boulangère avec le monsieur barbu* (P3E-Sonia)
9. *donc c'est c'est assez curieux comme contraste entre les mamies permanentées + avec leurs caddies et les touristes qui descendent des tuk-tuks* (P3-Denis)
10. *c'est arrivé plusieurs fois que des gens se + flippent surtout des filles pas pas des mecs* (P2-Guylaine)

3) les noms épïcènes, dont le genre est ou non marqué par un accompagnateur

11. *mais quand j'étais une jeune adulte ça m'était totalement égal* (B3-Joëlle)
12. *mais c'est je croyais que c'était ce photographe au coin d'abord* (B3-Maminou)
13. *au tout début euh j'allais voir la pédiatre ici en face* (B1-Carine)
14. *à part le fait qu'il y a mon médecin là-bas [rire] sinon euh* (P1-Mathieu)

4) les noms collectifs, soit féminins soit masculins, qui renvoient à des groupes mixtes

15. *mais il y a il y a une grosse communauté turque à Saint-Josse il y a il y a une énorme communauté marocaine sur plusieurs euh quartiers de Bruxelles et tout euh* (B3-Joëlle)

16. *et donc euh tous ces gens-là étaient avaient un peu le même profil que nous (rire) (B3-Maminou)*
17. *Florent était à la mer avec mes parents j'étais sortie la veille (B1-Carine)*
18. *à Bagnolet y a pas beaucoup de police on la voit jamais la police + et en fait c'est normal qu'on la voit pas parce que visiblement le maire avait des arrangements avec euh + les jeunes des cités (P2-Guylaine)*
19. *et comme mon père il a un commerce tu vois là lui il voit blindé que la clientèle a changé (B2-Sonia)*

5) les noms marqués en genre, qui peuvent renvoyer indifféremment à un homme ou à une femme

20. *c'est une personne âgée (?) (B2E-Sarah)*
21. *parce qu'il y a écrit dix euros de l'arc de triomphe à la tour Eiffel c'est dix euros par personne et ça personne le sait quand il monte dans le tuk-tuk (P3-Denis)*
22. *y a y a des gens qui ont dû redescendre du bus parce qu'il y avait trop de monde et que sous prétexte qu'un + qu'une personne handicapée avait la priorité pour accéder à bord ça a été une émeute (P3-Denis)*
23. *euh on avait pris le thermos pour aller euh devant Sainte-Gudule euh pour euh voir toutes les têtes couronnées défiler (B1-Carine)*
24. *elle était dans le clip et tout donc euh c'était une star tu vois quand on avait euh on était petits (B2-Sonia)*

B. Les emplois en discours : emplois exclusifs et emplois inclusifs

Dans les exemples que nous avons donnés ci-dessous des catégories 1 à 3, le genre grammatical des noms correspond au sexe des référents. L'emploi est dit *spécifique* ou encore *exclusif* : un féminin grammatical renvoie à une femme ou à des femmes uniquement, et un masculin grammatical renvoie exclusivement à un homme ou à des hommes.

On voit que ce n'est pas le fonctionnement des catégories 4 et 5, pour lesquelles il n'y a pas nécessairement de correspondance entre le genre grammatical et le sexe des référents. En effet, les noms collectifs renvoient à des groupes qui peuvent être mixtes, et les noms de la dernière catégorie peuvent renvoyer soit à des groupes mixtes (*les têtes couronnées*, dans notre exemple), mais aussi soit à un ou des homme(s) ou à une ou des femme(s) uniquement, et ce indépendamment de leur genre. C'est d'ailleurs dans cette catégorie que l'on trouve les traditionnels termes *sentinelle*, *estafette*, etc., mis en avant par les opposants à la féminisation des noms.

Si les termes de la 2^e catégorie, ceux pour lesquels le masculin et le féminin se forment par alternance lexicale, ne peuvent avoir que des emplois spécifiques, il n'en est pas de même

pour les catégories 1 et 3, qui ne sont pas enfermés dans ce fonctionnement. En effet, les locuteurs les utilisent en discours en neutralisant l'opposition homme-femme, dans des emplois au masculin, que celui-ci soit au singulier ou au pluriel :

25. *et on allait loger parfois chez l'habitant* (B3-Joëlle)
26. *non parce que vu que je travaille pas et tout et j'ai pas le temps pour euh faire un job d'étudiant et tout euh je vais pas m'habiller de marque quoi je vais au Zara* (B2-Sonia)
27. *déjà déjà la crèche c'est le parcours du combattant* (B1-Carine)
28. *c'est ultradur quand même de faire la passerelle parce que tu dois rattraper plein de matières et tout mais euh en trois ans après tu deviens juriste tu peux pas faire avocat tu vois* (B2-Sonia)
29. *c'était plutôt une ambiance d'école sympathique mais sans qu'il y ait euh + le maître de référence qui va te bouleverser* (P1E-Sonia)
30. *peut-être que l'occasion fera le le le larron* (P3-Christine)
31. *il y a il y a une vraie communauté là effectivement mais ça se sont plutôt des nouveaux arrivants j'ai l'impression de des des jeunes néerlandophones qui s'installent à Bruxelles et qui euh et qui font euh qui qui qui font bien vivre aussi Bruxelles culturellement parce qu'ils sont très dynamiques etcetera et donc euh il y a il y a une chouette vie néerlandophone jeune dans ces quartiers-là à Bruxelles dans mon quartier c'est plutôt des vieux flamands* (B3-Joëlle)
32. *il y a eu hein au tout début quand on est arrivé ici il y avait la fête des voisins* (B1-Carine)
33. *les enseignants seront les seuls d'ailleurs à travailler jusqu'à l'âge de septante-cinq ans* (B3-Maminou)
34. *ben les copains qui allaient à Maisons- à Créteil Soleil ou autre quoi enfin ou ou faire des billards dans le bar du coin enfin des choses comme ça je j'avais envie d'autre chose* (P1-Elisa)
35. *beaucoup d'amis de mes amis d'enfants de mes amis pardon sont aujourd'hui en colocation sont en + vivent d'échanges pour les vacances et autre* (P3-Denis)

C. Les reprises pronominales

Un nom de genre grammatical marqué est en général repris par un pronom de même genre et les accords respectent également les genres :

36. *ma mère est ma mère est née je crois dans le bas de Saint-Gilles euh possiblement rue de Mérode en tout cas elle a vécu là* (B3-Maminou)
37. *et comme mon père il a un commerce tu vois là lui il voit blindé que la clientèle a changé* (B2-Sonia)

Quand le genre n'est pas marqué grammaticalement, la reprise pronominale est quant à elle toujours marquée en genre :

38. *ok alors tu m'as parlé de tes parents euh ils étaient italiens et hum enfin ils sont de quelle origine enfin plus précisément (?)* (B2E-Sarah)

Or, on observe des cas de distorsion, en particulier avec un nom général comme *personnes*, grammaticalement féminin, repris par *ils* morphologiquement masculin, le masculin étant traité comme un neutre sur le plan sémantique de l'opposition de sexe.

39. *c'était des personnes âgées qui avait ça donc dans la rue ça sentait le hum ++ dès qu'on arrivait chez mes grands-parents ça sentait la la pâtisserie + ça sentait la tarte aux pommes et on passait notre vie avec mon frère dans ce ce ce truc-là donc on y allait ils nous connaissaient très bien* (P2-Guylaine)

D. Ils non introduit

Les données orales comportent de nombreux *Ils* non introduits. Ceux-ci suivent une séquence permettant de restituer un référent à valeur collective indéterminée

40. Sonia : *ce cimetière de Picpus tu y allais*
Matthieu : *j'y suis allé euh au moi- en fait j- j- s- j'ai dû y aller quand j'étais tout petit*
Sonia/Matthieu : *[1] mm [c'est l'école qui vous emmenait non [2] très très vague souvenir*
Mathieu : *je sais même pas [pause] peut-être peut-être avec mon père ça je je sais pas euh*
Sonia/Matthieu : *[1] ils ouvrent [2] mais*
Sonia : *une f- une ou deux fois par an* (P1E-Sonia)
41. *et entre autre les gardiennes d'écoles hein qui vous disent il vaut mieux être sans papier et logé dans un hôtel social parce qu'au moins quand ils vous expulsent ils vous relogent* (P3-Christine)
42. Sonia : *alors vous avez + d'autres utilisations de ces moyens franchisés le logement les échanges d'appartements des choses comme ça ou c'est juste pour le transport le XXX quoi* (P1E-Sonia)
43. – *ah non ça non non ça j'aime pas non -fin jusqu'à présent j'ai pas eu l'occasion et j'ai pas eu envie parce que bon on avait cherché pour New-York par exemple et je trouvais que les photos qu'ils nous montraient avec les les couettes les housses de couettes déjà avachies euh ça donnait pas très envie* (P3-Christine)

Dans ces exemples, les locutrices n'ont pas besoin de distinguer entre hommes et femmes. *Ils* renvoie indifféremment aux fonctionnaires de la mairie, ou aux New-Yorkais qui louent leurs appartements sans s'intéresser à leur identité sexuelle laquelle n'a aucune pertinence dans ce contexte.

E. Passage d'une forme neutralisée à un genre morphologiquement spécifié

44. *il y a un prof qui m'a vraiment + enfin que j'ai vraiment beaucoup beaucoup aimé + c'est une prof de + de français en quatrième (P1-Matthieu)*

On peut rapprocher ces va-et-vient entre formes englobantes (inclusives, dans laquelle l'opposition de sexe est neutralisée) et formes déterminées (exclusives) avec des exemples de divergences portant sur le nombre. Dans 45, on passe d'un singulier (*l'ami Picard*) à un pluriel (*eux*). La représentation mentale d'une entreprise associable à Picard permet d'évoquer les ouvriers qui y travaillent (*cf.* sur la construction des représentations mentales permettant l'anaphore Kleiber 1984) :

45. *d'abord notre ami Picard c'est le plus fidèle donc avec eux on n'est jamais déçus surtout quand on a de on a de gros problème de d'arthrose des mains ils vous pèlent les et vous nettoient les légumes d'une façon admirable (P3-Christine)*

La catégorie sexuée n'est pas toujours traitée comme pertinente par les locuteurs. Les relevés que nous présenterons montrent la fréquence avec laquelle elle est neutralisée, en particulier grâce aux épécènes et grâce à l'emploi du masculin pluriel, ou de pronoms. Ils montrent aussi que les dimensions de sexe peuvent être facilement réintroduites au fil du discours quand le besoin s'en fait sentir.

Bibliographie

- ALPHERATZ (2018). « Français inclusif : conceptualisation et analyse linguistique », <https://hal.archives-ouvertes.fr/halshs-02128443/document>
- BLANCHE-BENVENISTE Claire et MARTIN Philippe (2010), *Le français. Usages de la langue parlée*, Leuven-Paris, Peeters.
- BRANCA-ROSOFF Sonia, FLEURY Serge, LEFEUVRE Florence et PIRES Mat (2012), *Discours sur la ville Corpus du Français parlé parisien des années 2000, CFPP2000*. <http://cfpp2000.univ-paris3.fr/CFPP2000.pdf>
- CAPPEAU Paul et SCHNEDECKER Catherine (2014), « *Gens, personne(s), individu(s)*. Trois saisies de l'humain », 4^e *Congrès mondial de linguistique française*, Berlin, 19-23 juillet 2014, pp. 3027-2040. https://www.shs-conferences.org/articles/shsconf/pdf/2014/05/shsconf_cmlf14_01274.pdf
- CREISSELS, Denis (1999) « 'Genres' indo-européens et 'classes nominales' Niger-Congo ». in : *Faits de langues* 14. pp. 177-184. DOI : [10.3406/flang.1999.1280](https://doi.org/10.3406/flang.1999.1280)
- DISTER Anne et LABEAU Emmanuelle (2017), « *Le Corpus de français parlé à Bruxelles : origine, hypothèses et développements* », *Cahiers de l'AFLS* (Association for French Language Studies).
- GRINEVALD Colette. Typologie des systèmes de classification nominale. In: *Faits de langues*, n°14, Octobre 1999. La catégorisation dans les langues. pp. 101-122

- JOHNSEN Laure Anne (2019), *La sous-détermination référentielle et les désignateurs vagues en français contemporain*, Berne, Peter Lang.
- KLEIBER Georges (1994), *Anaphores et pronoms*, Louvain, Duculot.
- KLEIBER Georges (1987), *Rencontres avec la généricité*, Recherches linguistiques, CREM Editions-Université de Lorraine.
- KLEIBER Georges et VASSILIADOU Hélène (2012), « Histoire(s) de personnes : Qui est *je* ? Qui est *tu* ? Qui est *il* ? », dans *Les Cahiers philosophiques de Strasbourg*, N° 31, pp. 25-54.
- KLEIBER Georges et LAZZARO Henri (1987), « Qu'est-ce qu'un syntagme nominal générique ? Ou *Les carottes qui poussent ici sont plus grosses que les autres* » *Rencontres avec la généricité*, Recherches linguistiques, Metz, Crem-édition.
- LEROY Sarah (2004), *Le nom propre en français*, Paris, Ophrys.
- MATHIEU Cécile (2007), « Sexe et genre féminin : origine d'une confusion théorique », *La Linguistique* 2007/2, vol. 43, pp. 57-72.
- Vengoecha (de) Consuelo, *Catégorisation lexicale en muiname. Amazonie colombienne*, Thèse non publiée, Université Toulouse 2 Le Mirail, 2012.
<https://tel.archives-ouvertes.fr/tel-00796760/document>

Prolongements de la notion de paradigme : outil descriptif et théorique au-delà du français parlé

Otti Duvallon, Inalco SEDYL
Il-Il Yatziv-Malibert, Inalco CERMOM

1. Introduction

Les outils descriptifs élaborés par le Groupe aixois de recherche en syntaxe (GARS) pour l'étude des données orales ont permis d'aborder l'analyse grammaticale des textes sans découpage préalable (Blanche-Benveniste *et al.* 1990). La particularité de cette approche est de donner à voir les procédés par lesquels les locuteurs mettent en place des constructions syntaxiques dans leurs productions en situation de récit ou de conversation.

Les travaux du GARS ont

- 4) d'une part, mis en exergue la réalité des données orales non standard (amorces de mots, énoncés inachevés, répétitions, retouches, etc.), qui sont traditionnellement considérées comme une matière impropre à l'analyse grammaticale, et le problème de découpage, notamment la non-opérationnalité de la notion de phrase
- 5) et, d'autre part, contribué à la réflexion sur la syntaxe de la langue parlée, datant des années 1970-80 et développée dans le cadre de la linguistique fonctionnaliste discursive (*discourse-functional linguistics*) sous la forme d'une hypothèse concernant une
 - structure argumentale préférée » dans le discours oral (Givón 1975 ; Du Bois 1987 ; Lambrecht 1987).

La solution originale apportée par l'équipe du GARS dès la fin des années 1970 a trait à la contrainte de linéarité de la chaîne parlée et à l'aspect temporel de la production des énoncés. Fondées sur la théorie de l'approche pronominale (Blanche-Benveniste *et al.* 1987) et s'inspirant des séries paradigmatiques virtuelles évoquées par F. de Saussure (1966[1916]) et R. Jakobson (1956), les analyses en « grilles » amènent à la visualisation pluridimensionnelle des données (Blanche-Benveniste 1990).

Cette communication se propose de se pencher sur les prolongements qu'ont connus les outils élaborés par le GARS au-delà du français parlé. Le modèle d'analyse syntaxique en grilles » a pu être appliqué avec succès à d'autres langues, aussi éloignées les unes des autres que le sont par exemple le finnois (langue finno-ougrienne, de type agglutinant) et l'hébreu moderne (langue sémitique). L'objectif sera de montrer que la distinction fondamentale entre deux axes de déroulement de la chaîne parlée, syntagmatique et paradigmatique, auxquels s'ajoute le plan décroché des séquences incidentes, permet non seulement de mettre en lumière le caractère organisé et structuré des textes oraux et l'importance qu'ont les reformulations, les reprises et les parenthèses dans la construction du sens, mais constitue aussi une base solide et pertinente pour la comparaison des langues typologiquement différentes. Il s'agit là d'une vision de la grammaire non tributaire des catégories prédéfinies, les données, y compris les phénomènes de productions, étant prises en considération avec la logique qui leur est propre, comme la matière première pour l'observation des organisations syntaxiques.

Cette approche se caractérise comme non fonctionnaliste en ce sens qu'elle procède par une description fine des données en termes grammaticaux, syntaxiques et distributionnels sans chercher à avancer des motivations d'ordre pragmatique ou cognitif pour expliquer pourquoi les formes sont telles qu'elles sont. Cela donne la possibilité d'envisager, en parallèle avec l'analyse grammaticale, une analyse interactionnelle des données conversationnelles, cette dernière s'intéressant, entre autres, aux principes qui guident le changement de locuteur dans les situations d'interaction quotidiennes (Duvallon & Routarinne 2005 ; Yatziv-Malibert 2003, 2013).

2. Une vision dynamique de l'axe paradigmatique

L'originalité des travaux du GARS réside dans la distinction de deux axes de déroulement pour la chaîne parlée (ou écrite), l'axe syntagmatique et l'axe paradigmatique. On peut cependant constater un certain déséquilibre entre ces concepts. La description de l'axe paradigmatique se fait à l'aide des termes métaphoriques qui ont des connotations négatives. Dans les cas comme (1) et (2) :

5) *hier hier euh mon mari et moi sommes allées euh dans un dans - -* (Blanche-Benveniste et al. 1990 : 18)

hier
hier euh mon mari et moi sommes allées euh *dans un*
dans - -

(2) *ce qui était fantastique dans ce dans ce camping enfin dans ce cet hôtel c'est qu'on était carrément en face du Kilimandjaro* (Blanche-Benveniste 1997 : 17)

ce qui était fantastique *dans ce*
dans ce camping
enfin dans ce
cet hôtel *c'est qu'on était carrément en face du K.*

il est considéré que le déroulement syntagmatique est *brisé* en plusieurs endroits. La succession des éléments, identiques ou non (*hier hier ; dans ce dans ce camping - - dans ce cet hôtel*), qui instancient la même place syntaxique est interprétée comme un *piétinement* sur une seule et même place (Blanche-Benveniste et al. 1990 : 18-19). On parle aussi d'*entassement* d'éléments paradigmatiques (Blanche-Benveniste 1997 : 17).

On peut comprendre le choix de ces termes (*briser, piétinement, entassement*) en les situant dans le contexte où il s'agissait de trouver, pour les données orales, une présentation visuelle commode et lisible, répondant aux besoins d'une étude grammaticale. En effet, les analyses en grille étaient conçues dans le but de pouvoir intégrer tous les éléments de la chaîne parlée dans l'étude grammaticale, sans en modifier l'ordre, sans avoir recours à des notions telles que l'ellipse.

En dehors de ce cadre initial, la description de l'axe paradigmatique s'est montrée problématique. L'idée de piétinement sur un emplacement syntaxique a été assimilée à la suspension de la temporalité et de la progression du discours : les paradigmes, situés hors du temps, auraient un caractère statique, immobile (Traverso 2017).

Il s'agit là, nous semble-t-il, d'un malentendu. Nous donnons ici quelques exemples pour montrer que l'axe paradigmatique constitue une ressource grammaticale dynamique, non figée que les interlocuteurs exploitent dans la gestion de l'interaction.

Les deux premiers exemples en hébreu moderne viennent démontrer une progression de type « dialogique » dans un texte qui serait défini plutôt comme monologique. Ce qui donne la coloration dialogique ainsi que l'allure d'avancement à ces extraits est l'organisation en paradigmes, organisation collaborative tissée entre L1 et L2. Les paradigmes sont tout sauf figés, ils contribuent à la progression du monologue « dialogique ».

Dans l'exemple 1, à la ligne 1 (voir l'analyse en « grilles » encadrée), L1 entame deux syntagmes, le premier contenant l'article défini *ha* et le substantif *kol* « tout », et le deuxième que l'on peut entendre comme le début de l'adverbe *sam* « là-bas ». Le nom *kol* est repris par la même locutrice, de la ligne 2 jusqu'à la ligne 4, d'abord dans une liste paradigmatique de recherche de mots, et à la ligne 4 dans un nouveau syntagme complet

apportant une précision lexicale, *kol ha-havtsaa* « toute la conférence ». À la ligne 5, la même locutrice avance le discours à la fois sur l'axe paradigmatique (ajout d'une marque de pluriel dans le syntagme et reprise de *fam* « là-bas ») et sur l'axe syntagmatique (*haju be-* « étaient en »). Ensuite, à la ligne 6, l'axe syntagmatique est complété par L2 (*be-anglit* « en anglais »). Cet ajout forme « la séquence maximale » et il est confirmé, sur l'axe paradigmatique, par L1 (ligne 7).

(3) [hébreu moderne]

L1 *ve-ha-kol fa- ko- ko- kol ha-havtsaa ha-havtsa-ot fam*
 et-DEF-tout là-b- tou- tou- tout DEF-conférence DEF conférence-PL là-bas
 et le tout là-b- tou- tou- toute la conférence (toutes) les conférences là-bas

haju be-
 être/Passé/3PL LOC

étaient en

L2 *be-anglit*
 LOC-anglais
 en anglais

L1 *be-anglit*
 LOC-anglais
 en anglais

(CorpAfroAs, HEB_IM_CONV_3_SP1_278-281, HEB_IM_CONV_3_SP2_043-044)

1	L1	<i>ve-ha- kol</i> et -DEF-tout et le tout	<i>fa-</i> là-b- là-b-			
2		<i>ko-</i> tou- tou-				
3		<i>ko-</i> tou- tou-				
4		<i>kol ha-havtsaa</i> tout DEF-conférence toute la conférence				
5		<i>ha-havtsa-ot</i> DEF conférence-PL (toutes) les conférences	<i>fam</i> là-bas	<i>haju</i> étaient	<i>be-</i> en	
6	L2				<i>be-anglit</i> LOC-anglais en anglais	
7	L1				<i>be-anglit</i> LOC-anglais en anglais	

La locutrice principale (L1) et l' « enquêtrice » (L2) construisent un « récit » commun, y participent activement et progressent sur les deux axes par l'interaction. À la question sur la langue des conférenciers (que l'on ne voit pas dans l'extrait), la réponse est apportée (à l'aide du complément prépositionnel après *be*) par l'enquêtrice et non pas par L1 qui est censée y répondre. Ce complément est confirmé par la L1.

Dans l'exemple 2, l'« enquêtrice » (L2) demande à la locutrice principale (L1) si elle travaillait tous les jours fériés. Suite à la réponse positive de L1 (qui doit lui paraître surprenante), elle commence à énumérer les jours fériés.

Par cette énumération se forme une liste paradigmatique des jours fériés, qui est initiée par l'enquêtrice aux lignes 4 et 6. À la ligne 5, la locutrice 1 tente d'introduire la locution prépositionnelle *xuts mi* « à l'exception de » et de contribuer ainsi à la liste paradigmatique mais l'enquêtrice continue à compléter la liste paradigmatique à la ligne 6. À partir de la ligne

7, la locutrice 1 prend en charge la prolongation de la liste. Ce n'est qu'à la ligne 14, par l'introduction de l'adverbe *rak* « seulement », qu'elle arrive à nuancer l'intitulé de la liste, *les jours fériés auxquels je travaille*, et à mentionner les exceptions, c'est-à-dire les jours fériés auxquels elle chôme, comme tout le monde.

Tout sauf du statique dans cet exemple qui illustre la création dynamique d'une liste dans le cadre d'un contexte plutôt monologique, dans lequel l'enquêtrice joue un rôle de coproductrice de liste.

(4) [hébreu moderne]

L2 *klomaʁ ba-xofaf-im be-kol ha-xofafim at ovedet ?*
 c'est-à-dire LOC/DEF-congé-PL LOC-tout DEF-congé-PL 2SG.F travailler/Présent/SG.F
 c'est-à dire, pendant les congés, pendant tous les congés tu travailles ?

L1 *ken*

oui
 oui

L2 *xofef pesax*

congé Pâque
 le congé de la pâque juive

L1 *xuts mi-ju*

sauf du- jou

6) l'exception du jour...

L2 *xofef vof ha-fana*

congé du nouvel an
 le congé du nouvel an

L1 *be::: be-xol ha-moed sukot be::: e:::mm be-ruvim jom exad*

LOC LOC-deuxième jour férié Succot LOC HESIT LOC-fête de Pourim jour un
 le le deuxième jour de la fête de Succot le hum à la fête de Pourim un jour

be:xanuka be ::: xanuka anaxnu ovdim male ledaati
 LOC Hanoucca LOC Hanoucca 1PL travailler/présent/PL.M plein à mon avis

à la fête de Hanoucca à la fête de Hanoucca nous travaillons beaucoup à mon avis

be::: // vak be-favuot vak bi-flofet ha-vegalim
 LOC seulement LOC-la fête de Shavouot seulement LOC-trois jour férié/PL

le // seulement à la fête de Shavouot seulement lors des Trois Jours fériés

lo ovdim pesax lo ovdim
 NEG travailler/présent/PL.M Pâque juive NEG travailler/Présent/PL.M

on ne travaille pas (à) la Pâque juive on ne travaille pas

be xol hamoed ve-zehu

LOC deuxième jour (de la Pâque juive) et-c'est tout

le deuxième jour de la Pâque juive c'est tout

L2 *jo*

INTERJ
 ah bon ?!

L1 *ovdim kol ha-fana*

travailler/présent/PL.M tout DEF-année

on travaille toute l'année

(CorpAfroAs, HEB_IM_CONV_3_SP2_087-92, HEB_IM_CONV_3_SP1_513-536)

1 L2	<i>klomaḅ</i> c'est-à-dire c'est-à-dire	<i>ba-xofaf-im</i> LOC/DEF-congé-PL pendant les congés			
2		<i>be-kol ha-xofaf-im</i> LOC-tout DEF-congé-PL pendant tous les congés	<i>at</i> 2SG.F	<i>ovedet ?</i> travailler/PR/SG.F	
3 L1	<i>ken</i> oui oui				tu travailles ?
4 L2		<i>xofef pesax</i> congé Pâque le congé de la pâque juive			
5 L1	<i>xuts mi-ju</i> sauf du-jou à l'exception du jour				
6 L2		<i>xofef vof ha-fana</i> congé du nouvel an le congé du nouvel an			
7 L1		<i>be:::</i> LOC le			
8		<i>be-xol ha-moed sukot</i> LOC-deuxième jour férié Succot le deuxième jour de la fête de Succot			
9		<i>be::: e:::mm</i> LOC HESIT le hum			
10		<i>be-ruvim</i> LOC-fête de Pourim à la fête de Pourim		<i>jom exad</i> jour un un jour	
11		<i>be:xanuka</i> LOC Hanoucca à la fête de Hanoucca			
12		<i>be ::: xanuka</i> LOC Hanoucca à la fête de Hanoucca	<i>anaxnu</i> 1PL	<i>ovdim</i> travailler/PR/PL.M	<i>male</i> plein
13		<i>be::: //</i> LOC le			à mon avis nous travaillons beaucoup à mon avis
14	<i>vakbe-favuot</i> seulement seulement	LOC-la fête de Shavouot à la fête de Shavouot			
15	<i>vak</i> seulement seulement	<i>bi-flofet ha-vegalim</i> LOC-trois jour férié/PL lors des Trois jours fériés	<i>lo</i> NEG	<i>ovdim</i> travailler/PR/PL.M	on ne travaille pas
16		<i>pesax</i> Pâque juive (à) La Pâque juive	<i>lo</i> NEG	<i>ovdim</i> travailler/PR/PL.M	on ne travaille pas
17		<i>be xol hamoed</i> LOC deuxième jour (de la Pâque juive) le deuxième jour de la Pâque juive			<i>ve-zehu</i> et-c'est tout c'est tout
18 L2	<i>jo</i> INTERJ ah bon ?!				
19 L1			<i>ovdim</i> travailler/PR/PL.M on travaille	<i>kol ha-fana</i> tout DEF-année toute l'année	

Les deux exemples suivants, qui sont en finnois, illustrent une situation d'interaction différente, avec un interlocuteur plus passif. Ces extraits permettent de voir que les développements paradigmatiques sont aussi un moyen de stimuler l'interaction.

L'exemple (5) commence par un énoncé de L1 (ligne 1) qui est introduit par une particule-conjonction explicative (*ku* « parce que ») pour le situer à la suite des tours de parole précédents (qui ne sont pas montrés ici). Il est construit autour du verbe *kiittää* « remercier » qui régit un complément au cas partitif, instancié par le SN *kaikkia naisia* « toutes les femmes ». L'interlocutrice (L2) réagit à cet énoncé de façon minimale et neutre, en utilisant une particule - la forme rédupliquée, *joo joo* « ah bon » (ligne 2), après quoi L1 produit une séquence comparative (« comme il est d'usage », ligne 3) qui constitue un prolongement syntagmatique de son énoncé de la ligne 1. L2 continue de réagir de façon minimale, mais cette fois, elle emploie une autre particule, *niin* « oui » (ligne 4), qui sert à signaler qu'elle s'attend à ce que L1 poursuive le développement de son idée. Le tour de parole suivant de L1 (lignes 5-7 de l'analyse en « grilles ») est composé de trois SN au cas partitif (*Hanna-a ja äiti-ä-nsä ja ke-tä kaikk-i-a-nsa* « Hanna et sa mère et je ne sais qui ») qui sont en rapport paradigmatique avec la première instanciation (ligne 1) du complément du verbe *kiittää* « remercier ». Ce n'est qu'après ce développement paradigmatique du premier énoncé de L1 que L2 participe à l'interaction avec une réaction plus vive, *iha totta* « vraiment ? » (ligne 8).

(5) [finnois]

L1 *ku se esipuhe-ssa-an kato kiittä-ä kaikk-i-a nais-i-a*
 parce.que 3SG préface-INE-POS3 PTCL remercier-3 toute-PL-PAR femme-PL-PAR
 parce que dans sa préface tu vois il remercie toutes les femmes

L2 *joo joo*
 PTCL PTCL
 ah bon

L1 *kuten tapa-na on*
 comme coutume-ESS être.3
 comme il est d'usage

L2 *niin*
 PTCL
 oui

L1 *Hanna-a ja äiti-ä-nsä ja ke-tä kaikk-i-a-nsa*
 Hanna-PAR et mère-PAR-POS3 et qui-PAR toute-PL-PAR-POS3
 Hanna et sa mère et toutes ses [je ne sais] qui

L2 *iha totta*
 INTENS vrai-PAR
 vraiment ?

(Projets d'été 079/conversation téléphonique)

1 L1	<i>ku</i>	<i>se - - kiittä-ä kaikk-i-a nais-i-a</i>	
		parce.que 3SG remercier-3 toute-PL-PAR femme-PL-PAR	
		parce que - - il remercie toutes les femmes	
2 L2	<i>joo joo</i>		
	PTCL PTCL		
	<i>ah bon</i>		
3 L1			<i>kuten tapa-na on</i> comme coutume-ESS être.3 comme il est d'usage
4 L2	<i>niin</i>		
	PTCL		
	<i>oui</i>		
5 L1			<i>Hanna-a</i> Hanna-PAR Hanna
6		<i>ja</i>	<i>äiti-ä-nsä</i> mère-PAR-POS3
		et	sa mère
		et	
7		<i>ja</i>	<i>ke-tä kaikk-i-a-nsa</i> qui-PAR toute-PL-PAR-POS3
		et	toutes ses [je ne sais] qui
	et		
8 L2	<i>iha totta</i>		
	INTENS vrai-PAR		
	<i>vraiment ?</i>		

Il est à noter que dans cette liste paradigmatique, le lien syntaxique entre le verbe recteur *kiittää* « remercier » et les instanciations à distance de son complément est explicitement marqué par le suffixe partitif dont sont porteurs tous les SN aux lignes 5-7.

Dans l'exemple 6, L1 commence une séquence narrative par une liste de trois noms de nationalité (à la forme nominative) qui mettent en attente un verbe recteur (lignes 1-3). Cependant, au lieu de poursuivre son énoncé selon l'axe syntagmatique, L1 produit une séquence parenthétique interrogative (ligne 4, « ou faut-il dire un citoyen de la CEI ¹ ? ») qui porte sur le choix du dernier terme de la liste (« russe » ou « citoyen de la CEI »). Cette interrogation et son prolongement à la ligne 6 (« non mais à l'époque c'était la Rus(sie encore) ») peuvent être vus comme une stratégie à laquelle L1 a recours afin d'inviter son interlocutrice à participer plus activement à la réception du récit (ici une blague). Une fois que L2 a confirmé le choix initial du dernier terme de la liste (« russe », ligne 5), L1 poursuit son énoncé qui a commencé à la ligne 1 selon l'axe syntagmatique (lignes 6-7).

(6) [finnois]

L1	<i>italiala:ine r:anskala:ine ja venä:läinen,</i>	
	Italien Français et Russe	
	un Italien un Français et un Russe	
	^o > <i>tai pitä-ä=kö sano-a</i> < ^o <i>ivy:↑läine?</i>	
	ou falloir-3=Q dire-INF citoyen.de.la.CEI	
	^o >ou faut-il dire< ^o un citoyen de la CEI ?	
L2	<i>.mth</i>	
L1	<i>>^oei mut sillo ol-i Venä(jä viel)^o<=</i>	
	NEG.3 mais à.l'époque être-PRET.3 Russie encore	
	non mais à l'époque c'était la Rus(sie encore)	

38. La Communauté des États indépendants, fondée en 1991 après la dislocation de l'Union soviétique. L'exemple est tiré d'un enregistrement datant des débuts des années 1990.

- L2 = *ve↑nä:>läine<*,
Russe
un Russe
- L1 *nii;*, *ni< ol-i-vat taivaa-n-porti-lla ja*
PTCL PTCL être-PRET-3PL paradis-GEN-porte-ADE et
bon alors [ils] étaient aux portes du paradis et
siin jätkä-t juttel-i-vat Pietari-a odo:t-el-le-s°sa:-nsa°.
là-bas gars-PL discuter-PRET-3PL Saint.Pierre-PAR attendre-FREQ-INF-INE-POS3
ces gars étaient là en train de discuter en attendant Saint Pierre » (Blague, SR)

1 L1	<i>italiala:ine</i> Italien un Italien
2	<i>r:anskala:ine</i> Français un Français
3	<i>ja venä:läinen,</i> et Russe et un Russe »
4	<i>tai pitä-ä=kö sano-a ivy:↑läine?</i> ou falloir-3=Q dire-INF citoyen.de.la.CEI ou faut-il dire un citoyen de la CEI
5 L2	<i>.mth</i>
6 L1	<i>ei mut sillo ol-i Venä(jä viel)</i> NEG.3 mais à.l'époque être-PRET.3 Russie encore non mais à l'époque c'était la Rus(sie encore)
7 L2	<i>ve↑nä:>läine</i> Russe « un Russe »
8 L1	<i>nii;</i> , <i>ni ol-i-vat taivaa-n-porti-lla - -</i> PTCL PTCL être-PRET-3PL paradis-GEN-porte-ADE bon alors [ils] étaient aux portes du paradis - -

Dans tous ces exemples, les éléments qui relèvent de l'axe paradigmatique des emplacements syntaxiques contribuent à la progression de l'interaction. Les listes paradigmatiques ne suspendent pas la temporalité du discours. En revanche, la notion de paradigme permet de mettre en lumière le fait que les organisations syntaxiques qui sous-tendent le discours produit en situation d'interaction sont complexes et ne réduisent pas à la succession linéaire des unités composant la chaîne parlée.

3. Les particules discursives et la dimension paradigmatique des places de rection

Les travaux du GARS ont fait le choix de n'attacher *a priori* aucune interprétation particulière à la relation sémantique entre les éléments d'une liste paradigmatique. Comme on le voit dans les exemples cités, il peut s'agir d'une énumération additive, d'une relation de spécification, mais aussi de la mise en jeu des alternatives ou d'une opération de rectification. Il n'est pas toujours aisé de trancher entre ces options, car l'interprétation dépend de points de vue subjectifs, notamment de l'accès que l'on a aux contenus exprimés. La notion de paradigme est ainsi un outil théorique permettant de montrer que les phénomènes tels que la coordination, la recherche des « bonnes dénominations » et les phénomènes de réparation ont un fondement syntaxique commun.

Les locuteurs disposent cependant de moyens lexicaux pour organiser les liens entre plusieurs instanciations d'une même place syntaxique. L'équipe aixoise a proposé d'appeler *joncteurs* ces unités qui indiquent le type de jonction entre les éléments du même ordre. L'exemple le plus typique est sans doute le coordonnant *et* en français. On observe son équivalent finnois *ja* dans les exemples (5) et (6).

Les unités qui s'emploient en tant que joncteurs de liste ont aussi d'autres emplois, et on peut alors les considérer comme extérieures au principe même de l'instanciation d'une place syntaxique régie par un élément constructeur (verbal ou autre). Plus généralement, les joncteurs font partie d'une classe lexicale hétérogène dont les limites sont floues, qui a été diversement nommée comme *particules* ou *adverbes*, qui est souvent définie à l'aide de critères négatives : il s'agit d'unités invariables qui ne sont ni des verbes ni des noms (au sens large).

L'équipe du GARS ne s'est pas particulièrement intéressée à cette catégorie lexicale (Blanche-Benveniste et al. 1990 : 14). Mais il se trouve que la vision sur la syntaxe verbale selon les axes syntagmatique et paradigmatique permet de fonder la description d'une sous-partie des adverbes ou particules.

Dans le cadre de la linguistique française, Henning Nølke (1983, 2001) a étudié les unités telles que *même*, *aussi*, *surtout* et *seulement* et plus particulièrement leur emploi en tant qu'*adverbes paradigmatiques*.

39. *Même Pierre est venu.*

D'après Nølke, la principale caractéristique syntactico-sémantique des adverbes paradigmatiques, comme *même* dans (5), est d'introduire dans l'interprétation de l'énoncé une présupposition sur l'existence d'un paradigme. L'adverbe *même* présente l'élément constituant son foyer (*Pierre*) comme résultant d'un choix fait entre les éléments d'un paradigme que l'interprète doit rétablir (Nølke 2001 : 280). La prédication est valable aussi pour les autres éléments du paradigme qui restent implicites, mais elle est particulièrement significative, d'une manière ou d'une autre, dans le cas de l'élément foyer.

En plus des joncteurs de liste, il existe donc des adverbes ou particules qui s'adjoignent à l'instanciation d'une place syntaxique pour indiquer que cette instanciation doit être mise en rapport avec d'autres éléments d'un paradigme même si ces derniers ne sont pas matérialisés dans l'énoncé.

En dehors de la linguistique française, ce type de particules a été décrit par Ekkehard König (1991) qui, s'appuyant sur les données en anglais et en allemand, a proposé de parler de particules *focalisantes* (*focus particles*). L'analyse de König partage les points essentiels avec celle de Nølke, mais elle en diffère par le fait de mettre l'accent sur l'aspect discursif fonctionnel des particules en question, c'est-à-dire sur le phénomène de focalisation. Pour König (1991 : 31-33), la focalisation opérée par une particule établit une relation entre l'élément focalisé et un ensemble d'alternatives (*set of alternatives*). Une particule focalisatrice additive, telle que *also* dans (8), inclut ces alternatives dans l'interprétation de l'énoncé en tant que valeurs possibles à la place de l'élément focalisé.

(8) *FRED also bought a new car.*

Nous prenons ici un exemple finnois pour illustrer comment la notion de paradigme peut éclairer le fonctionnement des particules focalisantes (pour plus de détail, voir Duvallon 2021, Duvallon & Peltola 2014). L'extrait (9) comporte deux particules qui nous intéressent ici, la particule enclitique *-kin* et la particule autonome *vaikka*.

(9) [finnois]

– *Mi-tä=s sitä täällä, kun siellä on rautatie.*
 Q-PAR=PTCL PTCL ici quand là -bas être.3 fer.chemin
 À quoi bon rester ici, alors que là-bas, on a le chemin de fer.

1) *Rauta=ko tie?*

fer=Qchemin

Le chemin qui est en fer tu veux dire ?

2) *Se, joka kulke-e Oulu-sta Kuopio-on...*

3SG REL aller-3 NP-ELANP-ILL

Celui qui va d'Oulu à Kuopio...

ja pääse-e si-tä myöten ulkoma-i-lle=kin,

et pouvoir.aller-3 3SG-PAR le.long.de externe.pays-PL-ALL=KIN

et en l'empruntant, on peut aller même à l'étranger,

vaikka Helsinki-in, jos tahto-o.

VAIKKA NP-ILL si vouloir-3

à Helsinki par exemple, si on veut.

(J. Aho, *Rautatie* [Le chemin de fer], 1884)

Dans un énoncé sans contexte, l'enclitique *-kin* est porteur de deux valeurs possibles, soit inclusive-additive (10a), soit scalaire (10b) :

44. [finnois]

Äiti=kintule-e.

maman=KINvenir-3

(a) Ma mère aussi viendra. (b) Même ma mère viendra.

Dans l'exemple (7), la construction SN=*kin* (*ulkomaille=kin* « même à l'étranger ») a pour contexte immédiat un énoncé exprimant une éventualité (« on peut aller [quelque part] »). La particule *-kin* indique que le constituant focalisé, instanciation du complément de destination, est à situer par rapport à un paradigme d'éléments, tous susceptibles de valider la relation prédicative au même titre que l'élément actualisé. Il s'ensuit que la particule adjointe à l'instanciation effective invite l'interlocuteur à ne pas s'arrêter à cet élément, mais à *parcourir* le paradigme des destinations possibles, en passant de destination en destination (pour la notion de parcours, voir par ex. Culioli 2018 [2002]). Le contexte est tel que les éléments du paradigme s'organisent sur une échelle d'accessibilité. Le fait d'instancier la relation prédicative avec un élément inattendu pour l'interlocuteur a pour effet d'illustrer l'éventail des possibilités.

La construction SN=*kin* est suivie d'une deuxième instanciation effective des destinations possibles, introduite cette fois par la particule *vaikka* « par exemple » (*vaikka Helsinkiin* « à Helsinki par exemple »). Il s'agit là d'une autre particule focalisante inclusive, qui s'utilise pour faire des suggestions. Sa fonction est d'indiquer que l'élément focalisé est une possibilité parmi d'autres (d'où la traduction proposée par « par exemple ») et que de plus, l'actualisation effective de cette possibilité dépend de la volonté de l'agent concerné. Ce deuxième aspect de la sémantique de *vaikka* est explicité ici par la séquence qui suit, *jos tahtoo* « si on veut ».

Dans cet exemple, les deux instanciations successives d'une même place syntaxique, introduites par deux particules focalisantes différentes, ne s'organisent pas directement en une liste paradigmatique car chaque instanciation est soumise à une modalité propre en ce qui concerne sa relation avec le paradigme implicite.

Nous terminerons par un exemple où il s'agit de construire une référence indéfinie. Haspelmath (1997) qui étudie les pronoms indéfinis de type « quelque chose / something » dans une perspective typologique a mis en lumière le fait que dans beaucoup de langues, ces expressions sont de nature composée. Un des modèles attestés consiste à combiner une base interrogative-relative avec une particule focalisante. C'est le cas du pronom finnois *jokin*

« quelque chose », formé de la base pronominale *jo-* et de la particule *-kin*.

L'exemple (11) illustre l'emploi de cette expression qui permet au locuteur d'instancier une place syntaxique sans contenu lexical tout en impliquant l'existence de dénominations possibles.

(11) [finnois]

Liisa (...) *toivott-i* *hyvä-ä* *huomen-ta.*
 NP souhaiter-PRET.3 bon-PAR matin-PAR

Liisa (...) dit bonjour.

Mutta toinen ei vastan-nut, ei huomannut-kaan,
 mais autre NEG.3 répondre-PPA NEG.3 remarquer-PPA=PTCL

Mais l'autre ne lui répondit pas, elle ne la remarqua même pas,

kulk-i ohitse ja pyöritt-el-i jo-ta-kin
 passer-PRET.3 devant et faire.tourner-FREQ-PRET.3 INDEF-PAR-KIN

elle lui passa devant en faisant tourner quelque chose

– *ol-is=ko se ol-lut avain vai mikä? – etusorme-nsa ympäri.*
 AUX-COND.3=Q 3SG être-PPA clé ou quoi index-POS3 autour

– était-ce une clé ou bien autre chose ? – autour de son index.

(J. Aho, *Rautatie* [Le chemin de fer], 1884)

Le pronom *jo-ta-kin* (à la forme partitive) « quelque chose » matérialise le complément d'objet du verbe *pyörittellä* « faire tourner ». Ce faisant, il indique qu'il existe une entité participant au scénario décrit en tant que « ce que l'on fait tourner », mais il ne la nomme pas. Par là-même, il établit un paradigme qui correspond à la classe des désignations lexicales possibles de cette entité.

L'instanciation pronominale indéfinie de l'objet est suivi d'une séquence parenthétique interrogative qui met momentanément en avant une possibilité sans la valider (« était-ce une clé ou bien autre chose ? »). Cette séquence montre indirectement que la référence construite par *jokin* « quelque chose » met en jeu un paradigme à parcourir. Pour le dire autrement, la référence indéfinie implique l'activation de toute une classe de catégorisations possibles.

Pour expliquer la présence de l'élément *-kin* dans *jokin* « quelque chose », on peut supposer qu'à l'origine, cette forme a fonctionné comme pronom de libre choix. Les expressions de libre choix ont un sens de type « n'importe lequel / anyone ». Leur caractéristique principale est de référer à un ensemble dans lequel tout élément est susceptible d'être sélectionné, mais aucun ne l'est effectivement. L'idée de « libre choix » (*freedom of choice*, Vendler 1962) est fondée sur une situation d'interaction où le locuteur laisse le choix à son interlocuteur (*Take any one of them!*).

Le passage d'un pronom de libre choix vers un pronom indéfini de type « quelque chose » peut être vu comme un affaiblissement du sens et une extension des contextes d'emploi. Ce que les pronoms indéfinis et les pronoms de libre choix ont en commun, c'est de mettre en jeu dans la référence qu'ils construisent un paradigme implicite des instanciations possibles d'une place syntaxique.

4. Conclusion

Nous avons soutenu dans cet exposé que la notion de paradigme est un outil descriptif et théorique dont l'intérêt ne se limite pas à une « préanalyse » syntaxique des données orales.

45. Elle permet de faire ressortir les fondements syntaxiques de la progression du discours en situation de conversation.

46. Elle est pleinement inscrite dans la temporalité du discours et la dynamique des

interactions.

46. Elle permet de proposer une analyse grammaticale du fonctionnement de certaines particules dites discursives.

47. Elle aide à mettre en lumière, dans le domaine de la référence nominale, le lien qu'il y a entre des phénomènes habituellement considérés séparément, tels que l'indéfinition, la focalisation et les réinstanciations (sous forme de listes) d'un même emplacement syntaxique.

Abréviations

Exemples en hébreu moderne

DEF	défini
HESIT	hésitation
LOC	locatif
NEG	négation
PR	présent
PL	suffixe pluriel
SG.F	suffixe singulier féminin
PL.M	suffixe pluriel masculin
1PL	première personne du pluriel
2.SG.F	2ème personne singulier féminin
3PL	3ème personne du pluriel

Exemples en finnois

ADE	adessif
ALL	allatif
AUX	auxiliaire
COND	conditionnel
ELA	élatif
ESS	essif
FREQ	suffixe fréquentatif
GEN	génitif
ILL	illatif
INDEF	indéfini
INE	inessif
INF	infinitif
NEG	verbe négatif
NP	nom propre
PAR	partitif
PL	pluriel
PPA	participe passé
PRET	prétérit
POS	suffixe possessif
PTCL	particule discursive
REL	pronom relatif
SG	singulier
Q	suffixe interrogatif
1	première personne
2	deuxième personne
3	troisième personne

Références bibliographiques

- BLANCHE-BENVENISTE Claire, 1990, « Un modèle d'analyse syntaxique "en grilles" pour les productions orales », *Anuario de Psicologia*, 47, pp. 11-28.
- BLANCHE-B ENVENISTE Claire, BILGER Mireille, ROUGET Christine, VAN DEN EYNDE Karel, 1990, *Le français parlé. Etudes grammaticales*, Paris, Éditions du CNRS.
- BLANCHE-BENVENISTE Claire, DEULOFEU José, STEFANINI Jean, VAN DEN EYNDE Karel, 1987, *Pronom et Syntaxe. L'approche pronominale et son application au français*, Paris, SELAF.
- CULIOLI Antoine 2018 [2002], « *Nous parlîmes, qui à droite, qui à gauche* », in A. Culioli, *Pour une linguistique de l'énonciation, Tome IV : Tours et détours*, Limoges, Lambert-Lucas, p. 179-207.
- DU BOIS John, 1987, « The Discourse Basis of Ergativity », *Language* 63, p. 805-855.
- DUVALLON Outi, 2021, « *-kin-liitteen semantiikasta ja diskursiivisista tehtävistä (On the semantics and discourse functions of the Finnish suffix -kin)* », *Journal of Estonian and Finno-Ugric Linguistics* 12:1, 41-78.
- DUVALLON Outi, Peltola 2014, « Les énoncés concessifs en finnois : hypothèse sur le passage du virtuel au réel », *Travaux linguistiques du CerLiCO* 25, 207-226.
- DUVALLON Outi, ROUTARINNE Sara, 2005, « Parenthesis as a resource in the grammar of conversation », in A. Hakulinen, M. Selting (eds.), *Syntax and Lexis in Conversation: Studies on the use of linguistic resources in talk-in-interaction*, Amsterdam, John Benjamins, p. 45-74.

- HASPELMATH Martin, 1997, *Indefinite Pronouns*, Oxford, Oxford University Press.
- HENNING Nølke, 1983, *Revue Romane* no supplémentaire 23 : *Les adverbes paradigmatiques : fonction et analyse*, Copenhague, Akademisk Forlag.
- HENNING Nølke, 2001, « Les adverbes paradigmatiques révisés : non sur tout mais surtout sur surtout », dans *Le regard du locuteur 2. Pour une linguistique des traces énonciatives*, sous la direction de Nølke Henning. Paris, Éditions Kimé, « Linguistique », 2001, p. 271-298.
- GIVÓN Talmy, 1975, « Focus and the scope of assertion: Some Bantu evidence », *Studies in African Linguistics* 6, p. 185-205.
- JAKOBSON, Roman, 1956, « Two aspects of language and two types of aphasia », in Roman Jakobson & Morris Halle, *Fundamentals of language*, The Hague : Mouton, p. 53-82.
- KÖNIG Ekkehard, 1991, *The Meaning of Focus Particles: A Comparative Perspective*. London, Routledge.
- LAMBRECHT Knud, 1987, « On the Status of SVO Sentences in French Discourse », in R. S. Tomlin (ed.), *Coherence and Grounding in Discourse*, Amsterdam, John Benjamins, p. 217-261.
- SAUSSURE Ferdinand de, 1966 [1916], *Cours de linguistique générale*, Paris, Payot.
- TRAVERSO VÉRONIQUE, 2017, « Déambulations autour du paradigme et de la temporalité de l'interaction », *Signata* 8, 129-144.
- VENDLER Zeno, 1962, « Each and every, any and all », *Mind* 71, 145-160.
- YATZIV-MALIBERT II-II, 2003, « L'axe saussurien oublié et sa contribution à l'analyse de la langue parlée », [en hébreu], *Hebrew-a Living Language* 3, 153-166.
- YATZIV-MALIBERT II-II, 2013, « Entre séquentialité narrative et coopération entre locuteurs » [en hébreu], in M. Florentin (ed.), *Studies in Language: Studies in Contemporary Hebrew and its sources*, Israel, The Academy of the Hebrew Language 14, p. 160-176.

Enrichir la description du français parlé grâce à de nouveaux corpus : les constructions [CV₁ Ø CV₂] avec les verbes *croire* et *penser* dans le MPF

Auphémie Ferreira, LATTICE Université Sorbonne Nouvelle

1. Introduction

Les verbes *croire* et *penser* ont fait l'objet de multiples analyses portant sur leurs caractères sémantiques, pragmatiques et syntaxiques en français (Blanche-Benveniste, 1989, Andersen, 1999) et dans d'autres langues tels que l'anglais (Urmson, 1952, Thompson et Mulac 1991a et b, Bolinger 1968), l'espagnol, l'italien (Schneider, 2007) ou le portugais (Haßler, 2014).

Dans leurs travaux, Blanche-Benveniste (1989), Blanche-Benveniste et Willems (2007, 2016) et Willems et Blanche-Benveniste (2010, 2014), en s'appuyant sur des données de français parlé, ont mis en avant les spécificités syntaxiques de ces verbes qu'elles regroupent sous l'appellation « verbes faibles ». Elles décrivent la configuration des verbes tels que *croire* et *penser* et dégagent les propriétés syntaxiques caractéristiques de leurs emplois en recteurs faibles dont leur triple réalisation (suivi d'une *que*-phrase, en incise et disjointe) illustrée en (1) ;

- (1) je crois qu'un certain nombre de gens ont un peu peur
ce sont des gens qui respectent la pharmacie je crois et qui veulent être aussi respectés
L1 – tu crois que le français ça va se perdre pareil L2 – non je pense pas

[exemples issus du corpus de français parlé du GARS (Blanche-Benveniste, 1989 : 228)]

Des études récentes (Andersen, 1999, 2007 ; Avanzi, 2012 ou Gachet, 2012) relèvent l'existence d'une réalisation dans laquelle le verbe est en tête d'énoncé non suivi de l'élément *qu-* mais a une portée claire sur la construction verbale qui suit :

- (2) je crois elle est m'a même pas laissé deux jours [ofrom, unine1576m]

D'abord considérée comme « non existante » en français hexagonale dans les travaux de Blanche-Benveniste et Willems (2007 : 226, 230), cette construction est signalée par les mêmes autrices en 2016.

L'accès à de nouvelles données extraites du corpus MPF⁰, auquel nous avons contribué et que nous prendrons soin de présenter, permet de mener une analyse systématique sur de nombreux exemples comparables à (2) tels que l'énoncé présenté sous (3).

- (3) je pense je vais la prendre chez Ikea [MPF, Aristide6b]

Nous avons également identifié dans ce corpus, des réalisations qui semblent en apparence similaires, où le verbe *croire* ou *penser* est rattaché à une construction verbale sans élément *qu-*, mais qui présentent en réalité des propriétés différentes de celles relevées dans les études antérieures. Nous notons ainsi dans les exemples ci-dessous que le verbe n'est ni restreint à la première personne du singulier (4) ni au présent de l'indicatif (5).

⁰ <https://www.ortolang.fr/market/corpora/mpf>. Gadet et Guerin (2016) pour une présentation du corpus.

(4) le mec **il croit** il bosse au sénat [MPF_Zakia3]

(5) **j'ai cru** c'était Justin Bieber [MPF_JDB1b]

Nous nous intéressons aux exemples du type (2) à (5) que nous désignons, pour le moment, par le terme [CV₁ Ø CV₂] où CV₁ réfère à la construction verbale comprenant *croire* ou *penser* et CV₂ à la construction verbale qui suit.

Un des problèmes fondamentaux de ces verbes dans ce type de constructions a été formulé en ces termes par Blanche-Benveniste et Willems (2016 : 10) : « s'agit-il d'un cas d'ellipse du *que*, le verbe continuant à régir (faiblement) la P, ou faut-il plutôt les interpréter comme des cas de parenthétiques en position initiale ? ». Une observation minutieuse des constructions extraites du MPF, nous permet de participer au débat quant au lien de ces verbes à la construction verbale qui suit et de formuler une troisième hypothèse : les exemples (4) et (5) illustrent des cas d'une dépendance (forte) sans marque.

Nous précisons, dans une première partie, le cadre théorique choisis pour l'analyse de ces séquences⁰ et le corpus sur lequel nous avons travaillé. Nous rappellerons, dans un second temps, la définition de *verbe faible* proposée par Blanche-Benveniste et Willems et nous présenterons les différentes hypothèses formulées quant aux structures [CV₁ Ø CV₂] avec les verbes *croire* et *penser*. Enfin, nous formulerons nos hypothèses d'analyse pour ces structures. Nous concluons en réaffirmant la place essentielle des corpus de français parlé, et en particulier de données qui ont été recueillies dans des contextes de grande proximité, dans la description de la syntaxe du français.

2. Méthodologie

2.1. La micro et macro-syntaxe comme cadre d'analyse

Notre étude s'inscrit dans le cadre de l'Approche Pronominale (Blanche-Benveniste et *al.*, 1984) et son extension par la prise en compte des deux niveaux d'analyse que sont la micro-syntaxe et la macro-syntaxe (Blanche-Benveniste et *al.*, 1990 et 2010) telle que présentée dans Bilger et *al.* (2013). Ce cadre permet de décrire des séquences telles que [CV₁ Ø CV₂] sans pour autant parler d'ellipse d'élément *qu-*.

Le premier niveau d'analyse, la microsyntaxe, traite des relations de dépendances grammaticales. Blanche-Benveniste et *al.* (1984) repris par Bilger et *al.* (2013), montrent qu'un élément, verbal dans le cadre de la syntaxe verbale, peut exercer un certain nombre de contraintes sur un autre élément : ce phénomène est nommé rection. Des tests caractérisant les éléments régis (contraints), tels les tests d'insertion paradigmaticque, sont mis avant. Ces tests comprennent : la proportionnalité de l'élément régi avec une proforme, la possibilité d'extraire l'élément régi à l'aide d'une clivée (ou d'une pseudo-clivée notamment pour les constructions introduites par une conjonction de type *que*).

Nous illustrons l'utilisation de ces tests avec les énoncés regroupés sous (6) :

(6)Adelyne : Ils pensaient qu'on était ensemble [MPF_Jennifer2]

(6a) Adelyne : Ils *le* pensaient. Ils pensaient *quoi* ? qu'on était ensemble.

⁰ Suites d'éléments linguistiques soumis à l'analyse (Bilger et *al.* 2013).

(6b) *Ce qu'ils pensaient c'est qu'on était ensemble.*

S'ajoutent à ces tests, des manipulations concernant la portée des modalités de l'élément recteur sur l'élément régi : « L'élément inscrit dans le paradigme étant soumis à la modalité de relation du verbe, il pourra être opposé à un élément de ce même paradigme » (Bilger et al. 2013 : 70). Nous illustrons cette manipulation en l'appliquant, ci-dessous, à l'énoncé (6) :

(6c) Ils *ne* pensaient *pas* qu'on était ensemble *mais* qu'on était juste amis.

Le second niveau d'analyse, la macrosyntaxe, traite « [d]es relations qu'on ne peut pas décrire à partir des réactions de catégories grammaticales » (Blanche-Benveniste, 1990 : 113). Les dépendances macrosyntaxiques peuvent être définies par la négative : elles ne répondent pas aux contraintes qui s'appliquent à une relation microsyntaxique ; mais également par des propriétés positives telles que « la présence de modalités de construction (types de phrases) ou de modes d'énoncé (schémas prosodiques) contrastables et une liberté de composition en catégories ». Une dépendance à un niveau macrosyntaxique est désignée par le terme « association » par Blanche-Benveniste et al. (1990).

Suivant le cadre de Bilger et al. 2013, nous considérons que les deux composantes, micro- et macro-, interagissent.

2.2. Présentation des données

Nous avons travaillé sur le corpus MPF³ (Multicultural Paris French), corpus de français parlé en Ile-de-France qui correspond à environ 1 million de mots. Initialement, ce corpus vise l'étude de l'influence des langues de l'immigration sur le français pratiqué en région parisienne. L'ensemble du corpus est composé d'entretiens de proximité et de conversations naturelles entre pairs soit par auto-enregistrement de l'informateur (absence de l'enquêteur), soit que l'enquêteur fait partie intégrante du groupe. Une grande importance a été portée au lien et à la connivence enquêteur / enquêté (Gadet et Guerin, 2016 : 288). Les processus mis en œuvre pour la constitution de ce corpus permettant l'approche des pratiques langagières des jeunes de banlieue parisienne ont été présentés dans Gadet (2017). Nous souhaitons souligner que les structures [CV₁ Ø CV₂] ne semblent pas réserver aux locuteurs enregistrés pour le MPF, des occurrences ont été relevées dans le CEFC⁴ et dans des études antérieures.

2.3. Extraction des constructions

Après l'extraction de l'ensemble des verbes *croire* et *penser* à l'aide d'un concordancier, nous avons procédé à une classification syntaxique. Nous avons annoté [CV₁ Ø CV₂] les verbes qui sont directement suivis d'une construction verbale et « autre⁵ » les occurrences que nous avons choisies d'écarter de notre étude pour différentes raisons : le verbe constitue une amorce ou un abandon de construction, la qualité du son empêche l'analyse, l'absence de l'occlusive peut être confondue avec une agglutination malgré une lecture du spectrogramme, le verbe n'est pas lié à ce qui suit mais à ce qui précède (postposition). Nous avons relevé au total 52 occurrences du verbe *penser* et 123 du verbe *croire* en construction [CV₁ Ø CV₂].

³ <https://www.ortolang.fr/market/corpora/mpf> voir Gadet et Guerin (2016) pour une présentation du corpus.

⁴ [Corpus d'Etude du Français Contemporain \(CEFC\)](#), pour une présentation voir Debaisieux et Benzitoun (2020).

⁵ Cette analyse s'inscrivant dans le cadre d'un travail de recherche plus large, nous avons également annoté les constructions où le verbe en construction d'incise, disjointe ou suivi d'une *qu*-P.

3. Les verbes « faibles »

3.1. Présentation des travaux de Blanche-Benveniste et Willems.

Le terme de *verbe faible* a été proposé par Blanche-Benveniste (1989). Les analyses sur ces verbes ont été poursuivies dans Blanche-Benveniste et Willems (2007, 2016), Willems et Blanche-Benveniste (2010, 2014). Les chercheuses définissent une catégorie de verbe désignée par le terme de *verbe faible* ou *verbe recteur faible* en s'appuyant sur un ensemble de critères.

Les verbes faibles sont définis à la fois par une préférence en termes de personne, temps et mode : la première personne⁶ du singulier au présent de l'indicatif pour *croire* et *penser* :

« Un verbe comme « je crois bien », dans son emploi de recteur faible, est bloqué sur la personne « je ». « Je crois bien qu'il pleut » n'a pas pour équivalent : « tu crois bien qu'il pleut », ni « il croit bien qu'il pleut ». « Croire » est ici restreint à la fois dans sa modalité assertive, dans sa personne et dans son pouvoir rectionnel » (Blanche-Benveniste, 1989 : 66). Ces « blocages » (Blanche-Benveniste, 1989) sont associés à un affaiblissement du sens du verbe.

Les verbes faibles sont également caractérisés par un sens spécifique propre, distinct du verbe plein. Le sens faible « avoir l'impression », « à mon avis » est distingué du sens fort « accorder sa croyance », il s'agit dans ce dernier cas plutôt d'un emploi descriptif du verbe (le locuteur décrit sa croyance en quelque chose). Ce deuxième critère est directement lié au troisième critère.

La pronominalisation de leur « complément » en *le*, *cela* (*ça*) et *en* tout en conservant un sens faible est difficile voire impossible (Blanche-Benveniste, 1989 : 62). Ces verbes n'ont toutefois pas perdu toute leur capacité rectionnelle puisque leur complément sont équivalents à d'autres formes : *ce que* et *qu'est ce que* (Blanche-Benveniste et Willems, 2007 : 231-232 §4.2.1.1.).

Les verbes faibles constituent une famille de construction. Ils connaissent une triple réalisation (suivi d'une *qu*-P, en incise en milieu d'énoncé ou fin d'énoncé et en construction disjointe).

3.2. De l'existence des [CV₁ Ø CV₂]

L'existence des réalisations [CV₁ Ø CV₂] en français hexagonal⁷ a fait l'objet d'un débat. Dans les travaux de Blanche-Benveniste et Willems, l'existence de ces réalisations n'est reconnue que tardivement (Blanche-Benveniste et Willems, 2016). La supposition, selon laquelle il existerait un grand nombre de constructions avec effacement de *qu*-, est même rejetée par Blanche-Benveniste et Willems (2007 : 229-230). Ces constructions correspondaient, selon elles, à une rapide interprétation d'une incise postposée sans prise en compte d'un contexte suffisamment large ou une impression due à un phénomène d'agglutination. Les arguments de Blanche-Benveniste et Willems sont en partie formulés en opposition aux études

⁶ Elles relèvent quelques cas de verbes à la deuxième personne du singulier dans les structures interrogatives : « L1 : tu es tu penses que tu es plus maniaque à l'oral qu'à l'écrit L2 : beaucoup plus à l'écrit qu'à l'oral » (Blanche-Benveniste et Willems, 2016 : 1).

⁷ Des constructions [CV₁ Ø CV₂] ont par ailleurs été relevées en français hors hexagone : au Québec et à l'Ontario (Sankoff et Cedergren, 1971 ; Martineau, 1988 ; 1993), en Acadie (Wiesmath, 2002) mais aussi en Suisse et en Belgique (Avanzi, 2012 ; Gachet, 2012).

d'Andersen (1999 et 2007) qui, au contraire, relèvent de nombreux cas d'antéposition des verbes parenthétiques. Gachet (2012) et Avanzi (2012) confirment l'existence de ces réalisations en français, en s'appuyant notamment sur la l'écoute des enregistrements et une analyse prosodique minutieuse. Quelques énoncés semblent ambigus, comme en (7), (8), (9) ou (10) ;

(7) Ouais j'ai fait un entretien avec elle c'était à hmm **je crois** c'était à Neuilly ou quelque chose comme ça (.) c'était pas au sein du même euh c'était pas au sein du Monoprix c'était dans leur siège social mais il est pas dans c'était pas celui du quatre vingt douze. [MPF_Nacer4]

(8) Je sais pas ben en fait euh c'est euh **je pense** c'est les circonstances [PFC_Avanzi cité par Gachet 2012 p.224]

(9) quoi euh l'NNAAMMEE le veut pour le quinze du mois **je crois** je me suis fait relancer ce mois-ci le quinze [CEFC_reunions-de-travail > OF1_Reunion22Nov07]

(10) il n' y avait pas d' histoire euh en anglais on faisait euh **je crois** c' est deux heures euh deux heures en plus [CEFC_C-Oral-Rom > ffamd129]

La position de *je crois* dans l'énoncé (7) et (8) est ambiguë : s'agit-il d'un cas de [CV₁ Ø CV₂] ou bien d'une incise intraposée ? En effet, nous pourrions analyser « c'était » ou (22) c'est » dans ces énoncés comme un phénomène de bribes ou de « backtracking », c'est-à-dire le fait de répéter après l'incise un bout de la construction qui a été interrompu. Ce type d'ambiguïté a également été souligné par Gachet (2012 : 224) et Avanzi (2012).

En (9), sans indication prosodique, nous pouvons nous demander s'il s'agit d'un cas de verbe postposé (rattachement à gauche) ou de construction de type de [CV₁ Ø CV₂] avec un rattachement à droite.

Deux analyses sont en concurrence pour (10). La première voudrait que *je crois* corresponde à un cas d'en incise intraposée avec une greffe⁸ de l'énoncé « c'est deux heures en plus » sur le verbe « on faisait ». Ici, un groupe nominal de type « deux heures » était attendu, l'incise aurait favorisée ce phénomène de greffe d'une construction verbale. La seconde analyse ferait du verbe *je crois* la CV₁ d'une structure du type [CV₁ Ø CV₂] et de *c'est* la CV₂, le tout pouvant être également analysé comme un cas de greffe du verbe *faire*.

Les deux analyses peuvent être annotées ainsi sur l'énoncé (10) reproduit sous (11) :

(11) il n' y avait pas d' histoire euh en anglais on faisait euh [**je crois**] [c' est deux heures euh deux heures en plus]//
il n' y avait pas d' histoire euh en anglais on faisait euh [**je crois** c' est deux heures euh deux heures en plus]//

Dans la majorité des cas, une analyse prosodique suffit à trancher en faveur d'une hypothèse.

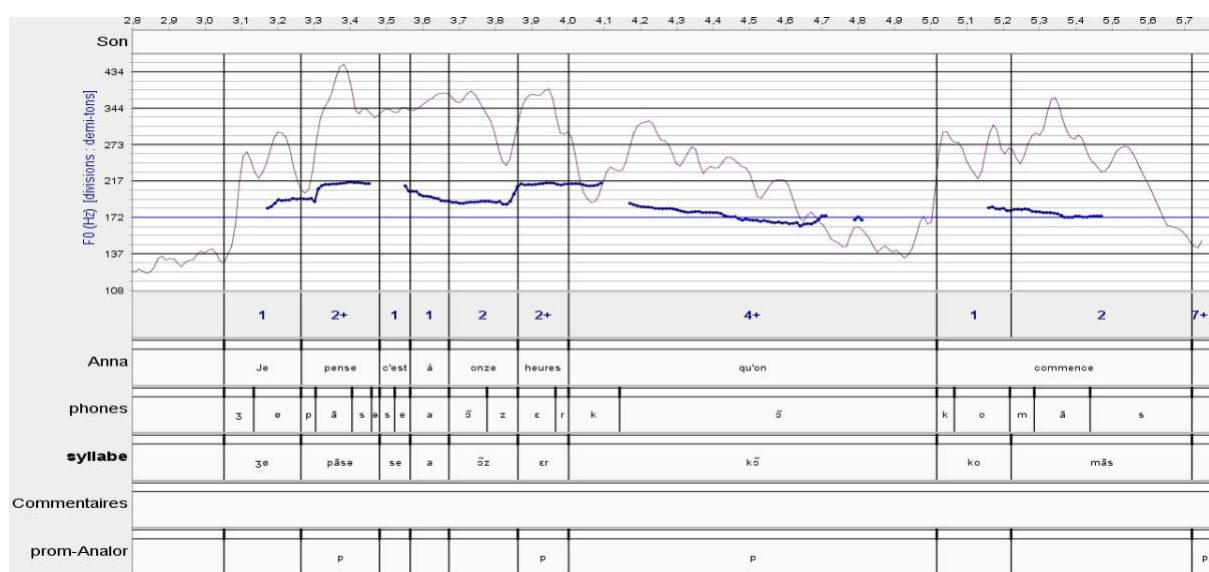
⁸ Le phénomène de greffe en syntaxe a été défini par Deulofeu (1999, 2010 : 178) comme « une violation de contrainte de sous catégorisation d'ordinaire imposées par une catégorie sur les éléments qu'elle régit ».

D'autres énoncés ne présentent aucune ambiguïté, c'est le cas des énoncés (12) ou (13) qui sont incontestablement des cas de [CV₁ Ø CV₂].

(12) La côte escarpée du mont Saint-Pierre qui mène au village euh qui même au village connaît des barrages chaque fois que des opposants de tous les bords mani- **je crois j'ai mal lu la phrase** manifestent leur colère [PFC_Avanzi 2012 p. 276]

(13) **Je pense** c'est à onze heures qu'on commence [MPF_Anna17a]

L'analyse des énoncés (12) et (13) est univoque. L'énoncé (12), extrait d'Avanzi (2012 : 276) et repris par Gachet (2012), illustre un cas de [CV₁ Ø CV₂]. Il est difficile de lier *je crois* à ce qui précède. Dans cet extrait, *je crois* ouvre une parenthèse. La locutrice commente sa lecture en soulignant une confusion « même » et « mène ». Une analyse prosodique, renforce une telle analyse.



Analyse de la séquence *je pense c'est à onze heures qu'on commence*

Suivant la méthodologie d'Avanzi (2012), une analyse prosodique de l'énoncé (13) à l'aide du logiciel Analor⁹ confirme que le segment *je pense c'est à onze heures qu'on commence* est prononcée comme une seule unité prosodique.

Nous avons relevé de nombreux énoncés tels que (12) et (13) dans lesquelles le verbe est bien en tête d'énoncé directement suivi d'une construction verbale. Nous nous intéresserons dans la suite, aux liens entre les deux constructions verbales dans ces structures.

3.3. Lien entre la CV₁ et la CV₂ : quelles hypothèses ?

Nous revenons dans les paragraphes subséquents sur les trois hypothèses majeures qui ont été formulées concernant le lien entre les verbes *croire* et *penser* et les constructions verbales qui les suivent.

La première hypothèse est en faveur d'une perte partielle de la capacité rectrice du verbe, idée défendue par Blanche-Benveniste et Willems ou encore Apothéloz (2003).

⁹ <https://www.lattice.cnrs.fr/ressources/logiciels/analor/>

La seconde hypothèse considère que les verbes *croire* et *penser* ne sont plus recteurs : il n'existe plus de lien de rection ou bien ce lien est inversé. Andersen (2007) note que « le verbe d'opinion joue un rôle syntaxique subordonné par rapport à la proposition à laquelle il se rapporte et où on peut parler de MDP » (Andersen, 2007 : 18). Pour Schneider (2007), il s'agit d'adjoints non régis : ils dépendent de la CV₂ mais ne sont pas régis par elle. Abeillé (2002) les traite comme des ajouts. Selon Gachet (2012), « la forme verbale *je crois* placée immédiatement devant une *P* n'est pas un verbe recteur, fût-ce par l'entremise d'un subordonnant Ø, mais un élément périphérique en position initiale, par analogie avec les adverbes assertifs » (Gachet, 2012 : 230).

Une autre hypothèse est en faveur d'une conservation du statut de recteur fort des verbes¹⁰, faisant de la structure *je crois P* une variante de *je crois que P* : « la structure est alors assimilée à une forme de subordination (subordination sans marqueur, subordination implicite, subordination marquée par Ø). Analyse proposée par Andersen (1993) » (Gachet, 2012 : 227).

Pour résumer, dans les séquences de type [CV₁ Ø CV₂], les constructions verbales sont liés :

(i) à un niveau microsyntaxique, la CV₁ régit la CV₂ (fortement ou faiblement), quelques études proposent d'inverser ces rapports de rections (la CV₂ régit la CV₁ pour Andersen, 2007 entre autres) ou (ii) à un niveau supérieur sans lien rectionnel comparable au niveau macrosyntaxique en nos termes (Gachet, 2012) prenant en compte notamment des critères pragmatiques/ sémantiques, (Schneider, 2007 et Abeillé, 2002). Nous présentons dans la suite de l'article, nos observations quant au lien entre les deux constructions verbales.

4. Les constructions [CV₁ Ø CV₂] : différents liens syntaxiques

Nous avons relevé trois cas de figure : (i) un cas où le verbe peut être qualifié de « verbe faible »¹¹ (ii) un cas où des indices syntaxiques permettent d'assimiler la relation à une rection forte et (iii) un troisième cas où il est difficile de statuer en faveur d'une analyse.

4.1. Le verbe est antéposé

Nous avons relevé des cas où les verbes *croire* et *penser* répondent aux critères dégagés par Blanche-Benveniste et Willems et que nous avons présentés plus haut (§3.1). Les verbes relevés ci-dessous, de (14) à (19), sont à la première personne du présent de l'indicatif.

(14) **Je pense** c'est à onze heures qu'on commence [MPF_Anna17a]

(15) Chokri : Et ce soir il y a une activité ?

Anna : Ce soir je sais pas. (.) **je pense** c'est une journée tranquille comme journée [MPF_Anna17a]

(16) Nacer : C'est où exact ?

Samir : Euh **je crois** c'était le quatorzième hein ou quelque chose comme ça (.) c'était pas au sein du même euh c'était pas au sein du Monoprix [MPF_Nacer8]

¹⁰ Cette capacité rectrice totale est défendue par des linguistes pour les constructions en *qu-P*, c'est le cas notamment de Gachet (2012).

¹¹ Nous ne reviendrons pas sur cette notion qui peut être discutée, l'objet principal de cet article étant d'exposer l'existence de cas de rection forte non marquée morphologiquement avec les verbes *croire* et *penser*.

- (17) Joanne : Ok tu as déjà fait peur à quelqu'un ?
Lupita Mais moi avoir peur ? (rires) Lupita : Ouais c'était en cinquième
% **je crois** c'est ça qui a fait que ma vie elle a changé complètement.
[MPF_Joanne11]

Dans les énoncés (14) à (17), les verbes sont compatibles avec un sens épistémique. Par l'emploi de ces verbes, les locuteurs donnent à voir la recherche d'une valeur (une heure *onze heures* en (14), une qualification *tranquille* en (15), un lieu *le quatorzième* en (16)). Ils indiquent leur incertitude quant au propos exprimé dans la CV₂. En (16) l'incertitude sur le lieu est renforcée par les interjections « hein » ou « euh » et par le syntagme nominal « quelque chose comme ça ».

- (18) Nacer Ben il y a trop de monde **je crois** il faut faire quelque chose hein
% sérieux.
Moustapha Ouais. [MPF, Nacer6]
- (19) Nacer Si si c'- **je pense** il y a un un petit bruit après il ça ça va gêner
pour l'enregistrement [MPF_Nacer8]

Toutefois, dans les énoncés (18) et (19), les verbes *je crois* et *je pense* s'écartent de l'emploi mitigateur (sens faible) ou descriptif (sens fort) mais correspondent davantage à un adoucisseur pragmatique (Kerbrat-Orecchioni, 2005 : 212). Ils permettent au locuteur de minimiser le propos avancé dans le segment verbal qui suit. Cette valeur a été relevée par Gachet (2012 : 199) proche de l'emploi formulaire ou de la fonction phatique.

Ces réalisations, dans lesquelles le verbe est en tête de la structure verbale non précédé de l'élément *qu-*, s'intègre à la famille de constructions des verbes recteurs faibles. Il s'agit d'une cinquième réalisation, que nous pouvons désigner par le terme d'« antéposition »¹². Dans ces énoncés, le verbe est *antéposé* à la CV₂ à laquelle il est rattaché. Suivant le cadre d'analyse de Blanche-Benveniste et Willems, nous pouvons considérer que le verbe régit toujours la construction verbale qui le suit mais faiblement.

Bien qu'il s'agisse d'une structure relevant d'un niveau microsyntactique, des éléments relevant du module de la macrosyntaxe conforte le regroupement de ces réalisations sous le terme de verbes faibles antéposés : la prosodie. D'un point de vue formel, nous observons également des régularités lexicales : la présence de *c'est* ou *il y a* dans le contexte de droite, la présence de nombreux *euh, je sais pas, à mon avis* dans l'environnement de ces occurrences.

4.2. Les hypotaxes asyndétiques

Au cours de notre recherche, nous avons relevé des cas qui ne peuvent pas être traités en termes de rection faible. Le premier critère définitoire des verbes faibles est leur restriction au présent et à la première personne du singulier¹³, or nous avons relevé des énoncés dans lesquels le verbe est à d'autres temps et personnes et dont le sens ou la valeur pragmatique en

¹² Le terme est employé par Andersen (1999) pour désigner des verbes qui ont perdu leur capacité rectrice, ce qui n'est pas le cas ici.

¹³ Nous rappelons que Blanche-Benveniste et Willems tout comme Gachet signalent qu'il s'agit principalement de la première personne du singulier du présent de l'indicatif qui est concernée par l'emploi faible : « Sans aller jusqu'à parler de figement, on peut noter que pour chaque verbe l'emploi de recteur faible est principalement réservé à une seule forme : *je crois, je pense* » (Gachet, 2012 : 227).

discours diffère du « sens faible ». C'est le cas des occurrences (20) et (21) où le verbe est à l'imparfait et (22) avec le verbe *croire* au passé composé.

- (20) Euh ouais ben voilà ouais **je pensais** c'était les filles et tout (.) les Antillaises et tout [MPF_JBD3]
- (21) moi **je croyais** tu disais que dans le film ils ont remporté un truc [MPF_Nawal1]
- (22) genre un un jour pour me dire qu'elle était énervée elle me dit je suis ionde ioumbe ba moi **j'ai cru** elle avait fait une faute de frappe genre tu vois sur le t- un texto [MPF, Auphelie1a]

Nous pourrions envisager que (20), (21) et (22) soient des cas d'assertion mitigée bien qu'aucun cas de ce type n'ai été relevé dans les études antérieures. Toutefois, la valeur d'accompli marquée par le passé composé dans *j'ai cru* en (22) paraît difficilement compatible avec une telle interprétation et va dans le sens d'un verbe décrivant un procès.

Outre une flexion en temps, nous notons également une flexion en personne. Dans les énoncés (23) et (24), le verbe est à la deuxième personne du singulier avec un *tu* spécifique (qui réfère à l'interlocuteur Ibrahimia en (23) et Miloud en (24)).

- (23) Amir : **tu crois** c'est la proche banlieue de Paris ils peuvent pas l'arranger s'ils veulent
Ibrahimia : Ils peuvent l'arranger parce que si ils ont voulu laisser ça comme <ça pour qu'ils restent> dedans <rester dedans>. [MPF_Wajih5]
- (24) Emir : Tu connais rien tu as vu qu'une vidéo et **tu crois** tu connais tout à sa vie ça y est [MPF_Wajih2]

Outre la flexion morphologique du verbe¹⁴, qui distingue cet usage de celui des recteurs faibles décrits comme bloqués sur la première personne, nous observons que la construction verbale est coordonnée à une suite de verbes également à la deuxième personne, *tu connais rien* ; *tu as vu qu'une vidéo* et *tu crois*, et semble bien être au même niveau d'assertion de discours¹⁵.

Dans les énoncés (25) à (29) les verbes sont à la troisième personne (singulier et pluriel). Pour ces structures, toute interprétation en termes d'assertion mitigée est impossible puisque le verbe à la troisième ne peut être en aucun cas la « manifestation de la subjectivité [qui] ne prend son relief qu'à la première personne » (Benveniste, 1966 : 264). Il s'agit bien ici de verbes descriptifs.

- (25) **elle a cru** j'allais pleurer parce que je faisais pas le truc de délégué [MPF_Wajih4]
- (26) Et quand mon grand frère il est pas là **il croit** trop c'est le grand frère en fait [MPF, Joanne11]

14 Nous n'avons pas relevé d'occurrence du verbe aux personnes *nous* et *vous* ni même à d'autres temps que ceux signalés ci-dessous. Il s'agirait peut-être de contraintes qui restent à explorer.

15 Il ne s'agit pas ici d'une assertion mitigée. Le verbe joue bien un rôle au plan propositionnel et ne peut être relayé qu'à un plan métadiscursif.

- (27) Caroline : Des vrais schlags.
Melissa : Et **il croit** genre c'est un bête de mec (rire) [MPF_Zakia3]
- (28) Ici **ils croient** ils sont à l'armée [MPF_Wajih4]
- (29) ma famille **ils croient** ils vont me marier avec ma cousine ils sont fous eux ou quoi [MPF_Wajih3_2]

Les verbes dans ces énoncés endossent une valeur pragmatique supplémentaire : le propos exprimé dans la CV₂ est attribué à une tierce personne et rejeté par le locuteur. L'intonation et des expressions telles que « en fait », « genre » ou « ils sont fous ou quoi », nous permettent de les interpréter dans ce sens. Le locuteur semble rejeter une idée qui lui semble aberrante.

Pour l'énoncé (26), nous notons également la présence de l'adverbe *trop*. Nous avons également signalé la présence d'adverbes dans Ferreira (2019) que nous illustrions avec les énoncés repris sous (30) et (31).

(30) **tu crois** vraiment il va mettre cinq cents balles dans un truc [MPF, Auphelie1e]

(31) **je pense** au contraire maintenant on commence à changer et à se dire bon voilà quoi [MPF, Sandrine2]

La présence de *trop*, *vraiment* et *au contraire* oblige à une analyse du verbe de plein statut¹⁶ et donc à une interprétation sémantique en termes de procès cognitif.

Nous notons également avec ces énoncés qu'une pronominalisation en *le* ou *ça* sans changement de sens est possible. Ce qu'illustrent par exemple la manipulation de l'énoncé (24) : « Tu connais rien tu as vu qu'une vidéo et **tu crois ça** ».

Les verbes dans ces occurrences n'entrent pas dans une famille de construction. Ils peuvent être comparés aux énoncés en *qu-P* :

(32) On avait dit ouais khari on était là on est descendus oh c'est le métro du futur comment faut descendre et pour euh y aller **on croyait que** c'était la ligne quatorze et tout le truc (.) [MPF_Salima2]

En revanche, nous n'avons pas relevé de cas où le verbe est en incise (en milieu ou fin d'énoncé) du type « [?]Euh ouais ben voilà ouais c'était les filles et tout, je pensais ».

Nous avons bien pour ce type de structure un lien rectionnel sans marquage morphosyntaxique, là où le *qu-* serait attendu. Il ne s'agit pas d'un phénomène isolé et propre à ces verbes puisque d'autres recherches, tels que Benzitoun (2010 : 158) ou Willems et Meulleman (2010) ont montré respectivement sur des énoncés du type « elle est arrivée en France elle avait trois ans » et « il y a des gens ils viennent acheter des aspirines pour faire de l'eau gazeuse » qu'il existe des constructions verbales régies non introduite par un terme dit « conjonctif » en français.

¹⁶ Gachet (2012 : 202) relève également un énoncé dans lequel le verbe est renforcé par *vraiment* et qui l'incite à penser qu'il est employé dans son sens fort. Néanmoins, tous les matériaux linguistiques exprimant la certitude ne vont pas nécessairement augmenter la force assertive de l'énoncé, ce que signale également Gachet (2012 : 196) au sujet des adverbes *sûrement*, *certainement* ou *probablement*.

4.4. Des cas ambigus

Toutefois, il reste des structures pour lesquelles il est difficile de statuer quant à la nature du lien entre les verbes *croire* et *penser* et la construction verbale qui suit. C'est le cas notamment pour des constructions dont le verbe à la première personne du singulier du présent de l'indicatif. Nous avons relevé dans une précédente étude l'énoncé repris sous (33).

- (33) Roberto : Mais c'est peut-être parce que tu connais du monde ?
Gaël : Ouais **je pense** c'est ça et que j'ai l'habitude de me balader ici
[MPF, Roberto1b]

Dans cet énoncé, le verbe *penser* est à la première personne du singulier du présent et il est suivi d'un présentatif, cooccurrence fréquente avec les verbes recteurs faibles (Avanzi, 2012 : 282), ce qui pourrait constituer des arguments pour une analyse du verbe en emploi recteur faible. Néanmoins la coordination avec une complétive marquée morphologiquement par *que* et l'équivalence avec les proformes, « je le pense », « je pense ça » où *le* et *ça* reprennent « c'est parce que je connais du monde », constitueraient des indices d'une construction en hypotaxe asyndétique.

Cette difficulté d'interprétation n'est pas réservée aux constructions du type [CV₁ Ø CV₂] puisque Willems et Blanche-Benveniste (2010) signalaient également les difficultés à trancher entre une interprétation en recteur faible ou fort pour les constructions suivies d'une *qu-P*.

Nous avons relevé, dans d'autres corpus, des cas où le verbe semble « flottant ». Nous observons dans ces séquences que le verbe est séparé de la CV₂ par une parenthèse comme en (34) ou un hanging topic en (35) et que la prosodie diffère des énoncés analysés précédemment. Nous notons une pause plus longue et un allongement de la dernière syllabe de la parenthèse « *après* » en (34) ou du hanging topic « *enfants* » en (35).

- (34) et **je crois** trois ou quatre mois après euh ben | je découvrais que Cage euh je découvrais plutôt les oeuvres pour pianos préparés de John Cage donc euh [CEFC_TCOF> explorationssonores]

- (35) euh oui **je crois** les enfants dès le primaire faut commencer à économiser pour euh l' université parce que à quinze mille d~ dollars par an [CEFC_TUFS > 15_LW_MG_100224]

Pour l'énoncé (36), il est difficile de dire si *je crois* porte en réalité sur ce qui précède ou ce qui suit.

- (36) Tu as vu après les euh tu as vu la génération de nous on a on a x guerre **je crois** on a plus le seum que nos grands-parents et tout ça là comment ils les ont mal traités et tout ça. [MPF_Nacer2]

Avanzi (2012 : 286) relève également des séquences qui présentent des ambiguïtés de rattachement. Nous pourrions analyser ces ambiguïtés comme d'autres cas de verbe « flottant », interposé entre deux constructions et pouvant porter sur les deux. Dans les énoncés (34) à (36), le lien entre la CV₁ et la CV₂ relèverait davantage de la macrosyntaxe.

Enfin nous terminons notre analyse en signalant qu'il semble exister une différence d'emploi entre *croire* et *penser*. Nous relevons davantage d'occurrences à d'autres temps et personnes pour le verbe *croire* que *penser*.

5. Bilan (provisoire) et éléments de conclusion

Nous avons dépassé le cadre des verbes « faibles », pour nous intéresser à l'ensemble des productions des verbes *croire* et *penser* rattachée (à un niveau micro- ou macro) à segment verbal sans éléments *qu-*. Après l'observation d'occurrences extraites de données de français parlée, nous avons mis en avant une diversité de fonctionnement syntaxique de ces verbes qui s'accompagnait également d'une diversité sémantique et pragmatique.

Nous proposons de recenser les fonctionnements syntaxiques des verbes analysés dans le tableau ci-dessous.

Nom de la construction	Lien	Exemples d'énoncés
Antéposition	Rection faible (microsyntaxe)	Je pense c'est à onze heures qu'on commence
Hypotaxe asyndétique	Rection forte (microsyntaxe)	elle a cru j'allais pleurer parce que je faisais pas le truc de délégué
« Flottant »	Macrosyntaxique (à explorer)	euh oui je crois les enfants dès le primaire faut commencer à économiser pour euh l'université parce que à quinze mille d~ dollars par an

Ce recensement est incomplet, d'autres occurrences sont encore en cours d'analyse. De plus, nous avons conscience que la prise en compte de nouveaux énoncés pourrait remettre en cause certaines propriétés et classifications proposées dans ce travail, tout comme nous avons pu mettre en lumière certaines réalisations du verbe *croire* et *penser* avec les données issues du corpus MPF.

Pour finir, il est clair que la diversité d'emplois de ces verbes dans des séquences de type [CV₁ Ø CV₂] a été mise à jour grâce à l'accès à de nombreuses données de français parlée, notamment des données « écologiques ». Nous souhaitons, avec ce travail, réaffirmer la place essentielle des corpus de français parlé, et en particulier de données qui ont été recueillies dans des contextes de grande proximité, dans la description de la syntaxe du français.

6. Bibliographie

- Abeillé, A. (2002). *Une grammaire électronique du français*. Paris: CNRS éditions.
- Andersen, H-L (1993) Les complétives non introduites en français parlé. *Subordinations, Travaux de linguistiques du Cerlico*(6), 5-14. Rennes.
- Andersen, H. L. (1999). Propositions adverbiales non introduites en français parlé. *Revue Romane*(34), 164-179.
- Andersen, H. L. (2007) Marqueurs discursifs propositionnels, *Langue française*, (154), 13-28.
- Apothéloz, D. (2003). La rection dite 'faible' : grammaticalisation ou différentiel de grammaticité ? *Verbum*(25), 241-262.
- Avanzi, M. (2012). *L'interface prosodie/ syntaxe en français. Dislocations, incises et asyndètes*. Bruxelles: Peter Lang.
- Benveniste, E. (1966). De la subjectivité dans le langage. Dans E. Benveniste, *Problèmes de linguistique générale* (Vol. 1, pp. 258-266). Paris: Gallimard.

- Benzitoun, C. (2010). Comment tirer profit de la parataxe ? *La parataxe. Entre dépendance et intégration. Tome 1*, 153-174. Bern: Peter Lang.
- Bilger, M., Debaisieux, J-M., Deulofeu, J., & Sabio, F. (2013). Chapitre 2 : Cadre descriptif. Dans Debaisieux, J-M. (dir.) (2013). *Analyses linguistiques sur corpus. Subordination et insubordination en français*, 61-95. Paris : Hermes –Lavoisier.
- Blanche-Benveniste, C. (1989). Constructions verbales 'en incise' et rection faible des verbes. *Recherches sur le français parlé*, 53-73.
- Blanche-Benveniste, C. (2010) : *Approche du français parlé*, Paris : Ophrys.
- Blanche-Benveniste, C., Deulofeu, J., Stefanini, J. & Van Den Eynde, K. (1984). *Pronom et syntaxe. L'approche pronominale et son application au français*, Paris : SELAF.
- Blanche-Benveniste, C., Bilger, M., Rouget, C., Van den Eynde, K. & Mertens, P. (1990). *Le français parlé : Études grammaticales*. Paris : éd. du C.N.R.S.
- Blanche-Benveniste, C., & Willems, D. (2007). Un nouveau regard sur les verbes faibles. *Bulletin de la Société Linguistique de Paris*(102), 217-254.
- Blanche-Benveniste, C., & Willems, D. (2016). Verbes faibles. *Encyclopédie grammaticale du français*. En ligne : http://encyclogram.fr/notx/010/010_Notice.php [consulté le 27/08/20]
- Bolinger, D. (1968). Postposed main phrases: an English rule for the romance subjunctive. *Canadian journal of linguistics*(14), 3-30.
- Debaisieux, J-M & Benzitoun, C. (2020) : Orféo : un corpus et une plateforme pour l'étude du français contemporain. *Langages*(219), 9-24.
- Deulofeu, H.-J. (1999). *Recherches sur la syntaxe de la conjonction 'que' dans l'ancien français*. Université de Paris III: Thèse d'Etat.
- Deulofeu, H.-J. (2010). La greffe d'un énoncé sur une construction : une combinaison originale de parataxe et de rection. Dans M.-J. Béguelin, M. Avanzi, & G. Corminboeuf, *La parataxe. Entre dépendance et intégration. Tome 1*, 175-208. Bern: Peter Lang.
- Ferreira, A. (2019). Influence du contexte communicationnel sur l'emploi des constructions syntaxiques de type [V.ØP.] avec les verbes croire et penser. Dans *Actes des Rencontres des Jeunes Chercheurs en Sciences du Langage 2019*. Paris : France.
- Gachet, F. (2012). *Incises de discours rapporté et autres verbes parenthétiques : une étude grammaticale*. Fribourg: Université de Fribourg.
- Gadet, F. & Guerin, E. (2016). Construire un corpus pour des façons de parler non standard : 'Multicultural Paris French'. *Corpus*, 15, pp. 285-307.
- Gadet, F. (dir.) (2017). *Les parlers jeunes dans l'Île-de-France multiculturelle*, Paris, Ophrys.
- Haßler, G. (2014). Étude comparée de l'usage parenthétique des verbes épistémiques dans trois langues romanes. *Discours*(14), en ligne <https://journals.openedition.org/discours/8888>.

- Kerbrat-Orecchioni, C. (2005). *Le discours en interaction*. Paris: Armand Colin.
- Martineau, F. (1988). Variable Deletion of que in the Spoken French of Ottawa- Hull. Dans J. P. Montreuil and D. Birdsong (éds), *Advances in Romance Linguistics*, 275-287. Foris, Dordrecht.
- Martineau, F. (1993). Rection forte et rection faible des verbes : l'ellipse de Que en français du Québec et de l'Ontario. *Francophonie d'Amérique, Le français, langue maternelle, en milieux minoritaires*(3), 79-90.
- Sankoff, D., & Cedergren, H. (1971). Some results of a sociolinguistic study of montreal french. *Linguistic diversity in Canadian society*, 61-87.
- Schneider, S. (2007). *Reduced parenthetical clauses. A corpus Study of Spoken French, Italian and Spanish*. Amsterdam: Benjamins.
- Thompson, S. & Mulac, A (1991a). A quantitative perspective on the grammaticalization of epistemic parentheticals in English. Dans Traugott, E. & Heine, B. (Éds), *Approaches to Grammaticalization* (vol.2), 313-329. Amsterdam/Philadelphie : Benjamins.
- Thompson, S. & Mulac, A (1991b). The discourse conditions for the use of the complementizer *that* in conversational English. *Journal of Pragmatics*(15), 237-251. North-Holland.
- Urmson, J.O. (1952). Parenthetical verbs. *Mind*(61), 480-496.
- Wiesmath, R. (2002). Présence et absence du relatif et conjonctif que dans le français acadien : tendances contradictoires ? Dans C. Pusch, W. Raible, C. Pusch, & W. Raible (Éds.), *Romanistische Korpuslinguistik-Korpora und gesprochene Sprache/ Romance Corpus Linguistics - Corpora and Spoken Language* (pp. 393-408). Tübingen: Gunter Narr Verlag.
- Willems, D. & Meulleman, M. (2010). « Il y a des gens ils viennent acheter des aspirines pour faire de l'eau gazeuse ». Sur les raisons d'être des structures parataxiques en « il y a ». Dans Béguelin; M-J. Avanzi; M., Corminboeuf, G (dir.) *La parataxe. Entre dépendance et intégration. Tome 2*,167-184.
- Willems, D., & Blanche-Benveniste, C. (2010). Verbes 'faibles' et verbes à valeur épistémique en français parlé : il me semble, il paraît, j'ai l'impression, on dirait, je dirais. Dans M. Iliescu, H. M. Siller-Runggaldier, & P. Danler (Éd.), *Actes du XXVe Congrès International de Linguistique et de Philologie Romanes. IV*, pp. 565-579. Berlin: De Gruyter.
- Willems, D., & Blanche-Benveniste, C. (2014). A constructional corpus-based approach of 'weak' verbs in French. Dans H. Boas, & F. González-García, *Romance Perspectives on Construction Grammar* (pp. 113-138). Amsterdam: J. Benjamins.

Une reconsidération du discours rapporté en langue parlée avec *être là*, *faire (à quelqu'un)* et *se dire*

Sophia Fiedler, Université de Neuchâtel/Universität Hamburg

1. Introduction

Le discours rapporté suscite depuis longtemps l'intérêt de nombreux.e.s chercheur.e.s, et ce dans de nombreux domaines de la linguistique (Mosegaard Hansen 2000; Rabatel 2004; Streeck 2002; Tannen 2007). Dans ces recherches, le rapport entre verbe "introduceur" et discours rapporté constitue souvent un point de discussion central, sous l'angle de plusieurs de ses particularités linguistiques : la conception du verbe "introduceur" comme verbe recteur ou incise, et, découlant de cela, le statut du discours rapporté comme complément du verbe ou non (Blanche-Benveniste 1988; de Cornulier 1978), mais aussi la conception du discours rapporté en soi, qui peut, suivant de Cornulier (1978), consister en de la langue parlée ou en un « mime » (p. 80) (c'est-à-dire une imitation prosodique ou gestuelle).

Le but de ma communication consiste à appliquer principalement les réflexions syntaxiques de Blanche-Benveniste (1988) et de de Cornulier (1978) aux verbes "introduceurs" (appelé ci-dessous introduceurs) que je trouve dans mon corpus de français (romand) parlé. A côté des verbes "traditionnels" comme *dire* ou *répondre*, ce sont là avant tout trois verbes : *se dire* (n=93), *faire (à quelqu'un)* (n=48) et *être là* (n=85), tous trois à la 1^{ère} pers. du singulier du présent et du passé, ainsi que, dans le cas de *faire* et *être là*, également à la 3^{ème} pers. du singulier du présent et du passé. Comme les trois introduceurs sont, à l'exception de neuf cas avec *se dire*, suivis par du discours rapporté direct, je me concentrerai sur ce format du discours rapporté (pour une vue d'ensemble de la distinction entre discours rapporté direct et indirect voir Mosegaard Hansen (2000)).

Les analyses principalement qualitatives se basent sur 12 heures d'enregistrements du corpus *Pauscaf* du Centre de Linguistique Appliquée de Neuchâtel. Ces enregistrements vidéo consistent en 34 interactions entre étudiant.e.s pendant leurs pauses café. Les enregistrements ont été faits en 2013 et 2019 et ont été transcrits selon les conventions de transcription de Jefferson (2004) (voir annexe).

Pour pouvoir montrer la distribution fonctionnelle de *faire*, *se dire* et *être là*, je traiterai non seulement des aspects syntaxiques mais également des questions séquentielles. En recourant à l'analyse conversationnelle, j'illustrerai le fait que l'utilisation de ces verbes dépend d'un contexte plus large, et notamment de ce qui précède dans la conversation. *Faire*, *se dire* et *être là* ne sont pas utilisées d'une manière aléatoire mais permettent aux locuteur.ice.s de réaliser des actions spécifiques dans la conversation en cours. Dans ce qui suit, j'essayerai donc de présenter les caractéristiques formelles et séquentielles de *se dire*, *faire (à quelqu'un)* et *être là* pour montrer comment l'analyse de la langue parlée peut compléter notre grammaire. Je suivrai ainsi l'idée de Claire Blanche-Benveniste pour qui « [l]a prise en compte de la langue parlée en Europe est en train de changer la conception que nous nous faisons de la grammaire de nos langues » (2007: 139).

Pour commencer, je ferai un bref panorama de la recherche qui a été effectuée au sujet de *se dire*, *faire* et *être là*. Je présenterai ensuite quelques réflexions théoriques, avant de passer aux aspects formels (avant tout syntaxiques) des trois formes. Ceci sera suivi d'une analyse qualitative de quatre extraits illustratifs. Je terminerai cet article par une brève conclusion.

2. Littérature

Deux des trois formes en question n'incluent pas de verbe portant une sémantique de *dire* ou *penser*. D'un point de vue sémantique, *faire* renvoie à une action tandis qu'*être là* dénomme la présence de quelqu'un à un certain moment ou lieu. Comme il a été tout récemment montré, *faire* et *être là* ne portent la fonction d'introducteur de discours rapporté que depuis peu (Cheshire & Secova 2018; Dostie 2020). D'une manière générale, ce développement d'une fonction d'introducteur de discours rapporté va de pair avec le constat de Tagliamonte et D'Arcy (2007: 211) selon lequel l'utilisation du discours intérieur (qui est rendu public en étant rapporté) a augmenté dans les dernières décennies et que de nouveaux introducteurs, comme en anglais *be like* 'être comme', ont rempli une lacune langagière pour marquer ce discours intérieur comme tel.

Concernant *faire* (à *quelqu'un*), Abouda et Redulić (2017) montrent dans leur étude contrastive de *dire* et *faire* que les deux introducteurs ne sont pas toujours interchangeables. Ceci est dû aux caractéristiques spécifiques de *faire* : quand le verbe est utilisé comme verbe introducteur, il est dans 85% des cas au présent et, avec seulement une exception, suivi par du discours rapporté direct (*ibid.*: 15) donc sans la conjonction *que*. Gardant son statut de verbe transitif, *faire* est – à un niveau strictement formel – suivi par son complément : du discours rapporté direct ayant « une prévalue pragmatique qui consiste à présentifier, en le mimant et en jouant, l'acte de parole représenté » (*ibid.*: 3). Dans leur étude, les auteures excluent un remplacement de *faire* par *dire* dans certains cas : quand *faire* introduit des sons, des interjections ou des présentifications, c'est-à-dire des imitations prosodiques ou gestuelles de la *façon* dont quelque chose a été dit. Abouda et Redulić (2017) montrent donc que certaines spécificités formelles du matériel rapporté représentent une condition nécessaire pour l'utilisation de l'introducteur respectif *faire* ou *dire*. Cette comparaison peut être étendue à d'autres introducteurs dont les citations remplissent des critères formels semblables que celles suivant *faire* : c'est notamment le cas d'*être là*.

Secova (2015), qui classe *j'étais là* comme une forme routinisée et non comme *verbum dicendi* traditionnel, décrit cet introducteur comme faisant partie des « termes déictiques variés, dont le but est d'attirer l'attention de l'énonciateur vers la source des propos cités, c'est-à-dire vers leur auteur » (*ibid.*: 142). Cette observation est confirmée par Dostie (2020), qui voit un lien entre le déictique *là* et la récurrence du comportement corporel dans les citations introduites par *être là*. Selon elle, *là* réfère d'une manière cataphorique à la citation (corporelle) qui suit (Dostie 2020: 86) – une observation qui rejoint ce que Blanche-Benveniste (2010) propose dans sa « grammaire de la période » (Blanche-Benveniste 2010: 188) : *Être là* est sous-spécifié tel quel et projette ainsi une spécification du déictique (local ou temporel) *là*. Secova (2015) voit dans le caractère mimétique des citations introduites par *être là* l'origine de leur fonction : elles permettent aux locuteur.rice.s d'exprimer leur attitude par rapport à ce qui a été dit. En m'appuyant sur les études susmentionnées, j'essayerai dans ce qui suit de décrire l'émergence de ces trois formes de leur contexte de parution, pour ainsi comprendre comment elles sont distribuées alors qu'elles introduisent toutes trois du discours rapporté. En effectuant des analyses détaillées de la séquentialité des récits contenant *faire*, *se dire* et *être là*, je montrerai que ces trois formes ne sont pas utilisées d'une manière aléatoire, mais qu'il y a une distribution fonctionnelle entre les trois formes qui dépend essentiellement de l'action que les locuteur.rice.s effectuent à l'aide du discours rapporté.

3. Réflexions théoriques

Avant de passer à la description des caractéristiques syntaxiques et sémantiques de chaque construction, j'aimerais proposer quelques considérations théoriques en partant principalement des réflexions de Blanche-Benveniste (1981; 1988) et de de Cornulier (1978). Ceci sera nécessaire pour comprendre les nouveaux défis découlant d'une intégration de *faire*, *être là* et *se dire* comme introducteurs de discours *et* de mimes dans une grammaire de la langue parlée.

Dans mes données, *être là* et *faire* introduisent toujours du discours rapporté *direct*, donc une unité qui n'est pas marquée comme étant syntaxiquement liée au verbe introducteur. Prenons d'abord trois exemples de mes données pour illustrer les relations syntaxiques entre introducteur et citation :

- (1) LAU: pis j'ai attendu pis **je me suis dit** bon ben: je verrai,
(Pauscaf 10, 19min30)
- (2) JOA: pis **elle me fait** <<((en imitant la locutrice originale)) ouais
je fais trop de sport avec la perrine,> (Pauscaf 05, 04min26)
- (3) ELI: et je m'suis relevée j'ai vu tout noir **j'étais là**, ((imite
sa position immobilisée pour ne pas s'évanouir pendant 1.1s))
(Pauscaf 20, 32min10)

Le premier exemple pourrait être décrit comme une clause complexe constituée du verbe recteur *se dire* et de son complément *bon ben je verrai* qui occupe la place de l'objet direct. Cette phrase pourrait être reformulée en ajoutant la conjonction *que*, mais cela impliquerait que les particules d'amorces (Guérin & Moreno 2015) soient enlevées : *Je me suis dit que je verrai*. Un double marquage est également possible. S'il n'y avait pas l'imitation corporelle, une telle reformulation avec *que* serait également possible dans l'exemple (2) – à condition que le pronom personnel soit changé dans le complément. Néanmoins, le fait d'avoir ici une citation mimique bloque cette reformulation en discours rapporté indirect car l'imitation doit être co-référentielle avec la locutrice originale de l'énoncé, qui est, dans l'exemple (2), imitée par Joana. La co-référentialité entre comportement corporel et information grammaticale ne serait plus assurée avec le changement du pronom. Dans l'exemple (3), l'ajout de *que* n'est pas possible car il ne s'agit pas d'une relation pronominale : *être là* est syntaxiquement déjà complet. Selon Blanche-Benveniste (1988), le discours rapporté direct – ou le mime, dans le cas présent –, « en l'absence d'une relation avec un pronom, [...] ne représente aucune catégorie grammaticale ; il est, du point de vue grammatical, "informe" » (*ibid.*: 56). Ceci peut être vrai pour *être là* mais doit être mis en question pour *faire* et *se dire*. Il convient dès lors de définir quel statut syntaxique ou pragmatique pourrait être attribué à de telles citations que Blanche-Benveniste décrit comme « style direct "pur" » (*ibid.*: 57).

Selon Blanche-Benveniste, si l'introducteur est suivi par un mime, il n'y a pas de relation pronominale entre les deux parties de la clause (verbe introducteur et citation). Pour *faire* et *se dire*, un mime seul ne constituerait donc pas l'objet direct du verbe, ce qui laisserait néanmoins le verbe introducteur incomplet (voir aussi Mosegaard Hansen (2000)). Ce débat quant à savoir si le mimo-gestuel fait partie de la grammaire d'une langue a été suscité récemment par Keevallik (2018), qui demande explicitement une inclusion de mouvements corporels dans la grammaire, entre autre dans le cas du discours rapporté direct où les imitations et mises en scène font souvent partie intégrante du format grammatical (Keevallik 2018: 14).

Si le discours rapporté qui suit n'est pas grammaticalement régi par la construction introductrice, le discours ou le mime pourrait être autrement décrit comme y étant « associé »

(Blanche-Benveniste 1981). De Cornulier conceptualise ce rapport syntaxique entre introducteur et mime en expliquant que le mime est « greffé sur l'acte linguistique » (1978: 84). Ces deux conceptions, dans leur façon d'être formulées, renvoient paradoxalement tout de même à l'idée que le verbe introducteur reste hiérarchiquement "au-dessus" de la citation, alors que cette idée est en même temps réfutée par Blanche-Benveniste (1981) et De Cornulier (1978) quand la citation est un mime. Toujours est-il que, partant de mes données, il semblerait que, au moins au niveau syntaxique, la distinction entre discours direct et mime devrait être abandonnée car dans la langue parlée les deux fonctionnent d'une manière très similaire. Mais, selon l'introducteur, différents formats de citations (mime ou discours) sont introduits qui accomplissent des actions distinctes dans la conversation en cours.

Après ces réflexions théoriques, il semble nécessaire de décrire brièvement les trois verbes introducteurs en question avant de passer à leur analyse qualitative.

4. Aspects formels

Être là ne peut être formellement décrit comme un verbe transitif, à la différence de *se dire* ou *faire*. L'introducteur est composé du verbe *être*, le plus souvent à la première personne du singulier de l'imparfait, et d'un adverbe qui fonctionne comme un déictique renvoyant à un lieu ou moment spécifique. *Être là* introduit toujours un discours corporellement ou prosodiquement animé (Dostie 2020; Secova 2015). La variété formelle des citations s'étend de sons, du langage accompagné de mouvements corporels jusqu'à un mouvement seulement (ce que Streeck (2002) appelle les *body quotes*). La configuration syntaxique de [*être là* + citation] se distingue elle aussi de celle des verbes transitifs *se dire* et *faire*. Syntaxiquement complet comme tel, *être là* ne projette, au niveau grammatical, aucun complément obligatoire, à la différence des verbes transitifs *faire* ou *se dire*. Le déictique *là* demande cependant à être spécifié ce qui établit une force projective interactionnelle : la construction projette une continuation de l'énoncé. Cette idée contredit de Cornulier qui, dans son article de 1978, explique que juste parce qu'« un mime peut être explicitement annoncé et désigné, [il] n'entre pas par cette seule raison dans la syntaxe grammaticale du discours » (*ibid.*: 83). Selon lui, un déictique comme *là* dans *être là* projetterait certes plus de parole, mais cette force projective ne serait pas suffisante pour l'intégrer dans une description syntaxique du discours rapporté direct. Un indice qui montre que ce constant mérite d'être (re)mis en question est le comportement corporel des interlocuteurs : un regard systématique de l'interlocuteur.rice au/à la locuteur.rice après *être là* – quand le regard n'est pas déjà établi – montre que la force projective semble être assez efficace pour attirer son attention.

Se dire et *faire* sont des verbes transitifs demandant un complément. Malgré cette configuration grammaticale claire, leur lien syntaxique avec la citation introduite n'est pas simple. Quand *faire*, sans l'objet indirect *me*, introduit du discours rapporté, le verbe ne *peut* pas être suivi par la conjonction parce qu'elle ferait de ce qui suit une action conséquentielle. Pour les formes avec objet indirect, un marquage est possible mais n'existe pas dans mes données ; la question de savoir s'il s'agit d'un non-marquage du lien syntaxique entre l'introducteur et la citation ou plutôt d'une juxtaposition constitue un sujet de discussion à part. Selon Abouda et Rendulić (2017), *faire* introduit, comme *être là*, fréquemment du discours rapporté qui est corporellement animé – une observation qui est également récurrente dans mon corpus ; parfois le verbe introduit, comme *être là*, seulement un geste ou son.

Se dire est la seule des trois formes à laquelle la citation peut être rattachée à l'aide de la conjonction *que*. En comparaison avec *être là* et *faire*, *se dire* est la seule forme qui introduit

des citations pratiquement non-animées au niveau mimo-gestuel. Dans mon corpus, la plupart des occurrences de *se dire* sont suivies de clauses complètes – et non d’une vocalisation ou d’un geste uniquement. Le cas de *se dire* est donc le moins complexe au niveau syntaxique car il est le seul introducteur des trois dont l’objet n’est pas composé uniquement d’unités non-linguistiques. Il servira *ex-negativo* à illustrer la pertinence de l’inclusion des mimes dans une description du discours rapporté direct, car ce sont les caractéristiques mimo-gestuelles et prosodiques qui permettent de différencier les actions accomplies avec *faire*, *être là* et *se dire*. Après cette vue d’ensemble d’un point de vue formel et une juxtaposition des conceptualisations de Blanche-Benveniste et de Cornulier, je vais montrer, à l’aide de quatre exemples illustratifs, qu’il y a une distribution fonctionnelle des trois unités lexicales qui se manifeste à la fois à travers leur configuration syntaxique, le matériel cité et le contexte séquentiel dans lequel le discours rapporté est utilisé. Tandis que *être là* permet aux locuteur.rice.s d’évaluer une situation ou un discours rapporté introduit auparavant, *faire* est utilisé pour rapporter des dialogues, avec une fréquence plus élevée pour les formes à la troisième personne du singulier – les deux formes apparaissant dans des séquences émotionnellement chargées ; enfin, *se dire* est un outil servant à partager une réflexion non-émotionnelle pendant des activités de description.

5. Analyses

Dans la partie présente, je vais présenter quatre extraits de données afin d’illustrer chaque introducteur dans son contexte séquentiel. Je montrerai que malgré certains parallèles formels, surtout entre *être là* et *faire*, on peut constater qu’il y a des différences interactionnelles qui caractérisent les introducteurs au-delà de leurs attributs syntaxiques.

Être là introduit, dans mes données, toujours une réaction évaluative à un événement ou une action préalablement décrite. Utilisé avant tout à la première personne du singulier (n=60), il constitue une évaluation (rapportée) du locuteur ou de la locutrice même. Cette évaluation à l’aide de *j’étais là* apparaît dans 44 cas à la fin d’une narration, permettant ainsi aux locuteur.rice.s de clore leur tour avec leur propre attitude – un phénomène récurrent, et qui a été décrit dans plusieurs études dans l’analyse conversationnelle (Couper-Kuhlen & Thompson; Goodwin 1986; Goodwin & Goodwin 1992).

Considérons d’abord un exemple avec cette forme. Trois ami.e.s parlent d’une newsletter qu’ils reçoivent par mail. Juste avant cet extrait, Pascal dit qu’il ne lit jamais cette newsletter. Liliane réagit :

Ex. 1: europa-park (Pauscaf_02_08, 02min47)

01 LIL: mais hier j’ai lu,

02 >parce que< l le titre de du mail c’était,

03 gagnez deux places pour europa-park,

04 et du coup [j’tais ’là° (.) #OH, #]

fig

#fig1#



#1

05 PAS: [ÇA c'est un beau titre de mail]

pour que tu lises ouais.

06 LIL: ouais,

J'étais là est très réduit morpho-phonologiquement et il est suivi d'une micro-pause pendant laquelle Liliane commence déjà à changer sa posture: elle bouge le torse vers l'avant et rassemble ses mains. Pendant le *oh*, elle met en plus sa tête en arrière et lève les yeux, comme montré dans la figure 1. La citation, qui apparaît à la fin de son tour et ainsi à la fin de son récit, lui permet, elle, de prendre position par rapport au contenu du mail et de l'évaluer, dans ce cas, positivement.

L'extrait 2 est similaire mais avec, cette fois-ci, plus de matériel lexical qui rend l'évaluation de ce qui précède encore plus évidente. Mathilde parle de sa mère qui est, selon elle, trop sceptique à l'égard de certains changements dans la société. L'extrait se réfère à une application que Mathilde lui a montrée.

Ex. 2: de la pub (Pauscaf_02_17, 08min20)

01 MAT: et pis euh direct elle est elle crie au complot t'sais.

02 mais voi↑là: c'est pour faire d'la ↑pub et ↑tout,

03 naNA+nanan,

fra +regarde en bas-->l.05

04 **ch'tais °là°** ↑o↓:h °putain°,

05 en:+core,

fra -->+regarde MAT-->l.08

Le juron à la ligne 04 évalue la critique de la mère qui est rapporté juste avant. Pendant le juron, Mathilde change de qualité de voix, ce qui lui permet de rendre son évaluation négative directement accessible via sa prosodie. Notez également que Francis, qui regarde en bas, commence à regarder Mathilde juste après ce juron, ce qui indique son attention portée sur le comportement mimo-gestuel de Mathilde à ce moment-là.

Ces deux premiers extraits ont montré que *j'étais là* est utilisé à la fin d'un récit pour évaluer des circonstances ou actions qui ont été décrites préalablement. Dans les deux cas les locutrices se servent non seulement du langage mais aussi des ressources corporelles et prosodiques pour exprimer leur attitude. Syntaxiquement, les exemples montrent que c'est effectivement le mime qui sert de spécification du déictique *là*. Il serait donc favorable de ne pas le paralléliser avec *se dire* et *faire* au niveau des relations syntaxiques entre introducteur et citation. Séquentiellement, les deux extraits montrent que la citation avec *j'étais là* répond dans le récit à une action (lire le mail) ou le tour rapporté de quelqu'un d'autre (le discours de la mère). En ceci, *j'étais là* se différencie aussi clairement des deux autres introducteurs malgré certains parallèles, surtout avec *faire*, que je présenterai par la suite.

Le troisième extrait va illustrer plusieurs occurrences de *faire* comme introducteur. *Faire* est plus fréquent à la troisième personne du singulier (n=35) qu'à la première personne du singulier (n=14). À l'exception de trois cas, toutes les occurrences sont au présent. L'introducteur est utilisé avant tout dans des dialogues reproduits lors desquelles les deux voix du dialogue sont introduites avec *faire*. Une caractéristique grammaticale de la troisième personne du singulier est que 25 des 35 cas sont utilisés avec le pronom d'objet indirect *me*, qui est, en l'occurrence, le locuteur ou la locutrice. L'extrait suivant illustre cette distribution de l'objet ainsi que les spécificités formelles des citations mêmes qui sont mises en scène

phonologiquement et corporellement. Dans cet extrait 3, Francis se plaint d'un professeur (pour qui il travaille aussi) avec qui il voulait faire un rendez-vous ce qui s'avère être une tâche difficile.

Ex. 3: appel téléphonique (Pauscaf_02_17, 16min40)

- 01 >il m'ignorait tout le temps<,
02 =il m'répondait même pas fi- il m'répondait à d'autres
03 trucs,
04 il m'disait juste oui non,
05 >mais il me répondait pas (.) pour le rendez-vous,<=
06 MAT: =mm,
07 FRA: =>du coup à la fin **je fais**< mais écoutez:,
08 est-ce qu'on peut au on fait un rendez-vous OU (.)
09 au moins (.) un appel téléphonique,
10 .h pis **il m'fait:t**^eu:h oui pour appel téléphonique,=
11 =voici mon numéro:,
12 appelez-moi: euh telle heure eu:h (.) vendredi?
13 (0.5)
14 FRA: vendredi j'l'appelle (.) Exactement à c't'heure,=
15 =**il m'fait** (.) +↑oui: #qui c'est, ((imitant une voix basse))+
fig #fig1



- 16 (0.7)
17 FRA: a:h ben:: super,
18 t'es super bien organisé quoi.

Dans cet extrait, *faire* introduit des tours de paroles d'un dialogue. Entre l'introducteur et le discours rapporté direct – qui constitue l'objet direct du verbe *faire* – sont insérées des particules d'amorces (Guérin & Moreno 2015) (lignes 07, 10 et 15) ou une pause (ligne 16), pratique fréquente dans mes données (n=27), qui « contribue [...] à aider l'interlocuteur à décoder la façon dont le locuteur conçoit le sens purement propositionnel exprimé et se positionne par rapport à celui-ci » (Dostie & Pusch 2007: 5). Cette pratique est aussi fréquente avec *j'étais là* (n=36), ce qui renforce la thèse selon laquelle les deux introducteurs sont utilisés quand les locuteurs.rice.s se positionnent émotionnellement par rapport à la scène rejouée. Contrairement à *j'étais là*, *faire* est davantage utilisé pour rapporter le discours des autres que celui du locuteur ou de la locutrice même. Dans l'extrait présent, le discours de l'autre, notamment du professeur, est mis en scène avec des imitations prosodiques de la voix de ce dernier (lignes 10-12 et 15) ainsi qu'avec un geste iconique du téléphone (ligne 15,

figure 1), ce qui est moins le cas avec l'emploi de la première personne du singulier (ligne 07). Comme *j'étais là*, l'introducteur *faire* est donc utilisé quand une scène rapportée est dramatisée. L'imitation de la prise du téléphone (figure 1) est un geste iconique qui contribue, avec l'imitation de la voix basse du professeur, à l'effet de la mise en scène.

Ce dernier extrait a montré qu'il existe certains parallèles entre *être là* et *faire* employés comme introducteurs : malgré leurs statuts syntaxiques différents, les deux formes projettent des citations qui sont mises en scène d'une manière mimo-gestuelle dans des narrations émotionnelles. Une différence est la distribution fonctionnelle : tandis qu'*être là* introduit des tours réactifs, souvent mis en scène comme étant spontanément produits, qui clôturent un tour de parole, *faire* est utilisé pour reproduire des dialogues avec des prises de paroles de deux participant.e.s. Il est également intéressant qu'*être là* rapporte avant tout du discours du locuteur ou de la locutrice même alors que *faire* semble être préféré pour rapporter le discours des autres.

Après avoir montré comment *être là* et *faire* sont utilisés dans la conversation, j'aimerais illustrer, à l'aide d'un dernier extrait, comment *se dire* est utilisé, notamment pour partager une réflexion avec l'interlocutrice. Elinda parle d'un cours de gym pendant lequel elle a failli s'évanouir. Avant le début de l'extrait, elle décrit la fin du cours où le prof leur a fait faire des étirements :

Ex. 4: je m'asseye (Pauscaf 20, 32min13)

01 ELI: et je suis <relevée> j'ai vu ↑tout noir,

02 **j'étais là** #(1.1)#

fig #fig1 #



03 EKT: >ah merde,<=

04 ELI: =>**je me suis dit**< (.) je m'asseye ou j'fais quoi,=

05 ==>>mais t'sais<< tout le monde était debout,

06 >alors **je me suis dit**< je vais pas m'asseoi:r,

07 je vais attendre °qu'° ça revienne,

08 si ça revient pas je m'asseye,

09 ((rire des deux,(0.9))

Le discours rapporté direct constitue l'objet direct du verbe *se dire*, qui est ici au passé composé (n=46), forme plus fréquente que l'imparfait (n=8) et que le présent (n=38). Les citations avec *se dire* permettent aux locuteur.rice.s de partager un raisonnement avec l'interlocuteur.rice. La différence avec *être là* est non seulement que la citation même n'est pas mise en scène corporellement, mais aussi qu'elle n'a pas la fonction d'évaluer ce qui précède. Au lieu de réagir à ce qui est préalablement décrit, *je me suis dit* introduit une conséquence, un pas suivant dans l'avancée du récit qui est présenté comme un raisonnement.

Les gestes qui accompagnent les citations (lever des sourcils, ligne 07 ; pointage, ligne 09) ne sont pas des gestes qui permettent à la locutrice d'exprimer son attitude. Il s'agit plutôt de gestes que Streeck (2009) nomme « *ceiving or ception* » (*ibid.*: 9, en italiques dans l'original). Il décrit ces gestes comme « a bodily form of conceiving, i.e. of conceptually structuring content to be articulated in speech [...] », qui sont utilisés « whenever speakers, without attending to the process and without wishing to depict anything, use their hands to give form to – i.e. construe – content » (*ibid.*). Par rapport aux gestes faisant partie intégrante des citations avec *j'étais là* ou *faire*, et qui peuvent eux-mêmes former un acte communicatif, *je me suis dit* introduit du discours rapporté direct dans lequel les gestes ont une fonction de structuration et ne fait pas partie intégrante de la mise en scène. Ceci pourrait être une raison pour laquelle il n'y a, avec *je me suis dit*, aucune occurrence avec des gestes seulement. Une deuxième raison pourrait être liée à la sémantique du verbe, qui renvoie, bien évidemment, à l'introduction d'un *discours* et non d'un mime. Une troisième raison qui semble pertinente pourrait être la fonction des citations avec *je me suis dit* dans la séquence : elles introduisent des raisonnements qui sont moins chargés émotionnellement que les citations avec *j'étais là*. Comme elles ne sont pas non plus réactives à quelque chose qui précède mais plutôt conséquentielles, une attitude envers une action précédente ou parallèle ne doit pas être mise en scène comme c'était le cas avec *j'étais là* ou *il/elle (me) fait*. *Se dire*, le troisième introducteur analysé, se distingue des deux autres par le fait qu'il est utilisé dans des contextes qui pourraient être décrits comme non-émotionnels. Il est utilisé pour rapporter des réflexions, ce qui rend ces citations, dans la plupart des cas, non-animées au niveau prosodique ou mimogestuel.

Cette analyse qualitative de quatre exemples a pu montrer que les trois introducteurs ne se distinguent pas uniquement en fonction de leurs statuts syntaxiques et du matériel cité. La considération des séquences dans lesquelles *être là*, *faire* et *se dire* apparaissent a également mis en évidence la distribution fonctionnelle spécifiques des trois introducteurs.

6. Conclusion

Nous avons vu dans ces analyses qualitatives que malgré la similitude des trois introducteurs concernant leur rôle grammatical, le fait d'introduire une citation, ces formes diffèrent quant à leur fonction dans l'interaction et quant au matériel qu'elles introduisent. Tandis qu'*être là* et *faire* permettent aux locuteur.rice.s de partager leur attitude à travers des citations corporellement animées, *se dire* est un moyen de rendre public un raisonnement avec du matériel linguistique, qui peut toutefois être structuré par des gestes. Malgré leurs parallèles actionnels, *être là* diffère de *faire* dans son implication dans l'organisation séquentielle : il est toujours réactif à une action précédente, comme dans les extraits 1 et 2, et contribue à la clôture d'une séquence – et donc à un changement de locuteur.rice –, tandis que *faire* est utilisé dans des dialogues rapportés (voir ex. 3) où la voix d'un tiers est avant tout reproduite. Au niveau syntaxique, nous avons pu observer une gradualité de relations plus ou moins étroites entre introducteur et citation. Tandis que *se dire* peut introduire du discours rapporté indirect et direct, *faire* introduit des paroles ainsi que des mimes. Dans les deux cas, la citation occupe la place de l'objet direct, et cela, comme je l'ai proposé dans ma partie théorique, indépendamment de la forme de la citation. D'un point de vue syntaxique, *être là* doit être classifié entièrement à part syntaxiquement. L'idée que la citation spécifie le déictique pourrait être reformulée en disant qu'*être là* "projette du discours à venir", suivant ainsi l'idée d'une grammaire *on-line* (Auer 2005) ou émergente (Hopper 1987; Hopper 2011). Ayant

recours aux idées de la projection et de la temporalité inhérente à la conversation, ces deux approches grammaticales partent du principe que les locuteur.rice.s produisent les formats grammaticaux en parlant, utilisant et formant ainsi une grammaire qui sert à des fins – et des demandes – interactionnelles de la conversation en cours. L'idée est que c'est la langue parlée *dans l'interaction* qui contribue à façonner notre grammaire, car c'est dans cet environnement-là que la grammaire est utilisée pour agir. Ces conceptions de la grammaire rejoignent l'idée de Claire Blanche-Benveniste selon laquelle, en regardant de plus près la langue parlée, notre conception de la grammaire changera. Et si on veut aller encore plus loin, en reprenant de Cornulier (1978), il faudrait, particulièrement dans le cas du discours rapporté direct, se faire à l'idée que le mimo-gestuel fait partie intégrante de l'interaction de tous les jours. Il pourrait donc être utile d'élargir, comme proposé par Keevallik (2018), au moins dans ce cas précis du discours rapporté direct, notre grammaire par des mimes.

7. Bibliographie

- Abouda, Lotfi & Nina Rendulić. 2017. Séquence d'introduction de discours représenté : *faire* ou *dire* ? Discours [online] 21.
- Auer, Peter. 2005. Syntax als Prozess. InLiSt 41.1-35.
- Blanche-Benveniste, Claire. 1981. Le complément verbal : valence, rection, associé. *Recherches sur le français parlé* 3.57-98.
- . 1988. Constructions verbales "en incise" et rection faible des verbes. *Recherches sur le français parlé* 9.53-73.
- . 2007. Corpus de langue parlée et description grammaticale de la langue. *Langage et société* 121-122.129-41.
- . 2010. Les pseudo-clivées et l'effet deux points. *La Parataxe. Structures, marquages et exploitations discursives. Tome 2*, ed. by M.-J. Béguelin, M. Avanzi & G. Corminobœuf, 185-217. Bern: Peter Lang.
- Cheshire, Jenny & Maria Secova. 2018. The origins of new quotative expressions: the case of Paris French. *Journal of French Language Studies* 28.209-34.
- Couper-Kuhlen, Elizabeth & Sandra A. Thompson. On assessing situations and events in conversation: 'extraposition' and its relatives. *Discourse Studies* 10.443-67.
- de Cornulier, Benoît. 1978. L'incise, la classe des verbes parenthétiques et le signe mimique. *Cahier de linguistique*.
- Dostie, Gaétane. 2020. Deixis à l'imaginaire et périphrases en être et faire préfaçant une scène recréée (j'étais comme, t'es là, j'ai fait...) *Le français innovant*, ed. by F. Diémoz, G. Dostie, P. Hadermann & F. Lefeuvre, 83-106. Bern: Peter Lang.
- Dostie, Gaétane & Claus Pusch. 2007. Présentation. *Les marqueurs discursifs. Sens et variation. Langue française* 154.3-12.
- Goodwin, Charles. 1986. Between and within: Alternative Sequential Treatments of Continuers and Assessments. *Human Studies* 9.205-17.
- Goodwin, Charles & Marjorie H. Goodwin. 1992. Assessments and the construction of context. *Rethinking context: Language as an interactive phenomenon*, ed. by A. Duranti & C. Goodwin, 147-89. Cambridge: Cambridge University Press.
- Guérin, Emmanuelle & Anaïs Moreno. 2015. Présence/absence de particules d'amorce et de particules d'extension dans le discours rapporté : peut-on parler de variation? *Langage et société* 4.67-82.
- Hopper, Paul J. 1987. Emergent Grammar. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, ed. by J. Aske, N. Beery, L. Michaelis & H. Filip, 139-57.
- . 2011. Emergent grammar and temporality in interactional linguistics. *Constructions: Emerging and Emergent*, ed. by P. Auer & S. Pfänder, 22-44. Berlin: De Gruyter.

- Jefferson, Gail. 2004. Glossary of transcript symbols with an introduction. *Conversation Analysis. Studies from the first generation*, ed. by G.H. Lerner, 13-31. Amsterdam/Philadelphia: Benjamins.
- Keevallik, Leelo. 2018. What Does Embodied Interaction Tell Us About Grammar? *Research on Language and Social Interaction* 51.1-21.
- Mosegaard Hansen, Maj-Britt. 2000. The syntactic and semiotic status of direct quotes, with reference to French. *Transactions of the Philological Society* 98.281-322.
- Rabatel, Alain. 2004. L'effacement énonciatif dans les discours rapportés et ses effets pragmatiques. *Langages* 4.3-17.
- Secova, Maria. 2015. Discours direct chez les jeunes : nouvelles structures, nouvelles fonctions. *Langage et société* 151.131-51.
- Streeck, Jürgen. 2002. Grammars, Words, and Embodied Meanings: On the Uses and Evolution of *So* and *Like*. *Journal of Communication* 52.581-96.
- . 2009. *Gesturecraft. The manufacture of meaning*. Amsterdam/Philadelphia: Benjamins.
- Tagliamonte, Sali A. & Alexandra D'Arcy. 2007. Frequency and variation in the community grammar: Tracking a new change through the generations. *Language Variation and Change* 19.199-217.
- Tannen, Deborah. 2007. *Talking voices. Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.

8. Annexe

Conventions de transcriptions suivant Jefferson (2004).

[début du chevauchement
]	fin du chevauchement
=	latching (pas de pause, pas de chevauchement)
(0.7)	pause mesurée en secondes et dixièmes de secondes
(.)	micro-pause (en dessous de 0.2 secondes)
wo-	mot tronqué
wo:rd	syllabe allongé
?	intonation finale montante
.	intonation finale descendante
,	intonation continuative
<u>word</u>	accentuation
°word°	moins fort que discours autour
<word>	discours ralenti
>word<	discours accéléré
WORD	plus fort que discours autour
↑word	montée de volume (se réfère à la syllabe suivante)
^	liaison
.h	inhalation
h.	exhalation
((laughter))	commentaire du/de la transcriv.eur.e

Décrire le ‘français tout court’ : pourquoi et à quelles conditions

Françoise Gadet, Université Paris-Nanterre et MoDyCo
Emmanuelle Guerin, Sorbonne Nouvelle et DILTEC

Parmi les nombreux apports de l’œuvre de Claire Blanche-Benveniste (CBB), nous nous arrêterons ici à son idée de faire de l’étude du français parlé une voie d’accès privilégiée au « français tout court » (Blanche-Benveniste, 1983). Cette expression, récurrente dans l’œuvre de CBB, renvoie à l’élargissement des données du français, en y intégrant la langue parlée sans souci de savoir si les formes sont ou non regardées comme standard : « Je voudrais défendre l’idée que la connaissance de la langue parlée fait progresser la description grammaticale de l’ensemble de la langue » (2007 : 129). L’appareillage grammatical traditionnel demande dès lors à être revisité, pour englober autre chose que des productions normées. Avec les données orales, elle élargit le champ d’application de la grammaire.

Tout un pan des sciences du langage est aujourd’hui consacré à des travaux sur l’oral et, appuyées sur des avancées technologiques, des études illustrent les spécificités de (certains) oraux. Pour autant, la perspective de connaissance du ‘français tout court’ ne vise pas à penser une autre grammaire pour ce type de données, mais à prendre en compte l’ensemble des productions. CBB affirmait : « pour dégager des systèmes de formes cohérents dans la description, il est très souvent nécessaire de rassembler des usages de la langue qui paraissent se situer à des ‘niveaux’ de normes hétérogènes » (1983 : 27).

Après avoir tenté de déployer la notion de ‘français tout court’, nous illustrerons ses effets avec des observations sur le fonctionnement de pronoms dans des énoncés issus d’un corpus de données ordinaires. Nous montrerons que la description traditionnelle ne permet pas de rendre compte de tous les faits, alors même que la communication se fait.

1. Le français tout court

CBB a ouvert la description grammaticale aux données orales en fournissant un cadre théorique et méthodologique pour traiter de formes jusque là au mieux ignorées, au pire invalidées. Elle a en cela tracé une voie aux chercheurs s’inscrivant dans le champ de la « linguistique de l’oral » (à supposer qu’une telle dénomination ne soit pas d’emblée invalidée par l’idée de ‘français tout court’). Après plus de 40 ans de travaux et de publications, il apparaît heureusement que les spécificités de l’oral (spontané en particulier)⁰ ne sont plus regardées comme des « déviances », mais analysées pour ce qu’elles sont, sans que soit mis en cause le système de la langue.

1.1. La conception du ‘français tout court’

Cette évolution dans les pratiques grammaticales et, par conséquent, dans les représentations (du moins celles des linguistes) est imprégnée de la considération des « différents façonnages de l’information » (Blanche-Benveniste, 1997 : 68)⁰, eux-mêmes tributaires des caractéristiques des contextes de production, entre autres des effets du medium

⁰ Dans son ultime écrit, elle résume les 4 caractères inhérents à la langue parlée (2010, p. 81) qui ne seront pas sans effet sur sa description : 1) la présence de bribes (qui sont donc à exploiter, et non à écarter), 2) la prosodie, 3) la restriction par le lexique et 4) la variabilité des constructions syntaxiques.

⁰ Le texte de 1994, relevant pourtant d’une période proche, parle non pas de « façonnage » mais de « façonnement ». Il ne nous a pas semblé que le sens visé différerait.

oral ou écrit. Il s'agit de rompre avec « les grands mythes séparateurs » (Blanche-Benveniste & Jeanjean, 1987 : 11), qu'elles présentent en ces termes :

« Quand on parcourt une documentation sur le français parlé depuis le début du vingtième siècle, on est frappé par la persistance de quelques grands mythes qui ont pour effet de “séparer” ce qu'on appelle le “français parlé” de l'ensemble de la langue ; on le voit retranché, mis à l'écart – pour le décrier comme pour l'encenser. Assimiler le parlé au populaire, c'est le retrancher du français légitime ; y voir la source des innovations ou des conservatismes, c'est le retrancher dans le temps ; opposer le parlé à l'écrit, c'est lui assigner une place bien à part ; l'accabler d'étiquettes et de “niveaux”, c'est vouloir le cantonner dans certaines activités de langage et l'exclure des autres. »

Le propos est ici à la fois plus original et plus complexe qu'il ne semble à première vue. Le travail sur l'oral n'est pas le pendant du travail sur l'écrit avec une approche différente, ce qui perpétuerait l'idée de deux langues différentes (Cappeau, dir, 2021 pour une critique). Au contraire, il s'agit de montrer qu'il est toujours question de français (tout court), dans la diversité de ses actualisations. Le ‘français tout court’ implique en effet de décrire le système de telle sorte que les « façonnages », plus ou moins contraints par la situation, soient partie prenante du cadre descriptif (Blanche-Benveniste, 2007 pour la relation entre données orales et description grammaticale). Il ne s'agit pas seulement d'affirmer un point de vue non normatif sur la langue mais d'approfondir la compréhension de celle-ci. Certains faits demeurent en effet peu analysables tant qu'on ne les rapproche pas de faits réputés non standard et, à ce titre, peu ou pas pris en compte dans les grammaires. L'intérêt d'élargir les données peut être illustré avec *donne-nous-en*. Si dans cette forme le [z] est motivé par la liaison, ce n'est pas le cas dans la forme isolée *donnes-en* (avec un -s à l'impératif), analysable grâce à des formes comme *donne moi z'en*, où *z'en* constitue une variante postverbale de *en* (Blanche-Benveniste, 1983) – forme accessoirement évaluée comme non standard.

Considérer que la langue se révèle dans ses actualisations où l'information est façonnée en relation avec les situations de communication dans lesquelles les locuteurs sont impliqués revient à parler de *variation* (Gadet, 2018). Bien que la démarche de CBB soit avant tout syntaxique, avec la notion de ‘français tout court’, elle alimente la réflexion pour une « sociolinguistique de la langue », un champ qui aborde le social par le prisme de la langue (Gadet, 2012). Observer les différents façonnages de l'information impliquant le repérage des contextes de production, cela revient à mettre en relation les possibles de la langue avec les éléments constitutifs de la situation de communication. On se donne ainsi un cadre qui ne repose ni sur des évidences linguistiques (standard vs non standard, écrit vs oral) ni sur des évidences sociales fondées sur des catégories sociodémographiques attribuant aux locuteurs des usages exclusifs.

Ainsi envisagée, l'étude du ‘français tout court’ fait écho à l'idée d'un continuum sur lequel se distribuent les actualisations du système, pour mettre en évidence que la langue française ne saurait être saisie à travers un unique état, fut-il socialement érigé en norme.

1.2. Les paramètres de la situation de communication

Partant, le tableau esquissé par CBB peut être complété avec la proposition de Koch & Oesterreicher (2001) visant notamment à déterminer des paramètres pour caractériser le comportement communicatif des interlocuteurs par rapport aux composants situationnels et contextuels. En s'appuyant sur le travail de Söll, qui distingue les plans médial (graphique/phonique) et conceptionnel (écrit/oral), ils théorisent l'opposition entre *immédiat* ou proximité et *distance* et caractérisent les comportements communicatifs à travers la combinatoire de paramètres des contextes de communication : toute interaction peut être

située quelque part entre le pôle de la proximité et celui de la distance. Dans Guerin (2017), ‘proximité’ et ‘distance’ communicatives sont mis en relation avec la connivence qu’entretiennent les interactants lors d’un échange.

Ainsi, un échange spontané en face à face entre intimes à propos d’une thématique partagée engendre un comportement communicatif illustrant la proximité : les locuteurs peuvent élaborer leurs énoncés en comptant sur une forte connivence, soit une somme de savoirs et d’expériences partagés. Et l’information est façonnée en conséquence. Les discours s’élaborent dans l’échange à partir de référents implicites (soit parce que perceptibles et accessibles de la même manière par les participants, soit parce que relevant de communs savoirs, expériences ou appréhension du monde) et l’interprétation repose sur des inférences. Une forte connivence des interactants permet ainsi d’analyser des productions même si elles échappent à une structuration grammaticale standard et de restituer des éléments informatifs non mentionnés parce que supposés saillants dans la mémoire discursive partagée ou récupérables en contexte. La coopération communicative (un des paramètres chez Koch & Oesterreicher) autorise l’absence d’éléments qui, dans un cadre standard, assurent la cohérence syntaxique, sémantique et pragmatique.

A l’inverse, lorsque les locuteurs sont contraints par le contexte à un comportement communicatif tendant à la distance, les productions, pour être efficaces, seront façonnées avec le souci d’accompagner l’interprétation en limitant les implicites et les jeux d’inférence. Ainsi des écrits littéraires : pour permettre l’interprétation d’un énoncé, son producteur doit fournir les éléments nécessaires à la restitution de l’information telle qu’il l’envisage, sans compter sur un partage de savoirs. Sachant que la norme prend pour caution les écrits littéraires, on voit les effets potentiellement négatifs du statut de « bon usage » absolu qui lui est accordé (Guerin, 2021). Le fait que les locuteurs aient pour repère une forme de langue façonnée pour des types de situation tendant à la distance peut être source d’insécurité dans les interactions courantes : ce qui est socialement valorisé ne correspond pas à ce qui est adapté à la proximité.

Ainsi, le ‘français tout court’, mettant en lumière la façon dont la langue est façonnée par les usagers pour être pertinemment actualisée en fonction du contexte de production, esquisse un cadre sociolinguistique. Données et métadonnées renseignant la situation de communication sont dès lors indissociables, alimentant un regard qui n’est plus celui de la grammaire traditionnelle, laquelle ne soulève pas la question du contexte de production puisque le corpus légitime relève de comportements communicatifs tendant à la seule distance.

1.3. Quelles données pour le « français tout court »

L’objectif de décrire le ‘français tout court’ interroge l’amplitude et la spécificité des données, le qualificatif « orales » (il en va de même pour « écrites ») ne suffisant pas à les caractériser (Blanche-Benveniste & Bilger, 1999). Comme nous l’avons vu, le comportement communicatif des locuteurs est aussi conditionné par d’autres paramètres. Certes, le médium a des effets sur les productions, qui peuvent être le fruit du temps de production et de réception, du non-partage du cadre a minima spatial ou encore du prestige de l’écrit dans les représentations conduisant parfois à user de formes standard à l’encontre du contexte ; autant de paramètres qui produisent des points communs entre écrits. Cependant, d’autres paramètres entraînent au moins autant de variabilité. Ce qui explique par exemple que deux productions

écrites comme un SMS et un texte de loi soient formellement très éloignées. Le même raisonnement vaut évidemment pour les productions orales.

Oral vs écrit ne constitue donc pas une modalité suffisante de constitution de corpus. Si l'on s'intéresse au 'français tout court' et non à un ensemble de productions circonscrit à certains paramètres comme la dimension médiale, alors les données ne vont pas sans des métadonnées déterminant la position sur le continuum proximité/distance. De ce point de vue, il n'y a pas lieu de privilégier des données *orales* ou *écrites*. Ainsi, le comportement communicatif des locuteurs dans un contexte de messagerie instantanée médiée par ordinateur est souvent plus proche de celui de locuteurs en échange oral spontané que de celui de correspondants par courriel. Guerin & Moreno (2016) ont montré des façonnages similaires avec le discours rapporté.

Il faut donc réfléchir aux critères sur lesquels appuyer le recueil de données et de métadonnées. Le cadre de la « sociolinguistique de la langue » incite à réévaluer la façon dont la sociolinguistique a traité des données. Les travaux les plus fréquents pour étudier la variation, de facture variationniste et visant au quantitatif, reposent sur des corpus construits sur le principe d'une supposée co-variance, qui repose sur la rencontre entre du linguistique/langagier et du social, sans que l'entrée en rapport des deux ordres pré-existants à la rencontre soit problématisée (Gadet & Guerin, 2015). Or, les échanges ordinaires laissent apparaître des différences dans les productions, dans une même situation, de locuteurs relevant d'une même catégorie socio-démographique.

Les métadonnées doivent donc tendre à une certaine précision, la granularité devant évaluer le degré de connivence entretenue par les interactants lors de l'échange. Autrement dit, les métadonnées doivent permettre de situer les discours sur le continuum proximité/distance. Un corpus homogène rassemblerait ainsi des données issues de contextes pour lesquels on reconnaît un même « relief communicatif » (Koch & Oesterreicher, 2001). Dans ces conditions, il est possible de mettre en lumière le comportement référentiel et/ou procédural des unités de la langue, qui se révèle seulement dans certains contextes. C'est ce que montrent de façon évidente, par exemple, les emplois métaphoriques : le sens d'une unité lexicale peut recouvrir des réalités différentes selon ce que partagent les interactants.

2. Observations sur le fonctionnement de certains pronoms à partir des données de MPF

La démarche de CBB l'amène donc à porter un soin tout particulier à la constitution de ses données, donc à une solide réflexion sur les corpus, même si elle n'a pas franchi le pas de ce que peut apporter une perspective sociolinguistique pour poursuivre son projet⁰.

2.1. Un corpus et ses métadonnées

Il est bien connu que tout corpus n'est pas adapté pour tout type d'analyse et que les modalités de sa conception le rendront plus adéquat pour tel type d'exploitation que pour tel autre. Le corpus MPF s'accommode bien à un cadre de sociolinguistique de la langue supposant que les métadonnées aideront à situer les comportements communicatifs des locuteurs. La liste des paramètres proposée par Koch & Oesterreicher constitue dans cette optique une base essentielle, adaptable à toute situation.

⁰ Au milieu d'une méticulosité sur les données rare même chez les linguistes de corpus, CBB n'a pas fait appel à des données de français hors de France, où l'on trouve pourtant de quoi conforter son analyse, comme cet exemple de Terre-Neuve cité par Gadet (2018), qui enrichit d'une profondeur diachronique l'analyse du pronom *z'en* évoquée plus haut : *ma mère me fait manger z'en* (= m'en fait manger).

Si le corpus MPF (*Multicultural Paris French*, voir Gadet dir. 2017 pour des détails sur le corpus)⁰ apparaît adapté pour illustrer notre propos, ce n'est pas parce que les formes qui s'y trouvent pourraient être caractérisées comme « orales » ou comme « non standard », mais bien parce qu'elles sont « situées » et produites dans des situations d'immédiat communicatif, ayant été recueillies selon un protocole destiné à entraîner la plus forte connivence possible entre les interactants (dont l'enquêteur).

Chaque enregistrement en effet offre des interactions pour lesquelles une fiche renseigne les éléments qui nourrissent le degré de connivence des intervenants : il peut s'agir d'échanges écologiques (auto-enregistrements) entre pairs ; dans d'autres cas, un enquêteur membre de l'équipe devait avoir un lien antérieur avec l'enquêté (membre de la famille, ami, éducateur scolaire, partage associatif...). Dans tous les cas, le protocole vise à obtenir une parole dans un contexte favorisant des comportements communicatifs tendant vers la proximité, l'hypothèse initiale ayant été que pourraient alors apparaître des formes ou des usages qui n'ont que peu de chances d'être observés dans des situations plus classiques pour l'obtention de données, en particulier dans les entretiens dits sociolinguistiques. Du fait que tous les locuteurs enregistrés sont des locuteurs du français (la plupart monolingues), ces formes et ces usages viendront enrichir le 'français tout court'.

2.2. Le choix des pronoms

Un relevé a été pratiqué manuellement à partir de 75 enquêtes représentant 51 heures d'enregistrement. Il était en effet impossible de procéder de façon automatique pour relever les 291 occurrences de pronoms que nous avons retenues, présentant divers intérêts de formes ou d'emplois « atypiques » aux yeux de la grammaire traditionnelle.

Par essence, les pronoms mettent en jeu une procédure interprétative supposant une collaboration entre les interactants. Alors que, dans des productions illustrant la distance, cette procédure conduit le récepteur à activer le référent évoqué par l'antécédent présent en cotexte sous forme de syntagme nominal, ce n'est pas toujours ce qui se passe dans les productions d'immédiat communicatif. Il arrive en effet très souvent que soit requise la récupération d'un référent implicite dans l'ensemble des savoirs supposément partagés (Jahandarie, 1999). Ce qui éclaire l'observation selon laquelle « certaines conversations orales à bâtons rompus n'ont aucun syntagme nominal pour sujet, tous les sujets étant des pronoms » (Blanche-Benveniste, 1997 : 105 ; voir aussi Blanche-Benveniste, 1994). Dit autrement, il est prévisible que les occurrences de pronoms seront plus nombreuses lorsque la connivence entre interactants est renforcée – au point qu'on pourrait envisager d'en faire un principe d'évaluation du degré d'immédiat d'un corpus. Sachant qu'un usage pertinent des pronoms requiert une coopération communicative suffisante pour que le référent soit activé, on peut s'interroger sur les effets quant à la diversité de l'emploi des pronoms que produirait un renforcement de cette coopération lors du recueil du corpus.

Notre objectif ici n'est nullement de synthétiser l'analyse de tel ou tel pronom, ce qui a déjà été fait et le sera sans doute encore, pas davantage de prétendre mettre à jour des faits jusque-là inédits. Il s'agit de réfléchir aux procédures d'interprétation des pronoms dans des énoncés qui ne répondent pas aux procédures interprétatives des grammaires traditionnelles. Certains énoncés de MPF nous conduisent en effet à émettre l'hypothèse que, lorsque le référent activé par le pronom est suffisamment saillant dans la mémoire discursive partagée

⁰ Toutes les enquêtes de MPF déposées sur le site d'Ortolang et librement consultables se composent de 3 éléments : le son en format .wav, une transcription effectuée sous Praat et deux types de fiches de métadonnées, l'un renseignant la situation d'enregistrement, l'autre les protagonistes et leurs liens.

lors de l'échange, alors le marquage du pronom qui guide sa récupération ne serait plus indispensable : le sens procédural ne se saisirait alors plus par le seul plan morphosyntaxique.

On comprend dès lors la possibilité de ces emplois atypiques de pronoms, sans qu'il soit besoin de convoquer les idées de fautes ou de ratés.

2.3. Des usages de quelques pronoms dans MPF

Une première série d'exemples montre la rupture de la chaîne d'accord en genre, qui s'opère toujours au profit d'un marquage au masculin (on n'a pas relevé de cas où un référent masculin serait repris par un pronom marqué du féminin). Le poids de ces cas n'est pas tout à fait anecdotique, puisqu'on a relevé 127 énoncés exemplifiant l'absence de marquage (en genre ou en nombre) sur les 75 enquêtes dépouillées, impliquant des locuteurs différents et des situations différentes.

Le fait qu'en français, le neutre soit évoqué par le masculin singulier, peut fournir une hypothèse : puisqu'il n'y a pas de guidage explicite vers le référent, alors c'est à une forme neutralisée qu'il est fait appel.

- (1) On dit narvali pour **les filles** (.) on fait wesh narvali et **ils** le prennent bien (.) mais ça dépend aussi des filles. (Anna2, 663, Aziz)⁰
- (2) **Mes copines** à moi **ils** sont sauvages hein (Emmanuelle4, 1450, Sandrine)
- (3) **Les meufs** des du Terroir **ils** vous cassent votre (bouche) hein. (Joanne11, 992, Lupita)
- (4) **Même ma grand-mère et mon arrière-grand-mère** comme je t'ai dit **ils** sont nés à Nanterre donc **eux ils** parlent français même l'autre (Nacer2, 503, Hanane)

Dans ces quatre exemples, un référent manifestement féminin est activé par un pronom marqué du masculin/neutre. L'exemple 5 présente l'intérêt, après un premier marquage au masculin, d'offrir un marquage féminin du clitique sujet lors de la seconde activation du référent :

- (5) Voilà ça ça c'est **les petites pouffes** qui qui qu'**ils** ont l'impression que quand on les regarde il faut il faut pas les admirer pour ce qu'**elles** sont mais pour leurs marques (Baligh1, 838, Lucile)

Risquons une hypothèse pour (5) : l'introduction d'un nouveau référent activé par *on* pourrait brouiller l'interprétation. Il faudrait alors désambiguïser, en rendant explicite, par la marque du féminin, le référent visé, *les petites pouffes*.

Comme pour le genre, le marquage du nombre ne semble pas toujours indispensable, comme dans les exemples de (6) à (9) :

- (6) **Ces gens-là** ils ils ils on va dire **il** va te peut-être donner la place pendant un an (Wajih5, 1189, Amir)⁰
- (7) Sinon des fois on dit des trucs euh (.) on a d- on dit des trucs en arabe comme ça

⁰ Les exemples de MPF conservent ici les notations originales : les enquêtes sont désignées au moyen du prénom de l'enquêteur suivi du numéro de l'enquête, puis est précisé le minutage de l'extrait. Le lecteur pourra ainsi mesurer la diversité des sources des exemples.

⁰ Il n'est pas exclu que les trois premiers *ils* aient pu être transcrits *il*.

les autres il comprend pas quand on fait un truc. (Wajih1, 1268, Salim)

- (8) Il y a **des gens** ils revendent des organes euh genre ils tuent des enfants (.) et **il** prend **il** prend les organes pour les redonner aux vieux riches (Nawal1, 2131, Olfa)
- (9) Après **s'ils** arrivent ben plus tard **il** pourra apprendre les mots de wesh wesh (Joanne10, 2410, Marie)

Si ces exemples ne remettent pas en cause la fonction d'activation d'un référent identifiable du pronom (en cotexte, en contexte ou dans la mémoire discursive partagée), ils permettent de relativiser l'importance à accorder à leur marquage morphologique. Le non-respect de la chaîne d'accords n'empêche pas la restitution du référent. La présence d'un pronom neutralisé pour le genre et/ou le nombre suffit à réactiver l'information et à assurer l'interprétation que le locuteur suppose partagée et saillante dans l'échange (ce que l'on peut inférer d'un énoncé comme (5)). Cette affirmation se renforce au regard de cas où le pronom (hors position sujet) est absent, comme dans les exemples suivants⁰ :

- (10) Il y a des gens ils mettent des tickets il y en a ils Æ^0 mettent pas. (Aristide5b, 29, Louis)
- (11) C'est comme Amsterdam je Æ suis allé il y a pas longtemps (Emmanuelle6, 137, Olivier)
- (12) Ils partageaient l'argent et moi j'ai pas voulu Æ prendre ils étaient là prends prends prends mais j'ai pas voulu Æ prendre. J' Æ ai donné à à mes potes et tout (Nawal2, 2447, Albert-Youssef)

En somme, regarder le fonctionnement (tout court) des pronoms sans se limiter à des données n'illustrant qu'un seul type de comportement communicatif, invite à revoir sa caractérisation. Les pronoms, qui n'ont qu'une référence actuelle non autonome (Milner, 1978), ont pour fonction de déclencher une procédure d'activation d'un référent nominal ou représentant (selon les termes de Benveniste, 1966) et, à ce titre, ils seront plus ou moins marqués selon qu'il est plus ou moins nécessaire de guider la restitution du référent. Ainsi, le marquage des pronoms est relatif à la connivence entre interactants : lorsque la coopération communicative est suffisamment intense pour éviter l'ambiguïté, les pronoms peuvent revêtir une forme neutralisée (le masculin singulier en français), voire être absents. En revanche, dans les contextes déclenchant des comportements communicatifs relevant de la distance, lorsque l'émetteur n'a pas la garantie que le référent est accessible, alors les pronoms sont marqués du genre et du nombre.

Comme pour toute description grammaticale, la limitation des observables aux seules formes standard invalident nombre de formes pourtant attestées dans les pratiques ordinaires de locuteurs français, à ce titre parties prenantes du 'français tout court'. Ce qui peut ainsi être dit des pronoms vaut pour toutes les unités de la langue : le sens qui peut leur être attribué compte tenu de leur emploi dans un contexte donné ne vaut pas nécessairement dans d'autres

⁰ En français, pour la position sujet, l'absence de pronom est marginale. Etant (avec le verbe) « la seule espèce de mots qui soit soumise à la catégorie de la personne » (Benveniste, 1966 : 225), il peut se comprendre comme un morphème préverbal et est souvent la seule marque de personne portée par le verbe, aux plans phonique comme graphique.

⁰ Un inconvénient de ce mode de notation est de sembler imager un manque, ce qui risque de nous ramener aux formulations que nous déplorons, en termes de manque.

contextes. Élargir le champ des observables permet alors d’appréhender l’étendue des potentialités de fonctionnement et donc le déploiement du système de la langue.

Pour ce qu’il en est des pronoms, leur marquage morphologique est relatif au contexte de production (marqués du genre et du nombre - marqués du genre ou du nombre - non-marqué - absence de pronom) mais pour autant leur emploi est contraint par le système (voir pour la position sujet). Ce qui confirme qu’il s’agit toujours bel et bien de français.

Conclusion

Nous avons ici cherché à poursuivre le fil de réflexions impulsées par les travaux de Claire Blanche-Benveniste quant au projet d’élaborer une grammaire du français riche de l’observation de phénomènes que de grands corpus de données ordinaires rendent désormais disponibles (voir Cappeau dir., 2021). On est alors conduits à revoir les cadres descriptifs traditionnels en admettant l’instabilité des catégories, compte tenu de ce que « l’intégrité des données semble mener ainsi inéluctablement à accepter l’hétérogénéité du système » (Blanche-Benveniste *et al.*, 1990 : 11). Il apparaît ainsi que l’œuvre de CBB constitue une solide base pour nourrir des réflexions sociolinguistiques, ses préoccupations avant tout syntaxiques rejoignant cette perspective, puisque le ‘français tout court’ permet la reconnaissance de locuteurs et d’usages français tout courts : si les productions standard ou non standard sont tout autant du français, alors les locuteurs non « légitimes » sont aussi français que les « légitimes ». Certes, un arrimage à la grammaire n’apparaît pas comme une orientation aujourd’hui très poursuivie en sociolinguistique, mais il nous semble productif de chercher à confronter l’étude du français parlé à la sociolinguistique de la langue.

Bibliographie

- Benveniste Emile (1966), *Problèmes de linguistique générale I*, Gallimard, Paris.
- Blanche-Benveniste Claire (1983), L’importance du ‘français parlé’ pour la description du ‘français tout court’, *Recherches sur le français parlé*, n° 5, GARS, Université de Provence, p. 23-45.
- Blanche-Benveniste Claire (1994), Quelques caractéristiques grammaticales des ‘sujets’ dans le français parlé des conversations, in *Subjecthood and subjectivity*, M. Yaguello (ed), London, Ophrys & Institut Français du Ruyau-uni, p. 77-107.
- Blanche-Benveniste Claire (1997), *Le français. Approches de la langue parlée en français*, Ophrys, Paris.
- Blanche-Benveniste Claire (2007), Corpus de langue parlée et description grammaticale de la langue, *Langage & société*, n°121-122, p. 129-14.
- Blanche-Benveniste Claire (2010), *Le français. Usages de la langue parlée*, Peeters, Leuven & Paris.
- Blanche-Benveniste Claire *et al.* (1990), *Le français parlé. Études grammaticales*, CNRS Éditions, Paris.
- Blanche-Benveniste Claire & Bilger Mireille (1999), Français parlé - oral spontané. Quelques réflexions, *Revue française de linguistique appliquée*, vol. IV-2, p. 21-30.
- Blanche-Benveniste Claire & Jeanjean Colette (1987), *Le français parlé, transcription et édition*, Paris, Didier-Érudition.
- Cappeau Paul (2021, Dir.), *Une grammaire à l’aune de l’oral ?*, Presses Universitaires de Rennes, Rennes.
- Gadet Françoise (2012) Une rencontre inachevée : la sociolinguistique de la langue et la syntaxe du français parlé, in *Penser les langues avec Claire Blanche-Benveniste*, S. Caddéo, M.-N. Roubaud, M. Rouquier et F. Sabio (dir.), Presses de l’Université de Provence, 35-43.
- Gadet Françoise (2018), *Langue et variation*, en ligne : <http://encyclogram.fr>
- Gadet Françoise (2017, Dir.), *Les parlers jeunes dans l’Ile-de-France multiculturelle*. Paris & Gap, Ophrys.
- Gadet Françoise & Guerin Emmanuelle (2015), Introduction, *Langage et société*, n°154, p. 7-15.
- Guerin Emmanuelle (2017), Éléments pour une approche communicationnelle de la variation, in *Variation en question(s)*, Bilger M., Cappeau P., Guerin E. & Tyne H. (dirs), Bruxelles : Peter Lang.
- Guerin Emmanuelle (2021), Quand la distance est dite proximité : le « FLM » en question..., *Humanités, Didactiques, Recherches*, n°1, p. 132-156.

- Guerin Emmanuelle & Moreno Anaïs (2016), Le "discours rapporté" dans les interactions orales et écrites : au-delà d'une opposition de surface, in *Actes du XXVIIe Congrès international de linguistique et de philologie romanes* (Nancy, Jacob D., Gadet F. & Lodge A. (éd.)15-20 juillet 2013). Section 9 : Rapports entre langue écrite et langue parlée, Nancy : ATILF, <http://www.atilf.fr/cilpr2013/actes/section-9.html>.
- Jahandarie Khosrow (1999), *Spoken and written discourse : a multidisciplinary perspective*, Ablex, Stamford.
- Koch Peter & Oesterreicher Wulf (2001), Langage parlé et langage écrit, in *Lexikon der Romanistischen Linguistik*, Tome 1, Tübingen, Max Niemeyer Verlag, p. 584-627.
- Milner Jean-Claude (1978), *De la syntaxe à l'interprétation*, Le Seuil, Paris.
- MPF (Multicultural Paris French) <https://www.ortolang.fr/market/corpora/mpf>

'Si vous avez quelqu'un sous la main' : les si-indépendantes en tant que format de requête

Anne-Sylvie Horlacher, Université de Neuchâtel

Simona Pekarek Doehler, Université de Neuchâtel

1. Introduction : la notion 'suspecte' de subordination

Avec l'intensification récente des approches centrées sur le langage en usage – dites '*usage based linguistics*' – l'œuvre de Claire Blanche-Benveniste est plus que jamais d'actualité. Dans les approches de ce type, les structures syntaxiques sont comprises comme des adaptations fonctionnelles à des problèmes récurrents liés à l'usage authentique du langage, qui sont conventionnalisées à travers cet usage (Bybee & Hopper 2001). L'un des résultats présumés de cette conventionnalisation est actuellement beaucoup débattu sous le terme d'insubordination (Evans 2007, Watanabe & Evans 2016 ; voir aussi les articles dans Beijering *et al.* 2019). Le phénomène peut être défini comme « l'usage conventionnalisé en tant que proposition principale de ce qui, *prima facie*, apparaît formellement comme une proposition subordonnée » (Evans 2007 : 367, notre traduction ; pour les subordonnées sans principales, voir déjà Deulofeu 1988 et les discussions dans Debaisieux 1994).

Nous nous proposons ici d'interroger le phénomène de l'insubordination dans des données interactives et dans une perspective résolument interactionnelle. Une telle approche nous semble répondre à plusieurs développements actuels dans le domaine. D'une part, les phénomènes de co-construction interactive sont considérés comme jouant un rôle important dans l'émergence des structures insubordonnées (Dwyer 2016, Evans & Watanabe 2016). D'autre part, de nombreux auteurs soulignent actuellement le besoin de modèles linguistiques mettant en avant la dimension intersubjective du langage pour pleinement comprendre le phénomène de l'insubordination (Watanabe & Evans 2016, Beijering *et al.* 2019) :

Preliminary indications suggest that we will need to draw on models of language use and semantic change that are more intersubjective than subjective if we are to fully grasp the phenomenon, and in particular those that focus on how speakers achieve intersubjective alignment. (Evans 2009 : 12)

Enfin, des recherches récentes (Lindström, Laury & Lindholm 2019, De Stefani 2021) démontrent comment une approche qui interroge l'usage du langage au sein d'échanges sociaux contribue à mieux saisir l'éventail fonctionnel, et par extension, les motivations interactives des structures insubordonnées.

Dans la présente contribution, nous nous proposons de poursuivre cette ligne d'investigation dans le champ de la linguistique interactionnelle (Ochs, Schegloff & Thompson 1996, Couper-Kuhlen & Selting 2018, Pekarek Doehler, De Stefani & Horlacher 2015), tout en rejoignant certaines réflexions de Blanche-Benveniste (1983) sur la notion 'suspecte' de subordination. Nous exposons ici une étude sur les structures en *si-indépendantes* (Corminboeuf & Jahn 2020) dans un corpus d'interactions authentiques, à savoir des appels d'urgence, en cherchant notamment à mieux comprendre la dimension fonctionnelle de ces structures et les possibles motivations interactives et intersubjectives de leur occurrence.

2. Les *si-indépendantes* dans la littérature

Sur le plan grammatical, la structure *si-indépendante* peut être décrite comme une protase qui n'est pas suivie d'une apodose – la 'reconstruction' complète de la structure étant habituellement comprise comme [si X, (alors) Y], p.ex. : "si vous avez quelqu'un sous la main, envoyez-le moi". Plusieurs auteurs considèrent toutefois qu'une telle reconstruction peut sembler « arbitraire », les données ne rendant pas toujours possible « la récupérabilité du matériel effacé » (Debaisieux, Deulofeu & Martin 2008 : 2 ; pour une critique de la notion d'ellipse dans une optique interactionnelle, voir Lindström, Laury & Lindholm 2019). En outre, la structure a parfois été analysée comme une erreur de performance, le locuteur n'arrivant pas à mener à terme son projet communicatif (pour des contre-arguments, voir ici encore Debaisieux, Deulofeu & Martin 2008 et plus récemment Debaisieux, Martin & Deulofeu 2019).

Dans la littérature, les *si-indépendantes* ont été étudiées dans de nombreuses langues (voir Günthner 2020 pour un état de l'art récent), et ont reçu des appellations diverses telles que *isolierte Nebensätze* (Buscha 1976 pour l'allemand), *isolated if-clauses* (Stirling 1999 pour l'anglais), *independent si-clauses* (Schwenter 2006 pour l'espagnol), *free conditionals* (Lombardi Vallauri 2004 pour l'Italien), *independent jos-clauses* (Laury 2012 pour le finnois), *insubordinated om-clauses* (Laury, Lindholm & Lindström 2013 pour le suédois), etc. En français, ces structures ont surtout été décrites dans une perspective typologique (Evans 2007), (macro)-syntaxique (Corminboeuf 2008, Debaisieux, Deulofeu & Martin 2008, Patard 2014, Corminboeuf & Jahn 2020) ou encore, plus rarement, interactionniste (De Stefani 2021, Horlacher 2021, Horlacher & Pekarek Doehler 2021).

En termes de fréquence, Corminboeuf & Jahn (2020) notent que les *si-indépendantes* représentent 6,7% (67 occurrences sur 1000) des constructions en *si* issues du corpus oral de français de Suisse romande *OFROM* (Avanzi, Béguelin & Diémoz 2012-2019). À noter que Corminboeuf & Jahn (2020) comptent aussi parmi les *si-indépendantes* la catégorie des hypothétiques co-construites, dont la partie [(alors) Y] est réalisée par un autre locuteur que celui qui a énoncé la partie [si X].

Sur le plan pragmatique, les constructions de type *si-indépendantes* ont été étudiées comme implémentant des actions aussi diverses que des souhaits/vœux, des invitations/propositions, des menaces (Beijering *et al.* 2019, Günthner 2020). Elles ont également été associées à des actes de requête, typiquement sous la terminologie de *if-requests* (Evans 2007, Lindström, Laury & Lindholm 2019) ou celle de *polite directives* (Ford & Thompson 1986). Récemment, les ressources mimo-gestuelles et les postures corporelles qui accompagnent ces *si-indépendantes* ont également attiré une certaine attention. Dans une étude menée dans des salons de coiffure, Horlacher (2021) montre que les coiffeurs mobilisent le format indépendant [si X] pour effectuer des mises en garde d'un résultat potentiellement négatif (*ben si j'coupe court là ; si j'fais trop près de la racine*). Le statut indépendant de certaines de ces structures hypothétiques peut toutefois être remis en cause puisque la partie [(alors) Y] se trouve souvent réalisée gestuellement (p.ex. quand le coiffeur mime un résultat négatif sur sa propre tête). La prise en compte des ressources non verbales dans la description des structures en *si* complexifie ainsi la définition initiale et ne rend pas toujours facile un décompte précis des occurrences dans les données.

En résumé, les travaux menés sur les corpus de langue parlée, dont les interactions sociales, ont permis de mettre en évidence le fonctionnement indépendant des structures en *si* et cela, dans de nombreuses langues. Ces études ont également généré une littérature importante sur leur rendement pragmatique, tout comme elles ont avancé diverses interprétations concernant leur statut grammatical : erreur de performance, ellipse ou insubordination. Outre leur réalisation prosodique, qui mériterait d'être analysée en détail, ce sont en somme le fonctionnement

pragmatique de ces structures en français (mais voir Débaisieux *et al.* 2019 sur les structures relativement figées en *si tu savais...*), ainsi que les contextes sociaux et interactifs de leur occurrence qui restent largement à être explorés. Pour les requêtes en *si* que nous analysons ici, se posent entre autres les questions suivantes : Qu'est-ce qui motive une requête formulée sous la forme d'une *si-indépendante* ? Quel type de requête est réalisée par ce format ? Dans quelle constellation interactive (asymétries institutionnelles) et sous quelles conditions intersubjectives (p.ex. savoirs partagés) apparaissent-elles ?

3. Corpus et méthodologie

Dans cette contribution, nous nous intéressons aux formes de réalisation et au fonctionnement pragmatique des *si-indépendantes* telles qu'elles sont formulées par des citoyens lors d'appels à une centrale d'urgence. Le corpus à notre disposition est composé de 497 appels (durée moyenne : 90 secondes) recueillis au numéro d'urgence de la police dans une région francophone d'environ 180'000 habitants. La tâche de la centrale est de trier les demandes légitimes des demandes illégitimes, et ensuite d'acheminer de l'aide le plus rapidement possible quand elle est nécessaire. Les raisons d'appel les plus fréquentes concernent le tapage nocturne, le parking sauvage sur des places privées, les accidents de la route, les vols et les cambriolages, ou encore la violence domestique et les bagarres de rue. Nous disposons également de 205 *dispatchs* (durée moyenne : 46 secondes) résultant des appels mentionnés ci-dessus, que nous convoquerons de manière moins centrale. Ces deux types d'appels, enregistrés en 2015, forment une durée totale d'environ 15 heures et ont été recueillis et anonymisés par Virginie Fasel Lauzon à l'Université de Neuchâtel.

La relation des citoyens à la police peut être qualifiée d'asymétrique dans le sens où un citoyen, demandeur d'une prestation, fait appel à une autorité professionnelle compétente pour un problème qu'il n'arrive pas à résoudre seul. Elle se double d'une asymétrie inversée à l'égard des savoirs, dans le sens où le citoyen, mais pas la centrale, a accès à la situation et détient donc l'autorité épistémique quant aux faits l'incitant à demander de l'aide à la police. En outre, un système de droits et d'obligations (Sacks 1992) est lié aux rôles institutionnels des participants : en tant que citoyens, ceux qui composent le numéro d'urgence ont le droit de solliciter de l'aide, alors que la police a l'obligation (contractuelle, morale) de garantir la sécurité de ses concitoyens et de maintenir l'ordre public. Néanmoins, demander l'envoi d'une patrouille à un opérateur de centrale peut être une action délicate pour le citoyen, qui doit convaincre de la légitimité de son problème. C'est dans ce contexte interactif marqué par de fortes asymétries, à la fois épistémiques (Heritage 2012) et déontiques (Stevanovic 2018), que nous nous sommes intéressées aux formats des requêtes (Horlacher & Pekarek Doehler, à paraître), dont les *si-indépendantes*. Nous avons comptabilisé 38 occurrences de *si-indépendantes*, dont 3 seulement surviennent dans les *disptachs*.

4. Analyses

Dans le corpus des appels à la centrale, les requêtes des citoyens concernent typiquement l'envoi d'une patrouille et prennent souvent la forme d'une *si-indépendante*. L'analyse de notre collection va révéler le fonctionnement socio-interactif du format cible.

4.1 Une première illustration

Par le biais du format *si-indépendant*, le citoyen (CIT) s'oriente vers la nature délicate de sa requête, ainsi que vers son propre statut déontique, i.e., sa faible légitimité à demander à la centrale (CEN) de faire quelque chose :

(1) VFL-CET - 15-02-15 - 00h57m25s - tapage nocturne

```
01 CIT y a::: tapage nocturne, eu:::h début de bagarre,
02 euh des petits jeunes qui devraient être au lit eu:h
03 plutôt que de traîner là-devant, et: ça fait un bordel
04 incroyable sur la route,
05 (0.6)
06 CEN [mhm, ]
07 CIT [eu:::h] 'fin voilà.
08 si vous avez quelqu'un sous la main euh
09 (0.5)
10 CEN okay
```

À l'instar de l'extrait 1, nous observons que les appelants formulent fréquemment une requête en se servant d'une *si-indépendante* à travers laquelle ils énoncent une condition *sine qua non* pour que la requête puisse être honorée (l'envoi d'une patrouille n'étant possible que si une patrouille est disponible). Le fonctionnement du format cible en tant que requête est confirmé par le fait qu'il est régulièrement interprété par la centrale comme accomplissant une requête : dans l'extrait 1, la centrale confirme par son *okay* (l. 10) l'envoi d'une patrouille. Cette 'transparence' mutuelle pour les participants du format en tant qu'action de requête est un premier indice de sa nature potentiellement routinière.

De par son sémantisme, la structure hypothétique en *si* se prête idéalement à la tâche qu'elle accomplit (voir plus bas pour ses propriétés lexico-sémantiques récurrentes). Par la formulation hypothétique en *si*, le citoyen donne à l'opérateur de la centrale toute la latitude possible pour décliner la demande d'aide, ce que ce dernier peut potentiellement faire justement en niant l'existence des conditions *sine qua non* évoquées dans la structure *si-indépendante*. La récurrence de ce format dans les données étudiées reflète les rôles sociaux des participants, et le système des droits et des obligations qui sont liés à leurs statuts (Sacks 1992) : le format traduit que le citoyen ne s'estime pas autorisé à réclamer ouvertement une aide immédiate, mais en évoque néanmoins les conditions indispensables, laissant au policier le soin de juger si l'envoi d'une patrouille est justifié ou non.

4.2 La composition lexico-syntaxique du format, et son indépendance prosodique et pragmatique

Nos analyses révèlent un certain degré de consistance lexico-syntaxique du format, comme le témoignent les occurrences suivantes :

- si vous avez une patrouille qui est par là.
- si vous avez quelqu'un sous la main
- si y a quelqu'un qui peut juste venir faire un constat?
- si quelqu'un peut passer euh
- s'ils peuvent passer jeter un œil.

La nature hautement routinisée de ce type de requêtes s'observe entre autres dans l'occurrence répétée du pronom de deuxième personne ou du pronom indéfini (parfois *ils*) en position de sujets, les deux renvoyant à priori à l'institution 'police' qui est appelée à l'aide. Elle apparaît également dans l'occurrence régulière des verbes modaux *vouloir* et *pouvoir* et d'atténuateurs (il s'agirait pour la police de *juste* passer ou de jeter un *rapide coup d'œil*), qui eux aussi indexent l'orientation du citoyen vers la faible autorité déontique qu'il détient. La routinisation du format transparait enfin dans la thématisation fréquente de la disponibilité ou de la proximité de l'assistance demandée (*une patrouille/quelqu'un pas loin, disponible, tout près, sous la main...*), notamment à travers des expressions existentielles du type *vous avez* ou *il y a*, qui justement énoncent les conditions indispensables pour une possible intervention de la police, et donc pour la satisfaction de la requête qui, elle, reste toutefois implicite. Par ailleurs, si certaines de ces requêtes possèdent un contour intonatif suspensif ou une hésitation en fin d'énoncé, beaucoup se terminent clairement par une intonation finale (dans les exemples cités, un « . » en fin d'énoncé signale un contour final descendant, alors qu'un « ? » indique un contour final montant) ; ces *si-indépendantes* ne projettent donc pas de suite sur le plan prosodique : elles sont produites (et d'ailleurs traitées par l'interlocuteur) comme des unités complètes. Enfin, ces structures en *si* sont également indépendantes sur le plan pragmatique : elles accomplissent en elles-mêmes l'acte de requête, et cela de manière reconnaissable pour autrui (voir ex. 1 *supra*). Ces observations sur le statut illocutoire et le profil prosodique des *si-indépendantes* fournissent des arguments solides qui vont à l'encontre d'une vision de la structure insubordonnée comme impliquant l'ellipse d'une clause-matrice récupérable par inférence (voir section 2 *supra*).

Le corpus ne présente qu'un nombre limité de formats hypothétiques 'complets' de type [si X, (alors) Y] fonctionnant comme des requêtes. Pour ces rares cas (= 3 occurrences), la partie [(alors) Y] exprime toujours une sorte de conséquence positive (ex. *s'ils peuvent venir, tant mieux quoi* ; *si (vous seriez) assez rapide, ça m'arrange bien* ; *si vous voulez (0.2) venir saler, comme ça ça évitera encore deux morts quoi*). On observe aussi dans les données la présence de structures interrogatives indirectes (Riegel *et al.* 2009 [1994] : 836) qui fonctionnent comme requêtes, du type : *je sais pas si vous avez quelqu'un, une patrouille qui est en route ou comme ça* ; *je sais pas si y a un service de vétérinaire qui peut venir* ; *je voudrais savoir si vous avez une patrouille qui est pas très loin*. Si on tentait d'appliquer une théorie de l'ellipse face à nos *si-indépendantes*, on ne saurait donc pas très bien si le matériel 'effacé' concernerait la partie [(alors) Y] ou alors le verbe introducteur d'une interrogative indirecte. Sous cet aspect encore, il nous semble préférable d'interpréter les *si-indépendantes* comme des adaptations routinisées aux problèmes interactifs pratiques que les locuteurs rencontrent, comme celui de formuler des requêtes tout en les adaptant aux contraintes situationnelles et aux droits et obligations liés à leur statut social, au lieu de tenter de reconstruire à tout prix une structure canonique originelle.

4.3 Variation sur le format de base et analyse en contexte

Nous trouvons parfois des cas de *si-indépendantes* qui comportent des traits les rapprochant davantage de requêtes 'standards'. Ici, *vous pouvez envoyer* (cf. Riegel *et al.* 2009 [1994] : 989 sur les énoncés formulés avec "pouvoir", dont la valeur de requête "s'est figée en langue") combiné avec le *s'il vous plaît* final, énonce plus explicitement une requête que d'autres *si-indépendantes* du corpus (cf. ex 1) :

(2) VFL-CET-188 – 29-03-15 – 03h41m38s – personne agressive

01 CEN police NOM-VILLE?
02 (0.7)
03 CIT oui:, bonsoir, j` m`excuse d` vous déranger, c`est l` videur de
04 NOM-DE-BAR, (0.7) (euh) y a mon collègue qui a viré un gars,
05 pis maintenant il es:t il est très agressif euh à l`extérieur,
06 là: euh (il veut) pas s` calmer là.
07 (1.4)
08 CIT **alors si vous pouvez envoyer une patrouille s`il vous plaît.**
09 (1.0)
10 CEN NOM-D'UNE-PLACE hein, on va regarder ça.
11 (0.7)
12 CIT merci beaucoup.

Une analyse contextuelle de la structure en *si* nous permet d'éclairer les motivations fonctionnelles, notamment sur le plan intersubjectif, de ce format de requête. Dès le début de l'interaction, le citoyen traite son appel comme étant dérangeant (*j` m`excuse d` vous déranger*, l. 03), instillant un certain doute sur sa légitimité. En outre, la salutation et l'identification de l'appelant à travers son activité professionnelle (l. 03) sont inhabituelles dans la mesure où les ouvertures d'appels en cas de vraie urgence sont typiquement réduites, sans salutations et sans identification de l'appelant (Wahlen & Zimmerman 1987 : 181). Dans le cas de l'extrait (2), la présence de ces mêmes éléments traduit ainsi potentiellement que la raison d'appel ne concerne pas une urgence vitale. Face aux événements présentés par la suite (ll. 04-06), on comprend que l'appelant décrit une situation que lui et son collègue devraient être capables de maîtriser, de par leur rôle de videurs. L'opérateur de la centrale ne traite d'ailleurs pas le récit du citoyen comme une demande d'assistance ; du moins, il ne répond pas pendant la pause relativement longue de la ligne 07. Il n'est toutefois pas rare que les citoyens décrivent leur problème, sans formuler de requête, laissant à l'autorité policière le soin d'inférer ce qu'il faut faire (Drew & Walker 2010). Devant le silence de la centrale, le citoyen s'auto-sélectionne à nouveau en 08 et formule une structure en *si*, en y ajoutant la formule de politesse *s'il vous plaît* (l. 08) qui ne laisse planer aucun doute sur le statut de requête de son énoncé. C'est donc à la lumière de ce contexte interactif précis, qui matérialise une légitimité déontique réduite du citoyen quant à une intervention de la police, que la formulation hypothétique en *si* (de surcroît, fortement modalisée) prend tout son sens : elle permet à l'appelant de demander de l'aide tout en évoquant la possibilité que cette demande ne soit pas honorée. En somme, le format de la requête traduit une orientation du requérant vers la nature délicate de sa requête, qui, pourtant, finira dans le cas présent par être honorée par la centrale (*on va regarder ça*, l. 10).

4.4 Un éclairage additionnel sur la dimension fonctionnelle du format : sa production en deuxième ressort, suite à une première requête explicite

L'interprétation de notre format cible en tant que requête se trouve par ailleurs étayée par le fait que, parfois, le citoyen ou la citoyenne énonce une requête au moyen d'un format 'standard' (du type "est-ce que vous pouvez...?", "(vous) pouvez...?"), avant de faire suivre, peu après, une seconde requête à travers une *si-indépendante*. En énonçant une seconde requête par le biais de ce format, le citoyen ou la citoyenne revoit ses prétentions à la baisse et diminue la contrainte exercée sur le destinataire.

L'extrait 3 est une version étendue de l'extrait 1. Il donne à voir une requête explicite surlignée en gris (ll. 06-08), qui précède le format *si-indépendant* (l. 16) :

3) VFL-CET-83 - 15-02-15 - 00h57m25s - tapage nocturne

01 POL (la) police NOM-VILLE,
02 (0.3)
03 CIT NOM à l'appareil à NOM-VILLE
04 bonsoir monsieur=
05 POL =oui bonsoir
06 CIT .hh .hh écoutez **est-ce que vous pouvez: eu:h**
07 **envoyer une patrouille ou deu:x euh,**
08 **devant le temple à NOM-VILLE s'il vous plait,**
09 parce que y a:: tapage noctu:rne, eu:::h début de baga:rre,
10 euh des petits jeunes qui devraient être au lit eu:h
11 plutôt que de traîner là-devant,
12 et:: ça fait un bordel incroyable sur la route,
13 0.6)
14 POL [mhm,]
15 CIT [eu:::h] 'fin voilà.=
15 **=si vous avez quelqu'un sous la main eu:h**
16 (0.7)
17 POL okay

Après l'ouverture comprenant l'identification réciproque et les salutations (ll. 01-05), l'appelant introduit à la première occasion une demande d'assistance (l. 06) en utilisant un format en *est-ce que?*, accompagné du verbe modal "pouvoir", et en terminant son tour de parole par *s'il vous plaît*. Ce n'est qu'après cette requête qu'il fournit la raison de son appel (cf. le *parce que*, l. 09), en signalant des problèmes de bruit parmi un rassemblement de jeunes (ll. 09-12). L'explication du citoyen rencontre d'abord une absence d'accord de la centrale (l. 14), ce qui semble inciter le citoyen à ajouter une conclusion plus explicite de son propos après '*fin voilà*'. (l. 15) qu'il fait déboucher sur une "nouvelle" requête par le biais d'une *si-indépendante* (l. 16). On notera que, dans sa première requête, le citoyen avait demandé l'envoi d'*une patrouille ou deu:x*" (l. 07), alors que dans sa seconde requête (l. 16), il demande plus modestement si quelqu'un est disponible. Il semble donc qu'à travers ce format, les locuteurs réintroduisent une requête, tout en diminuant leur droit à prétendre à une aide immédiate. En formulant une condition pour que la requête puisse être accordée, le citoyen s'offre déjà une raison possible en cas de refus (il n'y a pas de patrouille disponible) – et l'offre également à son interlocuteur. Le citoyen présente ainsi la possibilité à la centrale de ne pas entrer en matière sur la requête, non pas en la refusant directement, mais en affirmant que les conditions *sine qua non* pour l'honorer ne sont pas données : il n'y a personne 'sous la main'.

Dans l'extrait 4, le citoyen est un chauffeur de bus, qui formule lui aussi deux requêtes au moment de réclamer de l'aide pour des incivilités :

4) VFL-CET-148 - 08-03-15 - 03h10m52s - incivilités dans un bus

01 POL police secours NOM-VILLE?
02 (0.4)
03 CIT oui: bonsoir, BIP (routier)

04 (0.3)
05 POL bonsoir
06 (0.4)
07 CIT eu:h est-ce que vous pourriez: nous envoyer une patrouille,
08 on a un bus qui est arrêté au::: un bus noctambus au BIP
09 (0.6)
10 CIT y a des: incivilités, des gens qui fument, (.)
11 **donc euh:: (0.7) si vous av(i)ez une patrouille.**
12 (0.2)
13 POL alors noctambus:, arrêt du BIP
14 (0.3)
15 CIT ouais

Le pattern structurel est parallèle à ce que nous avons observé dans l'extrait précédent. Après l'ouverture (ll. 01-06), le citoyen demande une patrouille de police (l. 07) par le biais d'un format de requête courant : *est-ce que*, suivi du verbe modal "pouvoir". La raison de la requête viendra dans un second temps, dès la ligne 08. Après un manque de réaction de la centrale (l. 06), le citoyen ajoute une explication supplémentaire et réintroduit sa demande via une *si-indépendante*. Ainsi, il maintient la pertinence de sa requête en en produisant une seconde version, mais à travers laquelle il se positionne comme moins légitime, tout en projetant néanmoins une réponse du policier. A la ligne 13, le policier prend d'ailleurs les coordonnées géographiques de l'incident, ce qui projette un déplacement.

5. Conclusions et perspectives

A travers notre corpus d'appels d'urgence, nous avons montré que le citoyen se sentait investi du droit d'utiliser les services de la société pour régler ses problèmes mais en même temps, les appels examinés ne concernent pas des urgences vitales. Le format décrit semble traduire cette dualité, et notamment cette incertitude, de la part du citoyen, quant à la légitimité de son cas. Les requêtes examinées, énoncées sur le mode hypothétique, et centrées sur l'existence (ou non) des conditions permettant à la centrale d'organiser une intervention, s'expliquent donc par rapport à deux facteurs clés : a) la légitimité du locuteur (*entitlement*, Sacks 1984), c'est-à-dire son droit à formuler une requête face à son interlocuteur et b) les contingences (Curl & Drew 2008), c'est-à-dire les facteurs contextuels qui font que l'autorité policière est susceptible d'accorder une requête ou non. Alors que nos données ne soutiennent pas l'interprétation proposée par Evans (2007) de l'insubordination comme impliquant l'ellipse d'une clause-matrice, elles confirment l'association forte entre insubordination et gestion des relations interpersonnelles (Evans 2007, Beijering *et al.* 2019, Watanabe & Evans 2016). Nous espérons ainsi avoir pu éclaircir comment l'analyse d'un corpus d'interactions peut contribuer à élucider certaines motivations pragmatiques liées au choix que les locuteurs font en faveur d'un format linguistique spécifique pour accomplir un type d'action aussi omniprésent que la requête. La variation qui existe entre ces formats de requête, à laquelle nous n'avons pu toucher que brièvement (ex. 3 et 4), se révèle précisément "taillée" non seulement à la situation d'interaction, mais aussi au déroulement séquentiel même de cette interaction, le format *si-indépendant* étant parfois utilisé comme deuxième ressource lorsqu'un format de requête plus direct ne reçoit pas de réponse.

Dans l'optique du fonctionnement intersubjectif ainsi constaté, la rareté des *si-indépendantes* dans les *dispatchs* (c'est-à-dire dans les appels de la centrale aux patrouilles) par opposition aux appels impliquant les citoyens s'explique aisément : dans les *dispatchs*, et donc entre collègues, les enjeux de rapports hiérarchiques ne se posent pas à priori. Il est d'autant plus intéressant d'observer que les opérateurs de centrale énoncent une requête au moyen d'une *si-indépendante*

uniquement dans un contexte bien spécifique, à savoir lorsque le cas du citoyen ne tombe pas dans les catégories policières prototypiques, mettant en doute l'urgence de la situation et, par conséquent, l'obligation de la patrouille de se déplacer.

Notre interprétation des *si-indépendantes* rejoint des arguments récents sur la motivation interactive de routines langagières : la confrontation récurrente des locuteurs à des problèmes interactifs concrets, tels que l'organisation des changements de tours de parole (Detges & Waltereit 2011), la négociation de l'intercompréhension (Hopper & Traugott 2003 : 71), ou encore la gestion des droits déontiques mutuels, est susceptible de favoriser la routinisation (cf. Haiman 1994) de structures linguistiques à travers l'usage social du langage (Couper-Kuhlen 2011, Pekarek Doehler 2001). Nous proposons en somme que les *si-indépendantes* analysées relèvent de routines foncièrement interactionnelles – représentant un épiphénomène de l'usage (Hopper 2011, (Bybee & Hopper 2001) – et qui sont liées à la construction et au maintien de l'intersubjectivité dans l'interaction en tant que site *princeps* de l'utilisation du langage. C'est en ce sens, avec la particularité d'une approche interactionnelle du langage en usage, que nous espérons – avec cette contribution – rejoindre certains aspects du projet intellectuel de Claire Blanche-Benveniste : " Claire Blanche-Benveniste a, dès les années 1970, la certitude que les usages du français parlé dépendent de la situation dans laquelle il est produit " (Roubaud 2013 : 15).

Références

- Avanzi, M., Béguelin, M-J & Diémoz, F. (2012-2019). *Présentation du corpus OFROM – corpus oral de français de Suisse romande*. Neuchâtel : Université de Neuchâtel. <http://www.unine.ch/ofrom>.
- Blanche-Benveniste, C. (1983). « Examen de la notion de subordination ». *Recherches sur le français parlé*, 4, 71-115.
- Beijering, K. , Kaltenböck, G. & Sansiñena, M. S. (2019). *Insubordination. Theoretical en empirical issues*. Berlin / Boston : de Gruyter.
- Buscha, A. (1976). « Isolierte Nebensätze im dialogischen Text ». *Deutsch als Fremdsprache*, 13 (5), 274–279.
- Bybee, J. L. & Hopper, P. (éds.) (2001). *Frequency and the emergence of linguistic structure* (= Typological Studies in Language, 45). Amsterdam : Benjamins.
- Corminboeuf, G. (2008). « Entre détachement et intégration : la topographie des constructions en *si* et le marquage de la structure informationnelle ». *Congrès Mondial de linguistique française (CMLF-08)*. Actes du colloque, 2427-2439.
- Corminboeuf, G. & Jahn, T. (2020). « Taxinomie des constructions en *si* dans un corpus de français oral. L'exemple d'*OFROM* ». In G. Corminboeuf, J. Rotenbühler & M. Sauzet (éds.) : *Français parlé et français 'tout court'*, 4. DOI: <https://doi.org/10.25364/19.2020.4.10>
- Couper-Kuhlen, E. (2011). « Grammaticalization and conversation ». In H. Narrog & B. Heine (éds.) : *The Oxford Handbook of Grammaticalization*. Oxford : Oxford University Press, 424-437.
- Couper-Kuhlen, E. & Selting, M. (2018). *Interactional Linguistics. Studying language in social interaction*. Cambridge : Cambridge University Press.
- Curl, T. S. & Drew, P. (2008). « Contingency and action : A comparison of two forms of requesting ». *Research on Language and Social Interaction*, 41 (2), 129-153.
- Debaisieux, J.-M. (1994). *Le fonctionnement de « parce que » en français parlé contemporain : description linguistique et implications didactiques*. Thèse de doctorat, Université de Nancy.
- Debaisieux, J.-M., Deulofeu, J. & Martin, P. (2008). « Pour une syntaxe sans ellipse ». In J.-C. Pitavy & M. Bigot (éds.) : *Ellipse et effacement*. PU de St-Etienne, 224-246.

- Debaisieux, J.-M., Martin, P. & Deulofeu, J. (2019). « Apparent insubordination as discourse patterns in French ». In K. Beijering, G. Kaltenböck & M. S. Sansiñena (éds.) : *Insubordination. Theoretical and empirical issues*. Berlin / Boston : de Gruyter, 349-383.
- De Stefani, E. (2021). « If-clauses, their grammatical consequents, and their embodied consequence : Organizing joint attention in guided tours ». In L. Keevallik, S., Pekarek Doehler & X. Li (éds). *Frontiers in Communication*. Special Issue on the Grammar-Body Interface.
- Deulofeu, J. (1988). « La syntaxe de *que* en français parlé et le problème de la subordination ». *Recherches sur le français parlé*, 8, 79-104.
- Detges, U. & Waltereit, R. (2011). « Turn-taking as a trigger for language change ». In S. Dessi Schmid, U. Detges, P. Gevaudan, W. Mihatsch & R. Waltereit (éds.) : *Rahmen des Sprechens. Beiträge zu Valenztheorie, Varietätenlinguistik, Kognitiven und Historischen Semantik*. Tübingen : Narr, 75-190.
- Drew, P. & Walker, T. (2010). « Citizens' emergency calls : Requesting assistance in calls to the police. In M. Coulthard & A. Johnson (éds.) : *The Routledge handbook of forensic linguistics*, 95-110.
- Dwyer, A. (2016). « Ordinary insubordination as transient discourse ». In N. Evans & H. Watanabe (éds) : *Insubordination*. Amsterdam : Benjamins, 1-38.
- Evans, N. (2007). « Insubordination and its uses ». In I. Nikolaeva (éd.) : *Finiteness : theoretical and empirical foundations*. Oxford : Oxford University Press, 366-431.
- Evans, N. (2009). « Insubordination and the grammaticisation of interactive presuppositions ». Manuscrit d'une conférence donnée au Musée d'Ethnographie, Osaka, Japan, mars 2009.
- Evans, N. & Watanabe, H. (2016). « The dynamics of insubordination : An overview ». In N. Evans & H. Watanabe (éds) : *Insubordination*. Amsterdam : Benjamins, 1-38.
- Ford, C. E. & Thompson, S. A. (1986). « Conditionals in discourse : A text-based study from English ». In E. C. Traugott, A. Meulen, J. S. Reilly, & C. A. Ferguson (éds.) : *On conditionals*. Cambridge : Cambridge University Press, 353-372.
- Günthner, S. (2020). « From complex *wenn*-constructions to insubordinate ('stand-alone') conditionals in everyday spoken German ». In Y. Maschler, S. Pekarek Doehler, J. Lindström & L. Keevallik (éds.) : *Emergent Syntax for Conversation : Clausal patterns and the organization of action* [= *Studies in Language and Social Interaction*, 32]. Amsterdam : Benjamins, 185-220.
- Haiman, J. (1994). « Ritualization and the development of language ». In W. Pagliuca (éd.) : *Perspectives on Grammaticalization*. Amsterdam : Benjamins, 3-28.
- Heritage, J. (2012). « Epistemics in action: Action formation and territories of knowledge ». *Research on Language and Social Interaction*, 45, 1-29.
- Horlacher, A.-S. & Pekarek Doehler, S. (à paraître). *Faire agir autrui : la requête dans les interactions en milieu institutionnel*. Numéro thématique de *Langage et Société*, 179.
- Horlacher, A.-S. (2021). « Compound and independent if-clauses at the hairdresser's : negotiating the procedures of the treatment ». Présentation orale, *17th International Pragmatics Conference*, 27 juin - 2 juillet 2021, Winterthur, Suisse.
- Laury, R. (2012). « Syntactically Non-integrated Finnish *jos* 'If-Conditional Clauses as Directives ». *Discourse Processes*, 49, 213-242.
- Laury, R., Lindholm, C. & Lindström, J. (2013). « Syntactically non-integrated conditional clauses in spoken Finnish and Swedish ». In E. Havu & I. Hyvärinen (éds.) : *Comparing and contrasting syntactic structures : From dependency to quasi-subordination*. Helsinki : Société Néophilologique de Helsinki, vol. LXXXVI, 231-270.
- Lombardi Vallauri, E. (2004). « Grammaticalization of Syntactic Incompleteness : Free Conditionals in Italian and other Languages ». *Journal of Linguistics*, 17, 189-215.
- Lindström, J., Laury, R. & Lindholm, C. (2019). « Insubordination and the contextually sensitive emergence of if-requests in Swedish and Finnish institutional talk-in-interaction ». In K. Beijering, G. Kaltenböck, M. Sansiñena (éds.) : *Insubordination*. Berlin : De Gruyter.

- Maschler, Y., Pekarek Doehler, S., Lindström, J., Keevallik, L. (éds.) (2020) : *Emergent syntax for conversation : Clausal patterns and the organization of actions*. Amsterdam : Benjamins.
- Ochs, E., Schegloff, E. A. & Thompson, S. A. (Eds.) (1996). *Interaction and Grammar*. Cambridge : Cambridge University Press.
- Patard, A. (2014). « Réflexions sur l'origine de l'insubordination. Le cas de trois insubordonnées hypothétiques du français ». *Langages*, 196, 109-130.
- Pekarek Doehler, S., De Stefani, E. & Horlacher, A.-S. (2015). *Time and emergence in grammar : Dislocation, topicalization and hanging topic in French talk-in-interaction*. Amsterdam : Benjamins.
- Pekarek Doehler, Simona (2021). « How grammar grows out of social interaction : from multi-unit to single unit question ». *Open Linguistics*.
- Riegel, M. et al. (2009 [1994]). *Grammaire méthodique du français*. 4^{ème} Édition. Paris : Presses Universitaires de France.
- Roubaud, M.-N. (2013). « Préface ». In *Langue et enseignement. Une sélection de 22 manuscrits de Claire-Blanche Benveniste* (= TRANEL, 58), 7-18.
- Sacks, H. (1984). « On doing 'being ordinary' ». In J. M. Atkinson & J. Heritage (éds.) : *Structures of social action : Studies in conversation analysis*. Cambridge : Cambridge University Press, 413-429. Sacks, H. (1992). *Lectures on conversation* (2 vol.). In G. Jefferson (Ed.). Oxford: Blackwell.
- Schwenter, S. A. (2006). « Meaning and interaction in Spanish independent *si*-clauses ». *Language Sciences*, 58, 22-34.
- Stevanovic, M. (2011). « Participants deontic rights and action formation : The case of declarative requests for action ». *InLiSt, Interaction and Linguistic Structure*, 52, 1-37. <http://www.inlist.uni-bayreuth.de/>
- Stirling, L. (1999). « Isolated if-clauses in Australian English ». In P. Collins & D. Lee : *The Clause in English*. Amsterdam : Benjamins, 273-294.
- Whalen, M. R. & Zimmerman, D. H. (1987). « Sequential and institutional contexts in calls for help ». *Social Psychology Quarterly*, 172-185.
- Watanabe, H. & Evans, N. (éds) (2016). *Insubordination*. Amsterdam : Benjamins.

Corpus oraux et enseignement du français : présentation du projet « Enseignement » de la base OFROM

Laure Anne Johnsen, Institut de langue et civilisation françaises Université de Neuchâtel

Dans cette communication, nous souhaitons présenter le volet « Enseignement » du corpus OFROM (corpus oral de français de Suisse romande, Avanzi *et al.* 2012-2021 www.unine.ch/ofrom), projet en cours de réalisation qui a pour but de créer et de mettre à la disposition du public enseignant et apprenant des ressources pour l'enseignement/apprentissage de la langue française, en particulier dans ses dimensions orale et régionale.

La question sous-jacente à l'origine de ce projet est la suivante : sur quel modèle de la langue devrait se baser l'enseignement de l'oral ? Plus spécifiquement, quelles sont les approches les plus adaptées au développement des compétences orales attendues des apprenants ? Ces questions dépendent évidemment de nombreux critères, comme du profil des apprenants, du contexte et des objectifs spécifiques d'enseignement/apprentissage, du niveau de langue, etc.

À travers ce projet, nous ne souhaitons ni ne pouvons répondre exhaustivement à ces questions fondamentales, mais nous voulons proposer des pistes didactiques pour la prise en compte de plusieurs aspects de la langue souvent laissés pour compte dans l'enseignement. En effet, la langue quotidienne parlée en Suisse romande n'est que sporadiquement prise en compte dans l'enseignement du français (de Pietro 2008), à part dans le cadre de cours de sociolinguistique au niveau universitaire. Dans une perspective d'enseignement du FLM ou FLE, l'enjeu n'est pas tellement d'en faire un objet d'enseignement en soi (*idem*), mais plutôt un objet de réflexion et de sensibilisation des apprenants à la variation linguistique. C'est cet objectif que se donne le présent projet, en fournissant des outils didactiques aux enseignants et aux apprenants afin d'aborder la langue en usage en Suisse romande tout en prévenant des approches tendant à caricaturer, dévaloriser ou tout bonnement occulter tout un pan de ses caractéristiques fondamentales.

Depuis quelques années, on peut se réjouir du développement de projets didactiques fondés sur l'exploitation de corpus oraux en vue de l'enseignement/apprentissage du FLM ou FLE qui partagent l'objectif général de confronter les apprenants à une image plus réaliste des pratiques langagières. Parmi ceux-ci, on peut citer le projet PFC-EF <https://www.projet-pfc.net/le-projet-pfc-ef/>, la base CLAPI-FLE <http://clapi.icar.cnrs.fr/FLE>, la ressource FLORALE <https://florale.unil.ch/>, le site FLEURON <https://apps.atilf.fr/fleuron> et le projet MLE-MPF (<http://www.mle-mpf.bbk.ac.uk/Resource.html>). Les objectifs spécifiques, le public cible et les types de ressources peuvent néanmoins sensiblement différer d'un projet à l'autre. Reste que les moyens d'enseignement pour le FLM ou les manuels de FLE ne renvoient pas spécifiquement, à notre connaissance, à ces ressources. Ceci peut paraître paradoxal, du moins pour le FLE, sachant que l'élaboration de corpus comme le *Français fondamental* (Gougenheim *et al.* 1956/1964) et la linguistique de corpus sont nées d'un besoin lié à l'enseignement des langues étrangères et qu'on reconnaît l'apport des corpus dans ce domaine (Boulton & Tyne 2014). Parmi les explications, outre la distance notoire qui sépare le milieu de la recherche et celui de l'enseignement (Gagnon & Benzitoun 2020), on peut invoquer les difficultés de maîtrise de ces outils (moteurs de recherche, terminologie à maîtriser, etc.) mais aussi le risque de leur obsolescence, ainsi que l'insuffisance de la formation des enseignants à la description du français parlé et encore les représentations bien ancrées, dans le milieu de l'enseignement comme ailleurs, associant langue écrite à langue correcte et langue orale à langue familière et fautive (Blanche-Benveniste 1997).

Le présent projet a précisément pour ambition de permettre et favoriser ce lien entre recherche et enseignement/apprentissage en proposant à travers une interface facile d'accès des explications claires accompagnées d'illustrations. L'approche s'efforce d'adopter le point de vue de l'enseignant ou de l'apprenant, d'où le choix d'un accès aux informations via un outil de *foire aux questions* (FAQ) évolutif. Pour chacune de ces questions, une réponse brève est ou sera rédigée par un spécialiste dans un style adapté au public-cible.

De par la nature même du corpus OFROM, deux axes sont représentés, chacun pourvu d'une FAQ comme point d'entrée.

L'axe régional vise à présenter les spécificités du français parlé en Suisse romande. Cet axe fournira dans le même temps une documentation inédite pour un public élargi intéressé par ces questions. Or, parmi la foule des ouvrages ou supports sur le français romand destiné au grand public, peu donnent une image de la vitalité du français de Suisse romande sous tous ses aspects, la description étant souvent réduite à la dimension lexicale, parfois axée sur un folklore désuet (de Pietro 2008). Actuellement, outre le Dictionnaire Suisse Romand (Thibault 1993) numérisé dans la BDLP¹ et le Glossaire des patois de la Suisse romande², il n'existe pas à notre connaissance de ressource de vulgarisation scientifique pérenne qui se concentre spécifiquement sur le français régional de Suisse romande et en illustre les caractéristiques à différents niveaux linguistiques. Ce projet se veut donc en partie combler ce manque et servir de source de documentation théorique et sonore reposant sur des bases scientifiques pour un public de non spécialistes.

La FAQ *Le français en Suisse romande* se présente comme suit :

The screenshot shows the website interface for the 'Corpus Oral de Français de Suisse Romande'. At the top, there are logos for 'unine' (Université de Neuchâtel), 'UNIFR' (Université de Fribourg), and 'ILCF' (Université de Neuchâtel). Below the logos is a navigation menu with 'ENSEIGNEMENT' highlighted. The main content area is titled 'Le français en Suisse romande'. A paragraph explains the purpose of the rubric. On the left, there is a list of 14 FAQ questions. On the right, there is a text block starting with 'Le français fait partie des quatre langues nationales de la Suisse...' and a graphic titled 'NOUS PARLONS « SUISSE »' showing that 24.3% of Swiss residents do not have a primary language, and that Switzerland has 4 national languages.

1. Qu'est-ce que le français de Suisse romande?

2. D'où vient le français parlé en Suisse romande?

3. Langue, patois, dialecte: quelles différences?

4. Que signifie le mot « dialecte » en Suisse?

5. Qu'est-ce qu'un helvétisme?

6. Quelles sont les particularités linguistiques romandes?

6.1. Lexique

6.2. Prononciation

6.3. morphosyntaxe

7. Y a-t-il un accent romand?

8. Les Suisses romands parlent-ils lentement ?

9. Qu'est-ce qu'un germanisme?

10. Qu'est-ce qu'un statalisme?

11. Qu'est-ce que le « français fédéral »?

12. Comment les Romands jugent leur manière de parler ?

13. Comment les enseignants jugent le français de Suisse ?

14. Qu'est-ce que le bolze?

Le français fait partie des quatre langues nationales de la Suisse. D'après l'Office fédéral de la Statistique (OFS), le français est déclaré langue principale par 23% de la population, vivant majoritairement dans l'ouest du pays, à savoir la Suisse francophone appelée Suisse romande. L'allemand constitue la langue majoritaire du pays, il est cité comme langue principale par 62% de la population. Ce chiffre éclipe en fait une réalité plus complexe : c'est le dialecte alémanique, se déclinant lui-même en de nombreuses variétés, qui est essentiellement parlé au quotidien, tandis que l'allemand standard est employé principalement à l'écrit surveillé et oralement dans certains contextes institutionnels ou publics. Ensuite, l'italien est parlé principalement par 8% des habitants, vivant au sud-est du pays. Ici également, il faut mentionner la pratique du dialecte tessinois dans certains contextes du quotidien, mais dans une mesure bien moindre que la situation du dialecte alémanique. Enfin le romanche, qui rassemble lui aussi différents «idiomes» d'origine rhéto-romane des Grisons, est déclaré langue principale pour 0,5% des résidents de la Suisse. D'autres langues, dites étrangères, sont parlées en Suisse, la plus citée comme langue principale étant l'anglais (6%), devant le portugais (3,5%) (les personnes pouvant indiquer plusieurs langues principales, le total dépasse 100%).

NOUS PARLONS « SUISSE »

24,3% des habitants suisses n'ont pas comme langue principale l'une des 4 langues nationales.

4 LANGUES En Suisse, il existe 4 langues nationales.

22,9% de français

Dans la colonne de gauche figurent des questions générales (e.g. questions 1, 2) ou terminologiques (e.g. question 4, 5, etc.) susceptibles de se poser des enseignants, apprenants ou personnes intéressées,

1 <https://www.bdlp.org/>

2 <https://www.unine.ch/gpsr>

recouvrant différents aspects du français parlé en Suisse romande. Dans la colonne de droite figure la réponse proposée par le spécialiste (en moyenne 700 mots), sous forme de texte explicatif accessible au public auquel s'ajoutent quelques références pour aller plus loin (par exemple pour préparer un exposé, une exposition, un article de presse) et sous forme d'illustrations sonores tirées de la base OFROM, d'éventuelles images comme des cartes géographiques (Avanzi, blog <https://francaisdenosregions.com/>) :

L'interface permet ainsi de rassembler pour chaque sujet spécifique et dans un environnement convivial différents médiums explicatifs, à savoir textuels, sonores et visuels de source fiable, recueillis en amont sur des bases empiriques. En particulier pour les apprenants (FLM ou FLE), cet axe vise à contextualiser les notions de variété, dialecte, langue et à questionner la légitimité de leur existence et à réfléchir à la participation des aspects linguistiques régionaux dans la construction identitaire.

Le **second axe** aborde le français dans sa dimension orale en contexte et vise à documenter des phénomènes langagiers aux niveaux diamésique et diaphasique. L'objectif est de proposer une approche nuancée sur l'opposition entre oral et écrit (Béguelin 1996, Blanche-Benveniste 1997, Koch & Oesterreicher 2007), de mettre en évidence différentes situations de l'oral selon des indices explicites et de susciter une réflexion sur les critères d'adéquation des pratiques langagières dans un contexte donné.

La FAQ le français parlé en contexte se présente sous la même forme, en voici ici uniquement la partie « questions » :

1. Oral vs écrit: quels points communs, quelles différences ?
2. Quels rapports entretiennent les lettres avec les sons?
3. L'écrit comme modèle de l'oral, une bonne idée?
4. Y a-t-il des phrases à l'oral?
5. Quels sont les différents genres de l'oral?
6. Que sont les registres ou les niveaux de langue?
7. Quelles sont les caractéristiques de l'oral en conversation?
 - 7.1. Au niveau interactionnel
 - 7.2. Au niveau morphologique
 - 7.3. Au niveau lexical
 - 7.4. Au niveau syntaxique
 - 7.5. Au niveau phonétique
8. Qu'est-ce qu'un argot?
9. Comment caractériser les parlers «jeunes»?
10. Est-ce que les jeunes parlent moins bien qu'avant?

Ici également, aussi bien des questionnements « naïfs » que d'ordre terminologique sont proposés. La réponse est expliquée dans une forme de technicité adaptée à un public de non spécialistes, en particulier enseignants et apprenants, et illustrée par des exemples sonores tirés de la base OFROM. Afin de répondre

aux objectifs susmentionnés, un élargissement du corpus OFROM à de nouveaux genres est en cours. En plus des trois types d'interaction actuels (conversation, entretien, conférence), la base sera dotée d'enregistrements d'interviews médiatiques, de mémos vocaux, d'enseignement, etc. Sur la base de cette diversification en types d'interaction, cet axe souhaite illustrer la variété des pratiques langagières et ce, à différents niveaux d'analyse et notamment proposer aux enseignants des pistes et outils adaptés pour aborder la question complexe de la production orale, non limitée au registre dit « familier », sans passer par le modèle sous-jacent de l'écrit ni mettre dos à dos les deux médiums.

Nous parcourons quelques-unes des stratégies de didactisation et de vulgarisation pour ouvrir la discussion et présenterons les pistes de développement futur du projet.

En somme, le présent projet souhaite prolonger les acquis de la recherche sur le français parlé, largement initié par Claire Blanche-Benveniste et son équipe, et exploiter les avantages des corpus oraux pour en faire bénéficier les champs d'application comme l'enseignement/apprentissage du français.

Références

- Avanzi M., Béguelin M.-J., Corminboeuf G., Diémoz F. & Johnsen L. A. (2012-2020). Corpus OFROM – *Corpus oral de français de Suisse romande*. Université de Neuchâtel, www.unine.ch/ofrom
- Béguelin, M.-J. (1996). « Le rapport écrit-oral: tendances dissimilatrices, tendances assimilatrices ». *Cahier de linguistique française*, 20, 229-253.
- Blanche-Benveniste, C. (1997), *Approches de la langue parlée en français*, Paris, Ophrys, France.
- Boulton, A. & Tyne, H. (2014). « Des Documents Authentiques aux Corpus : introduction ». In A. Boulton & H. Tyne, *Des Documents Authentiques aux Corpus : Démarches pour l'Apprentissage des Langues*, Paris : Didier.
- Gagnon, R., & Benzitoun, Ch. (2020). « Le français parlé comme objet d'enseignement ? Regards croisés d'un linguistique et d'une didacticienne ». *Formation et Pratiques d'Enseignement en Questions - Revue des HEP de Suisse romande et du Tessin*, 26, 37-51
- Gougenheim, G., Michea, M., Rivenc, P., Sauvageot, A. (1956/1964), *L'élaboration du Français fondamental (1er degré). Étude sur l'élaboration d'un vocabulaire et d'une grammaire de base*, Paris : Didier.
- de Pietro, J.-F. (2008). « L'école face aux variétés du français: réflexions à partir de la situation en Suisse francophone ». In *La langue française dans sa diversité*. Québec : Secrétariat à la politique linguistique, p. 181-208.
- Koch, P. & Oesterreicher, W. (2007). « Schriftlichkeit und kommunikative Distanz ». *Zeitschrift für germanistische Linguistik*, 35, 346-375.
- Thibault, A. (1997). *Dictionnaire suisse romand*. Carouge : Éditions Zoé.

Tu comprends, vous comprenez, je comprends **Quelles fonctions ont-ils dans le discours?**

Kaoru Kunisue, Université des Langues Étrangères de Tokyo

1. Introduction

Cette étude porte sur les fonctions pragmatiques et discursives de *tu comprends, vous comprenez et je comprends* employés en « construction parenthétiques » (terme utilisé par Bolly 2009). Certaines caractéristiques de ces expressions ont déjà été étudiées par d'autres auteurs. Nous voulons apporter un nouvel éclairage en les réexaminant en tant que « marqueurs pragmatiques » - des emplois ayant beaucoup attiré l'attention des linguistes particulièrement ces dernières années.

Dans ce qui suit, nous décrivons ces trois constructions parenthétiques en nous appuyant sur des corpus de français parlé. Après un bref rappel de ce qui a été acquis dans les recherches antérieures dans la section 2, nous présenterons, en 3, les corpus et la méthodologie que nous utiliserons. Nous montrerons ensuite notre résultat d'analyse en 4, et enfin, en 5, nous en tirerons une conclusion sur les particularités des constructions au niveau discursif et pragmatique.

2. État de l'art

2.1. Constructions parenthétiques

Il existe diverses appellations pour la construction parenthétique telle qu'*incise* (Blanche-Benveniste 1989), *reduced parenthetical clauses* (Schneider 2007), *les marqueurs discursifs propositionnels* (Andersen 2007). Ces termes désignent une unité linguistique construite par la suite de « sujet + verbe », qui est grammaticalement indépendante de l'énoncé auquel elle est rattachée, que nous appelons « énoncé hôte ». Suivant Kaltenböck (2009) et Bolly (2009), nous définissons l'énoncé hôte comme un énoncé auquel la construction parenthétique est sémantiquement et pragmatiquement liée. En ce qui concerne la relation qui s'établit entre l'énoncé hôte et la construction parenthétique, Bolly (2009) note ceci :

« Sur le plan sémantico-pragmatique, on considère généralement que l'argument absent est récupérable dans le contexte immédiat, à savoir dans l'unité « hôte »: les constructions parenthétiques réduites seraient donc reliées à leur hôte par une relation qui n'est pas d'ordre (morpho-)syntaxique, mais bien sémantico-pragmatique. » (Bolly 2009 : 107)

Par ailleurs, d'après l'Approche Pronominale de Blanche-Benveniste (1984, 1990, 2010), un « énoncé » constitue une unité minimale de discours, dont on peut distinguer deux types : le « noyau », qui constitue une partie centrale et l'« affixe » (Blanche-Benveniste 1990 : 126) ou le « satellite » (Debaisieux 2013: 74), qui forme un élément périphérique dépendant du « noyau ». Nous pensons que l'« énoncé hôte » correspond au « noyau ».

La construction parenthétique constitue un élément périphérique par rapport à l'énoncé hôte, en y ajoutant différentes modalités (inter)subjectives. Elle est détachable comme le montre l'exemple de Blanche-Benveniste cité ci-dessous.

- (1) **je crois bien**⁰ que c'était signalé dans le journal.
c'était signalé dans le journal, **je crois bien**.
c'était, **je crois bien**, signalé dans le journal.

(Blanche-Benveniste 1989: 60)

Dans notre étude, nous adaptons le terme de Bolly (2009), « construction parenthétique » pour désigner cette unité. Nous préférons utiliser le terme « construction », qui est défini, par Goldberg (2019: 2), comme un appariement entre les formes linguistiques et ses significations fonctionnelles.

⁰ C'est l'auteur de l'article qui met en gras. Il en va de même pour les exemples de corpus qui suivent.

Les particularités de la construction parenthétique sont: (i) la liberté de position, elle peut être antéposée, postposée ou insérée par rapport à l'énoncé hôte (voir l'exemple 1) ; (ii) l'invariabilité au niveau morphologique; (iii) la possibilité de la suppression sans changer de conditions de vérité de l'énoncé; (iv) l'indicateur de la modalité du locuteur entraînant un sens subjectif ou intersubjectif.

Notons que ces caractéristiques s'observent également pour des marqueurs pragmatiques.⁰ En effet, beaucoup de chercheurs ont travaillé sur les fonctions pragmatiques de constructions parenthétiques de différents verbes tels que *voir*, *savoir*, *dire* en les considérant comme des marqueurs pragmatiques. Dans cet article, nous nous intéresserons particulièrement à la construction parenthétique du verbe *comprendre*, relativement peu étudiée dans ces recherches antérieures.

2.2. Verbe *comprendre* en construction parenthétique

Voici quelques exemples de la construction parenthétique du verbe *comprendre* attestés dans le Corpus d'études du français contemporain du projet ORFÉO:

- (2) L2: de en f- eh oui, mais c~ **vous comprenez** quand ils étaient quelque part ça ça accrochait pas le sol ça si c' était mouillé
(CRFP, PRI-SAI-1)
- (3) il y aura pas de chasse il y aura pas de chasseurs, mais ils seront obligés là d' aller faire nettoyer les chemins eux faudra qu' ils coupent les arbres couper les branches des orties les ronces couper l'herbe euh tailler les branches et ainsi de suite quoi **tu comprends**
(TCOF, Employeducnrs)
- (4) et puis il y a il y a des des paroles qui partent il y a des gestes qui partent et ça va pas ç **je comprends** hein le le le fait de de d' être agressé dans le train des trucs comme ça, mais il faut quand même un tout petit peu revoir un peu sa copie par rapport au service qu' on qu' on qu' on fait hein
(CFPB, CFPB-1000-5)

Dostie (2004) remarque que le verbe *comprendre* fonctionne comme un marqueur pragmatique sous la forme de *tu comprends / vous comprenez, comprends-tu / comprenez-vous, comprenez et je comprends*. D'après cette étude, *je comprends* est considéré comme un *marqueur de la réalisation d'un acte illocutoire* et *tu comprends* comme un *marqueur d'appel à l'écoute*.

Andersen (2007: 23) affirme qu' « il n'est pas possible dans le cas de *comprendre* de parler de deux significations selon les emplois, mais bien d'un certain affaiblissement sémantique quand il s'agit du MDP ». On pourrait distinguer cette différence en ayant recours au test formel de zéro-pronominalisation, mais l'efficacité du test ne semble pas être évidente particulièrement pour la construction parenthétique du verbe *comprendre*, car il est parfois possible d'insérer le pronom déictique *ça* à la place de zéro-pronominalisation comme dans l'exemple ci-dessous :

- (5) la vache c' est une c' est une bête à lait donc nous on ne travaille pas cette viande-là **vous comprenez** → *vous comprenez ça*
(CRFP, PRO-QUI-1)

⁰ Dostie et Pusch (2007) et Fédriani et Sansó (2017) présentent la liste des caractéristiques des marqueurs pragmatiques.

Le statut de marqueur pragmatique de *comprendre* est ainsi difficile à délimiter par le test formel.⁰ Nous considérons que le verbe *comprendre* peut fonctionner comme un marqueur pragmatique tout en gardant le sens cognitif de son emploi en tant que verbe plein.

2.3. Objectif d'étude

En considérant les études antérieures sur *comprendre*, nous pouvons soulever deux questions : (i) quelle serait la relation entre la position de marqueur pragmatique *comprendre* et les fonctions discursives qu'ils exercent ? (ii) quels traits peuvent être dégagés en les analysant quantitativement ? Nous pensons donc qu'il serait nécessaire d'examiner les fonctions discursives de *comprendre* selon ses positions et de montrer quantitativement ses caractéristiques. De nombreux travaux ont été réalisés récemment sur la relation entre les marqueurs pragmatiques et leurs positions (tels que les études dans Beeching and Detges 2014 et Crible 2018). En nous référant à ces études, nous voulons montrer que selon leurs positions, la construction parenthétique du verbe *comprendre* peut remplir des fonctions plus variées, que nous énumérerons dans les tableaux 5 et 6 à la fin de cet article.

3. Corpus et méthodologie

3.1. Corpus et données

Le but de notre recherche étant d'examiner les fonctions pragmatiques, nous nous intéressons aux conversations interactives. Pour obtenir le plus grand nombre d'exemples possible, nous utilisons le Corpus d'étude du français contemporain (CEFC) mis en ligne sur la plate-forme ORFÉO⁰ qui contient 4 millions de mots et 12 sous-corpus au total.

Parmi ces sous-corpus, nous avons sélectionné ceux qui contiennent des conversations interactives. Pour le corpus TUFUS, qui est un des sous-corpus du CEFC, nous y ajoutons d'autres données, qui ont été recueillies en 2015 (dans ce qui suit, nous l'appelons TUFUS 2015.⁰)

En utilisant le moteur de recherche d'ORFÉO, nous avons tout d'abord extrait tous les exemples à la première et à la deuxième personne du présent de l'indicatif. Puis, nous avons dépouillé manuellement les données en respectant les critères définis ci-dessous :

- (i) nous avons choisi les exemples qui sont figés à *je comprends, tu comprends et vous comprenez*. (sans compléments, sans adjoints)
- (ii) parmi les exemples sans compléments, nous avons enlevé les exemples anaphoriques comme (6)

(6) CG: c' est super pr-~ hm plus tu lis et moins **tu comprends** fait
(TUFUS, 10CJTD110913)

→ plus tu lis et moins tu comprends ce que tu lis

Le résultat de notre sélection des données figure dans le tableau 1.⁰

⁰ Le test de zéro-pronominalisation fonctionnerait mieux pour une construction parenthétique d'autres verbes comme par exemple, *tu vois*, pour laquelle on observe une distinction nette entre l'emploi en tant que verbe plein *tu vois ça* et l'emploi en tant que verbe faible *tu vois*.

⁰ La plate-forme est accessible à partir du lien suivant:

<https://repository.ortolang.fr/api/content/cefc-orfeo/11/documentation/site-orfeo/index.html>

⁰ Corpus TUFUS est un corpus élaboré par l'Université des Langues Étrangères de Tokyo. Il recueille des conversations libres et informelles des étudiants de 18 à 25 ans. TUFUS 2005 et 2011 sont accessibles sur la plateforme ORFÉO. Nous avons ajouté à notre corpus le TUFUS 2015 qui a été enregistré à Paris en 2015.

⁰ Nous avons corrigé certaines erreurs de transcription et exclu les enregistrements difficiles à entendre.

Tableau 1: Les sous-corpus et les données de *comprendre*

SOUS-CORPUS	nombre total de mots	tu comprends	vous comprenez	
CFPP	500000	3	2	
CRFP	440000	1	13	
CORALROM	300000	1	0	
CLAPI	210000	0	0	
REUNION	200000	1	0	
VALIBEL	450000	5	37	
TCOF	400000	7	3	
OFROM	330000	0	1	
TUFS	728000	15	1	
TUFS 2015	541213	1	0	
CFPB	10 fichiers	1	0	
TOTAL	4099213	35	57	

3.2. La méthodologie de l'analyse

Nous avons défini la position de *comprendre* par rapport à l'énoncé hôte. Ainsi, nous pouvons distinguer quatre positions.

i. Zone initiale: *comprendre* se situe à la frontière gauche de l'énoncé.

(7) L2: non, mais **tu comprends** c' est un budget prévisionnel

L7: oui non, mais bien sûr

(reunions-de-travail, OF1_CA_3Dec07)

ii. Zone médiane : *comprendre* est intégré à l'intérieur de la construction de l'énoncé hôte.

(8) on en avait peur c' était la terreur un homme imposant assez grand assez fort assez assez un mot en wallon que je vais dire assez cagnès rude **vous comprenez** ouais qui n' était pas beaucoup apprécié par personne

(Valibel, chaBR1r)

iii. Zone finale: *comprendre* se situe à la périphérie droite de l'énoncé.

(9) sous ah d' accord oui c' est aujourd'hui tu as un peu de peine à voyager dans le monde entier hein chaque cinq ans seulement un petit voyage d' une semaine en vacances à cinq cents kilomètres oui c' est limité **je comprends**

(OFROM, unine11c06d)

Dans les exemples attestés, il existe également une position dans laquelle *comprendre* est à la frontière de deux unités de discours autonome. Bolly (2009)⁰ a nommé cette position « transitionnelle » et l'a rangée comme une sous-catégorie parmi les positions initiales. Dans le cas de *comprendre*, qui se place après l'énoncé hôte, nous catégorisons les exemples comme (10) dans la zone finale.

(10) ah oui tiens oui oui c' est c' est un plateau **tu comprends** c' est vaste c' est euh

⁰ Cette étude a effectué une analyse diachronique de *tu vois*.

tu as les montagnes du NNAAMMEE autour

(TUFS, 30_JD_GD_100224)

- iv. Indépendante: *comprendre* est séparé de l'énoncé hôte pour former lui-même un énoncé autonome, pour prendre un tour de parole. C'est dans cette position que *je comprends* est attesté le plus souvent (voir le graphique 2 plus bas).

(11)CT: je veux pas démissionner

MB: ouais **je comprends** ben moi j'aimerais au moins trouver un boulot quoi avant de dire que je veux démissionner

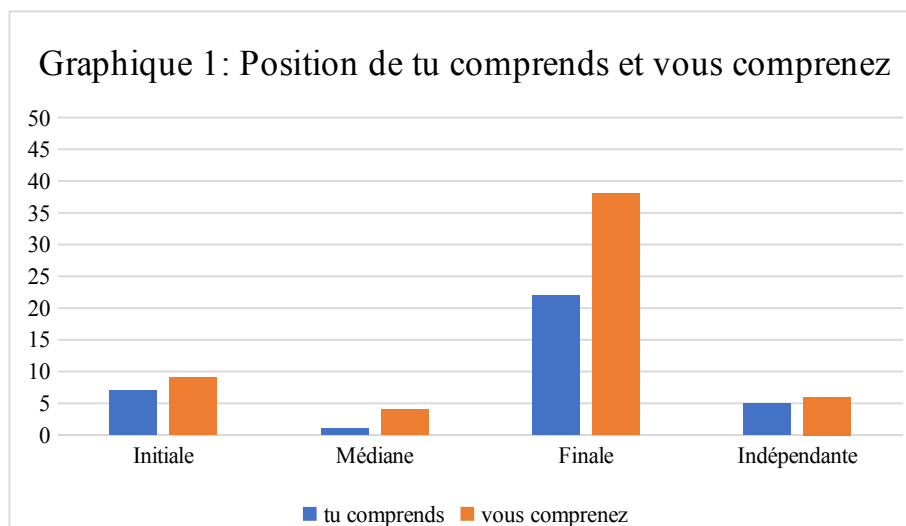
(TUFS, 14_CT_MB_100224)

4. Analyse des données

Dans ce qui suit, nous décrirons d'abord *tu comprends* et *vous comprenez*, les constructions parenthétiques à la deuxième personne, qui ne présentent pas de différences significatives entre elles, sauf au niveau de la politesse. Ensuite, nous examinerons la construction à la première personne, *je comprends*.

4.1. *Tu comprends / vous comprenez*

Voici le résultat de l'analyse. La répartition n'est pas équilibrée. Plus de la moitié des exemples de *tu comprends* et de *vous comprenez* sont en zone finale. Nous avons identifié beaucoup moins d'exemples de la zone initiale et très peu d'exemples de la zone médiane et indépendante. Nous examinerons plus bas leurs fonctions selon les trois positions. Le pourcentage signifie la proportion par rapport à la totalité des occurrences de *tu comprends* (35 cas) et de *vous comprenez* (57 cas).



4.1.1. Les fonctions selon leur position

Placés à la zone initiale, *tu comprends* et *vous comprenez* fonctionnent comme un marqueur annonçant le début d'un énoncé. Ils apparaissent souvent après d'autres connecteurs tels que *parce que*, *puis*, *et*, *mais*. Ils peuvent marquer également l'ouverture d'une citation comme dans (13). Detges et Waltereit (2014) soulignent le fait que la périphérie gauche (i.e. zone initiale) est une zone privilégiée pour des marqueurs pragmatiques jouant un rôle d'ancrage qui marque le début du message. Nous pouvons le constater également dans le cas de *tu comprends* et *vous comprenez*.

La position initiale favorise également les marqueurs pragmatiques exprimant des modalités intersubjectives, qui modalisent les relations interpersonnelles entre le locuteur et

l'interlocuteur. La construction parenthétique *tu comprends / vous comprenez* placée dans cette zone a tendance à inciter l'interlocuteur à comprendre ce que le locuteur lui communique. Elle sert ainsi à attirer l'attention de l'interlocuteur et à appeler sa compréhension de ce qui va être dit.

- (12) LF: est-ce que vous pourriez arrêter de laisser des cadenas
BM: parce que **vous comprenez** c' est et c' est pas que c' est pas joli ça peut être pris comme de l' art, mais
LF: mais on est pas dans une euh une expo
(TUFS, 08LFBM110912)
- (13) je lui explique euh **vous comprenez** j' ai un travail la vie est dure euh j' ai pas beaucoup de temps euh vous savez euh
(OFROM, unine08a26m)

Les exemples de *tu comprends* et *vous comprenez* qui s'insèrent à l'intérieur de l'énoncé sont considérés comme un marqueur appartenant à la zone médiane. Les unités linguistiques qui suivent sont souvent des informations ajoutées ou des précisions pour aider la compréhension de l'interlocuteur ou bien pour articuler une information importante.

- (14) quatre roues et celles de devant elles étaient directrices il y avait un en dessous il y avait un grand rond comme ça en fonte et tout le devant pour su~ suivant que les bêtes tiraient sur l' avant il pivotait il prenait les tournants mieux que l' autre **vous comprenez** ayant quatre roues et alors vous en aviez deux qui étaient là-bas comme ça
(CRFP, PRI-SAI-1)

Lorsque *tu comprends* et *vous comprenez* se placent à la zone finale, ils indiquent la fin d'un message. Detges et Waltereit (2014) remarquent que la périphérie droite (i.e. zone finale) est une position privilégiée pour faire un commentaire sur ce qui a été dit à l'interlocuteur. Il en va de même pour *tu comprends* et *vous comprenez*. Le locuteur commente le dit pour vérifier que l'interlocuteur suit bien la conversation.

- (15)L2: il y avait tout un tas de types là-bas ils attendaient alors je t' y foutais une bouteille de vin et une pièce ils me chargeaient les sacs **tu comprends**
L1: d' accord oui oui oui eh oui
(CRFP > PRI-SAI-1)

Tu comprends et *vous comprenez* peuvent apparaître au milieu d'une phrase en interrompant l'énoncé hôte. Dans (16), l'énoncé hôte est abrégé en *vous comprenez*, qui marque la fin du message impérativement et qui fonctionne pour vérifier si l'interlocuteur comprend l'intention du locuteur sans qu'il l'exprime jusqu'à la fin.

- (16) et puis deuxièmement euh ça un côté snob que **vous comprenez** non moi je
(Valibel, norBB1r)

Nous avons trouvé quelques exemples de *tu comprends* et *vous comprenez* à la position indépendante. Comme on le voit dans (17), la conversation a été interrompue par la réaction de l'interlocuteur. Ici, *vous comprenez* a pour fonction de constater la compréhension de l'interlocuteur comme « je vois que vous comprenez ».

- (17) Pauline_de_Bordes: entre autres entre autres non les propriétaires privés ont continué à louer, mais je veux dire que c' est quand même comme idée
Sonia_Branca-Rosoff: oui

Pauline_de_Bordes: **vous comprenez**

(CFPP2000, Pauline_de_Bordes_F_67_7^e)

4.1.2. *tu comprends, vous comprenez introducteur*

Il arrive que *tu comprends* et *vous comprenez* apparaissant à la zone finale ou transitionnelle introduisent un autre énoncé comme dans l'exemple (19) et (20). Ce sont des énoncés qui constituent une reformulation, un exemple ou une conclusion de ce qui a été dit, afin de renforcer la compréhension de l'interlocuteur.

(18) Reformulation

parce que la la mode est à l'agression comme pour l'oreille pour la musique
vous comprenez les gens n'écouent les gens n'écouent plus de l'harmonie ils
écouent du bruit eh bien l'harmonie c'est pareil dans l'imprimerie

(CRFP, PRO-AIX-1)

(19) Conclusion

ils vont dire ah ben non euh tu dégages ton matériel et puis on va mettre un autre
commerce à ma place **vous comprenez** tout ça c'est un peu le business

(CRFP, PRO-QUI-1)

Dans l'exemple (18), le locuteur reformule ce qui a été dit dans l'énoncé hôte et dans l'exemple (19), le locuteur le résume et conclut. Dans les deux exemples, *comprendre* a pour fonction de thématiser l'énoncé hôte et d'introduire un autre énoncé pour aider l'interlocuteur à comprendre la conversation. Bref, *comprendre* fonctionne comme un introducteur d'un autre énoncé. Nous appelons ce type, *tu comprends / vous comprenez introducteur*⁰. Ce type se place à la zone finale par rapport à l'énoncé hôte, mais on peut aussi dire qu'il se place à la zone médiane entre deux énoncés qu'il relie pour constituer une autre unité discursive plus grande.

Voici la répartition des types d'énoncés qui suivent *tu comprends* et *vous comprenez*. Le pourcentage montre la proportion par rapport à la totalité des occurrences attestées dans la zone finale.

Tableau 2 : Types d'énoncé introduit par *compre*

	tu comprends	vous con
conclusion	0 (0%)	5 (13)
exemple	2 (9%)	1 (3)
reformulation	8(36%)	6 (16)

Le tableau 2 montre que la reformulation est le type d'énoncé le plus fréquent, surtout dans le cas de *tu comprends* ; 36% des exemples figurant dans la zone finale introduisent une reformulation. *Tu comprends* et *vous comprenez* à la zone finale ont tendance à ajouter une modalité intersubjective (i.e. pour vérifier si l'interlocuteur suit bien la conversation) à

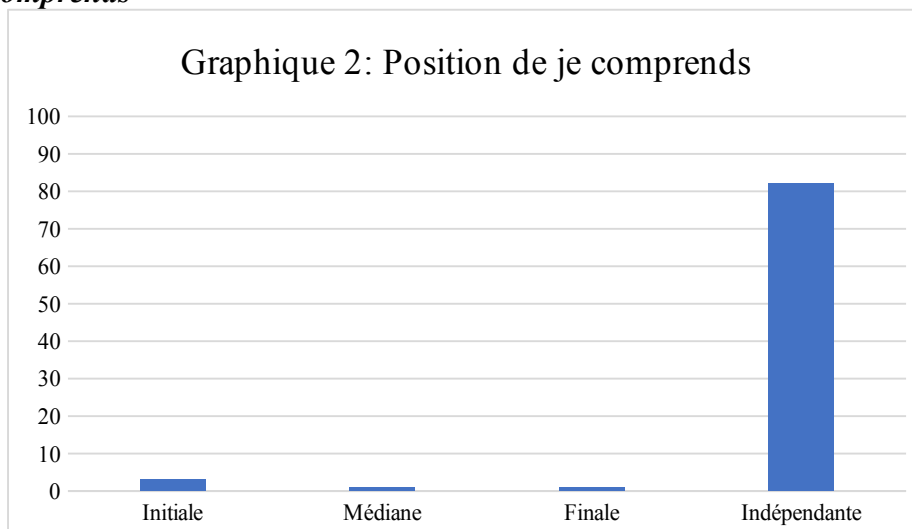
⁰ Nous avons relevé des exemples apparaissant avec d'autres connecteurs. Dans l'exemple ci-dessous, *tu comprends* suivi d'un connecteur *alors* qui sert à introduire une reformulation. Ces exemples sont inclus dans le tableau 2.

on voit tout de suite quelqu'un qui n'est pas de de Charleroi quoi **tu comprends alors** ils ont un un accent soit
euh parce que j'ai des amis qui ont eu un en une fille en p en pension et c'était une flamande même ou j'en
je m'entendais bien et tout ça

(Valibel, accPH1r)

l'énoncé hôte en introduisant un autre énoncé pour aider l'interlocuteur à comprendre ce qui a été dit.

4.2. *Je comprends*



Le graphique 2 présente les positions de *je comprends* par rapport à l'énoncé hôte. La grande partie des données est analysée comme énoncés indépendants. En effet, il est rare que *je comprends* apparaisse à la périphérie de l'énoncé hôte pour communiquer son état de compréhension à l'interlocuteur comme on le voit dans (9) et (20).

(20) **je comprends** alors euh maintenant cela dit vous savez je cela dit je p je crois aussi qu' il y a une certaine comment dirais-je une certaine unité à tous les par exemple nous parlions des Verviétois il est évident que là pff je dirais pour les Canadiens aussi mais les Canadiens ça fait un ensemble enfin n' est-ce pas qui est plaisant et puis on sait qu' ils sont très lointains mais mais un Verviétois euh je je voudrais qu' il se corrige parfois de de sa dénasalisation de sa de sa de des choses comme ça

(Valibel, norHJ1r)

Dans ce qui suit, nous examinerons *je comprends* constituant un énoncé indépendant, qui représente 94% de l'ensemble des données attestées.

4.2.1. Deux types de *je comprends*

Dans l'exemple (21), *je comprends* est séparé de l'énoncé hôte et constitue un tour de parole indépendant.

(21)A: ouais, mais tu sais, mais il paraît normalement c' est pas trop saigné c' est pas soigné les acouphènes
B: quoi ah bon ils savent pas comment les soigner
A: non, mais ils savent pas comment les soigner, mais
B: ah
A: euh ils disent beaucoup ça les mène au suicide
B: ben **je comprends** et je ouais ouais

(TUFS, Kathy)

La fonction de *je comprends* est d'approuver ce que l'interlocuteur a dit. Nous pouvons en observer deux types. Le premier est un cas observé dans (21), qui exprime l'approbation à ce qui a été dit par l'interlocuteur, « ça les mène au suicide ». Le locuteur affirme son état de compréhension du propos de l'interlocuteur.

Le deuxième type constitue plutôt un comportement langagier du locuteur, qui assure à l'interlocuteur qu'il suit bien la conversation et qui incite l'interlocuteur à continuer sa parole, comme on le voit dans (22) ci-dessous :

- (22) BA: euh bien donc ça ça allait hm hm ensuite euh après notre séjour à Porto on est remontés vers Calvi c' est au niveau de Bastia, mais de l' autre côté de l' île
 PE: hm d' accord en fait tu as fait les quatre coins de l' île
 BA: voilà c' est ça
 PE: **je comprends**
 BA: euh à Calvi très sympa, mais euh là on y est surtout restés pour euh décompresser on va dire du voyage et appréhender la suite
 (TUFS, 09BAPE110913)

Cet emploi a une fonction similaire à *je t'écoute* dans le sens où le locuteur essaye de ne pas déranger l'interlocuteur et l'encourage indirectement à continuer son discours.

La répartition des deux emplois figure dans le tableau 3.

Tableau 3: type contenu et comportement de *je coi*

contenu	65 (79%)
comportement	17 (21%)

4.2.2. *je comprends introducteur*

Comme *tu comprends* et *vous comprenez-introducteur* (voir 4.1.2 plus haut), *je comprends* peut aussi être employé pour introduire différents actes de langages. En (23), *je comprends* introduit une reformulation : le locuteur reformule ce qui a été dit par l'interlocuteur sur l'office du tourisme.

(23) Reformulation

BM: quoi, mais tu sais que c' est c' est chiant quand on n' a pas la monnaie par exemple quand je travaillais à l' office de tourisme c' était assez souvent qu' on n' avait pas la monnaie parce que bon on n' avait pas énormément de monnaie

LF: oui ben oui

BM: je sais pas exactement comment ça se passe comment ils renflouaient en monnaie et

LF: tout oui non, mais ouais bien sûr, mais

BM: et du coup, mais les gens ils venaient pour acheter des cartes on était là vous vous avez pas exactement l' à-point non bon ben vous avez pas de carte non s' il vous plaît

LF: non, mais **je comprends** l'office du tourisme tu vois c'est pas un truc euh exclusivement fait pour ça quoi

(TUFS, 08LFBM110912)

Et en (24), *je comprends* introduit un commentaire sur le dit de l'interlocuteur. Dans cet exemple, en disant *je comprends*, le locuteur fait une concession par rapport à ce que l'interlocuteur a dit, tout en ajoutant son commentaire « c'est relatif bien sûr ».

(24) Commentaire

norHJ1: prononciation allant jusqu' à mal prononcer des mots puis alors les tous les accents que nous donnons c' est oui euh le le le no no quelqu'un qui dit moitié par exemple

norFA0: ah oui

norHJ1: bon il y a d il y a des choses comme ça qui sont graves
 norFA0: oui oui oui oui oui oui **je comprends** c' est relatif bien sûr oui
 (Valibel, norHJ1r)

Dans le cas de reformulation en (23), *je comprends* semble avoir une fonction similaire à *tu comprends* et *vous comprenez introducteur* que nous avons vus plus haut. *Je comprends* affirme à l'interlocuteur que le message a été bien compris.

En revanche, *je comprends* de type concession (24) montre un caractère que nous n'avons pas observé dans les exemples de *tu comprends* et *vous comprenez*. *Je comprends* sert comme amortisseur pour donner son opinion. *Je comprends introducteur* est indépendant par rapport à l'énoncé hôte, mais il fonctionne comme un préambule de l'énoncé qui suit. Nous pensons donc qu'il est placé à la zone initiale dans une unité discursive plus grande.

Voici la répartition des types de proposition qui suivent *je comprends*. Le pourcentage note la proportion par rapport aux 65 occurrences de *je comprends* exprimant une approbation sur ce qui a été dit par l'interlocuteur.

Tableau 4: Types de proposition que *je comprends* introduit

reformulation	3 (5%)
commentaire	16 (25%)

Nous pouvons observer dans le tableau 4 qu'au total 25% des exemples de *je comprends* exprimant une approbation fonctionnent en tant qu'*introducteur*. Ces exemples de *je comprends* ne fonctionnent pas comme simple signe d'approbation, mais comme stratégie discursive introduisant un acte de langage, soit pour affirmer sa compréhension à l'interlocuteur, soit pour ajouter son commentaire sur le dit.

5. Conclusion

Dans ce qui précède, nous avons analysé les caractéristiques de *tu comprends*, *vous comprenez* et de *je comprends* selon leur position par rapport à l'énoncé hôte. Les tableaux 5 et 6 présentent le récapitulatif des fonctions de *comprendre*.

Tableau 5: Bilan de *tu comprends* / *vous comprenez*

	Initiale	Médiane	Fi
Fonction	ouverture d'une information	ajout d'une information	fermeture, abréviation
Modalité intersubjective	inciter l'interlocuteur à comprendre la conversation	vérifier la compréhension de l'interlocuteur	vérifier si l'interlocuteur

Tableau 6: Bilan de *je comprends*

	type contenu		
Fonction	approbation	introducteur	
Modalité subjective ou intersubjective	affirmer son état de compréhension	exprimer sa compréhension, une concession	

Pour conclure, notons qu'il existe des différences et des similitudes entre *tu comprends / vous comprenez* et *je comprends*. *Tu comprends* et *vous comprenez* sont relativement libres dans leurs positions. En revanche, *je comprends* est plus fixe. Il constitue normalement un énoncé autonome. Les trois constructions partagent pourtant un caractère en commun : elles affirment la compréhension de ce qui a été dit, soit par le locuteur, soit par l'interlocuteur.

La construction parenthétique du verbe *comprendre* contribuant à la structuration discursive est considérée comme un marqueur pragmatique, mais elle n'a pourtant pas connu la grammaticalisation. Elle partage toujours son sens cognitif de base avec son emploi en tant que verbe plein.

Il serait maintenant intéressant de comparer *comprendre* avec d'autres marqueurs pragmatiques. Bien que *tu comprends* et *tu vois* aient un sens cognitif similaire, *tu comprends* apparaît très souvent à la zone finale, mais *tu vois*, principalement à la zone initiale,⁰ nous pourrions dégager plus clairement les particularités de *comprendre* en le comparant avec d'autres marqueurs pragmatiques ayant des fonctions similaires. Nous voudrions aborder ce sujet dans une prochaine étude.

Bibliographie

- ANDERSEN, H. (2007). Marqueurs discursifs propositionnels. *Langue française*, 2(2), 13-28. <https://doi.org/10.3917/lf.154.0013>
- BEECHING, K., DETGES, U. (2014). *Discourse functions at the left and right periphery. Crosslinguistic investigations of language use and language change*, Leiden: Brill.
- BENZITOUN, C., DEBAISIEUX, J.-M., DEULODEU, J. (2016). Le projet ORFÉO: un corpus d'étude pour le français contemporain, *Corpus [En ligne]*, 15, mis en ligne le 15 janvier 2017, consultés le 12 mars 2021. URL: <http://journals.openedition.org/corpus/2936>; DOI : <https://doi.org/10.4000/corpus.2936>.
- BLANCHE-BENVENISTE, C., DEULOFEU, J., STÉFANINI, J., VAN DEN EYNDE, K. (1984). *Pronom et syntaxe. L'approche pronominale et son application au français*, Paris: SELAF.
- BLANCHE-BENVENISTE, C. (1989). Construction verbale « en incise » et rection faibles des verbes, *Recherche sur le français parlé*, 9, 53-74.
- BLANCHE-BENVENISTE, C., (1990). *Le français parlé: Études grammaticales*, Paris: Éditions du CNRS.
- BLANCHE-BENVENISTE, C., (2010). *Le français: Usages de la langue parlée*, Paris / Leuven/Paris: Peeters.
- BOLLY, C. (2009). Constructionnalisation et structure informationnelle. Quand la grammaticalisation ne suffit pas pour expliquer *tu vois*, *Linx [En ligne]*, 61, mis en ligne le 01 juin 2013, consulté le 07 juillet 2021. URL : <http://journals.openedition.org/linx/1342> ; DOI : <https://doi.org/10.4000/linx.1342>
- BOLLY, C. (2012). Du verbe de perception visuelle au marqueur parenthétique "tu vois": Grammaticalisation et changement linguistique. *Journal of French Language Studies*, 22(2), 143-164.
- CRIBLE, L. (2018). *Discourse markers and (dis) fluency: Forms and functions across languages and registers*. Amsterdam / Philadelphia: John Benjamins.
- DEBAISIEUX, J.-M. (éd.). (2013). *Analyses linguistiques sur corpus: subordination et insubordination en français.*, Paris: Lavoisier.
- DETGES, U., WALTEREIT, R. (2014). *Moi je ne sais pas vs Je ne sais pas moi: French disjoint pronouns in the left vs. right periphery. Discourse functions at the left and right periphery. Crosslinguistic Investigations of Language Use and Language Change*. Leiden: Brill.
- DOSTIE, G. (2004). *Pragmaticalisation et marqueurs discursifs: Analyse sémantique et traitement lexicographique*. Bruxelles: De Boeck Duculot.

⁰ Voir l'analyse de Bolly (2009).

- DOSTIE, G. & PUSCH, C. (2007). Présentation. Les marqueurs discursifs. Sens et variation. *Langue française*, 154, 3-12. <https://doi.org/10.3917/lf.154.0003>
- FEDRIANI, C., SANSÓ, A. (2017). Pragmatic markers, discourse markers and modal particles: What do we know and where do we go from here?, *Pragmatic markers, discourse markers and modal particles: New Perspectives*, Amsterdam/Philadelphia: John Benjamins, 1-33.
- GOLDBERG, A-E. (2019). *Explain me this: Creativity, Competition, and the Partial Productivity of Constructions*, Princeton University Press.
- KALTENBÖCK, G. (2009). English comment clauses: Position, prosody, and scope. *AAA-Arbeiten Aus Anglistik Und Amerikanistik*, 34(1), 51-78. Retrieved July 19, 2021, from <http://www.jstor.org/stable/26430887>
- SCHNEIDER, S. (2007). *Reduced parenthetical clauses as mitigators: A corpus study of spoken French, Italian and Spanish*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- URMSON, J. (1952). Parenthetical verbs, *Mind*, 61, 480-496.

Formules parenthétiques avec *dire* dans le discours scientifique oral

Chaeyoung Lee, LIDILEM Université Grenoble Alpes

Agnès Tutin, LIDILEM Université Grenoble Alpes

1. Introduction

Dans la continuité de nombreux travaux sur le lexique et la phraséologie transdisciplinaires du discours scientifique écrit en français (Grossmann & Tutin : 2014, Jacques & Tutin : 2018), ce travail s'inscrit dans une étude plus large sur les « routines sémantico-rhétoriques »¹ qui remplissent certaines fonctions discursives et/ou rhétoriques dans le discours scientifique oral avec l'objectif de guider les interlocuteurs vers une meilleure compréhension de la communication. Dans cette perspective, nous souhaitons étudier les formules parenthétiques construites avec le verbe *dire*, fréquentes dans les communications scientifiques orales, qui semblent particulièrement intéressantes pour observer le fonctionnement dialogique du discours scientifique oral.

Encore peu étudié par manque de données disponibles en français, le discours scientifique oral se distingue très nettement du discours scientifique écrit par un ensemble de paramètres : la réalisation du discours en temps réel, qui se caractérise souvent par des contraintes de temps, la présence directe des interlocuteurs pendant la production du discours, ou encore l'interaction implicite entre le locuteur et son auditoire avec l'utilisation du support visuel (Jacques : 2017).

À l'intersection des études du français parlé, de la phraséologie et du discours scientifique, notre travail part de l'hypothèse que l'environnement discursif oral détermine l'utilisation des expressions préfabriquées qui sont spécifiques au genre de discours. Dans cette communication, nous tentons d'analyser plus précisément les formules parenthétiques construites avec le verbe *dire*, récurrentes dans les communications scientifiques orales, en les comparant avec celles des interactions en entretiens. Ce travail poursuit un triple objectif : 1) mettre en évidence la productivité discursive de ces formules dans le discours scientifique oral, 2) déterminer leurs différents rôles rhétoriques, au-delà de la fonction principale qui est « métalinguistique ou métaénonciative », 3) démontrer la diversité des sous-registres du genre oral, par une analyse contrastive des corpus, et plus particulièrement la particularité du discours scientifique oral.

Après une définition de notre objet d'étude, nous présenterons la démarche méthodologique que nous avons suivie, de la description de deux corpus jusqu'à l'annotation des fonctions que remplissent les formules, en passant par les critères du repérage et des difficultés rencontrées.

¹ Relevant de la « phraséologie étendue » (Legallois & Tutin 2013), les routines sémantico-rhétoriques sont définies ici comme des formules phrastiques récurrentes étroitement liées à des fonctions discursives et/ou rhétoriques spécifiques à un genre discursif (Kraif & Tutin : 2017, Tutin : 2018, Ji & Tutin : 2018).

Ensuite, nous montrerons les résultats de notre analyse fonctionnelle autour des formules les plus productives dans les deux corpus, ce qui permettra de mettre en évidence des pistes intéressantes. Pour terminer, nous présenterons quelques perspectives de l'étude qui nous paraissent prometteuses.

2. Critères définitoires des formules parenthétiques avec *dire*

La notion de « construction parenthétique » a été traitée sous plusieurs dénominations et selon différents points de vue en linguistique (Schneider : 2007) et reste jusqu'ici l'objet de nombreux débats entre linguistes. Les verbes « parenthétiques » (Urmson : 1952) ou « recteurs faibles » (Blanche-Benveniste : 1989, Andersen : 2007), propres à ces structures parenthétiques (*je crois, tu vois, semble-t-il, etc.*) ont été décrits sous l'angle syntaxique, sémantique et prosodique, ainsi qu'autour des questions de grammaticalisation et de pragmaticalisation.

Nombre de structures parenthétiques peuvent aussi être envisagées comme des sortes de phrases préfabriquées, dans la mesure où elles apparaissent récurrentes et préconstruites dans certains genres discursifs. À titre d'exemple, les structures formées sur le verbe de constat *voir* (*on le voit, nous le voyons, etc.*) sont très caractéristiques du discours scientifique écrit, où elles remplissent une fonction évidentielle de co-constat dans la mise en scène du dialogue avec le lecteur dans le texte, mais aussi une fonction de cohérence discursive (Grossmann : 2014, Grossmann & Tutin : 2010).

Dans cette communication, à la suite des travaux de Blanche-Benveniste (1989) sur la notion de « rection faible » des verbes, nous définissons les formules parenthétiques comme des formules phrastiques (assertives, injonctives, interrogatives) sans rection, ni subordonnant, insérées dans une autre proposition en position initiale, médiane ou finale².

3. Démarche méthodologique

Pour cette étude, nous avons constitué un corpus du discours scientifique oral en français, nommé CComSci (Corpus de COMmunications SCIENTIFIQUES), et avons exploité, à titre de comparaison, le Corpus du Français Parlé Parisien des années 2000 (CFPP2000) (Branca-Rosoff et al. : 2012). Nous avons ensuite repéré les formules parenthétiques avec le verbe *dire* dans les deux corpus, puis les avons annotés en analysant leurs fonctions rhétoriques en fonction du contexte dans chaque corpus.

3.1. Corpus

Pour mener notre étude sur le discours scientifique oral en français, nous avons élaboré un corpus de communications scientifiques orales (CComSci) dans trois disciplines des Sciences

² Notre définition est ainsi plus étroite que celle d'autres auteurs comme Delais-Roussarie (2006) ou Gachet & Avanzi (2008).

Humaines et Sociales : la linguistique (dont un sous-ensemble est issu du corpus EIIDA³), les sciences de l'information et de la communication et la didactique des langues. Les communications ont été manuellement transcrites, puis segmentées selon la convention de segmentation proposée par André Valli et José Deulofeu sur la plate-forme ORFEO (Outils et Ressources sur le Français Ecrit et Oral) et adoptée pour le « corpus Gold » du CEFC (Corpus d'Etude pour le Français Contemporain).⁴ Ici, il est à noter que cette segmentation s'inspire de l'approche de Blanche-Benveniste (2002), pour laquelle les unités de segmentation du corpus oral sont des constructions verbales, non des phrases.

Nous avons ensuite effectué l'alignement sonore sur le logiciel ELAN avec objectif de diffuser un sous-ensemble du corpus, dont l'autorisation a été obtenue, avec la version audio. Toutefois, pour l'annotation fonctionnelle des expressions dans ce travail, nous nous sommes contentées de travailler principalement sur les transcriptions, mais en faisant un aller-retour entre le texte et l'audio pour prendre en compte les éléments sonores autour des expressions. La composition du corpus est résumée dans le tableau suivant.

Sous-discipline	Nombre de communications	Durée
Linguistique	20	8h 51m
Sciences de l'Information et de la Communication	20	8h 12m
Didactique des Langues	20	9h 01m
Total	60	26h 04m

Tableau 1. Description du corpus de communications scientifiques (CComSci)

Par la suite, pour réaliser une analyse comparative, nous avons choisi un sous-ensemble du CCFP2000, corpus qui a également été étudié par Steuckardt (2018) sur les marqueurs de reformulation formés sur *dire*. Ce corpus d'entretiens semi-directifs individuels et collectifs a été constitué dans le cadre de l'enquête menée par le programme de recherche de la ville de Paris depuis l'année 2008. Parmi les 48 entretiens, nous en avons sélectionné 22, tout en conservant la diversité de la composition du corpus entier dans différents aspects (sexe, âge, quartier d'habitation des interviewés). Cela nous a permis d'obtenir un sous-corpus d'environ 26 heures, d'une durée comparable à notre corpus CComSci.

	CComSci	Sous-corpus du CCFP2000
Type	Communications scientifiques orales en Sciences Humaines et Sociales	Entretiens semi-directifs individuels ou collectifs
Unités	60 communications	22 entretiens
Durée totale	26h 04m	25h 59m

Tableau 2. Description des deux corpus comparés

³ Le projet EIIDA (Études Interdisciplinaires et Interlinguistiques du Discours Académique) a été mené dans le cadre du laboratoire LATTICE entre 2012 et 2017 avec l'objectif de travailler l'impact de trois paramètres, qui sont le mode de transmission (écrit/oral), la langue (anglais, français, espagnol) et la catégorie disciplinaire (sciences exactes-géochimie/sciences humaines-linguistique), sur le paysage du discours scientifique. Ils ont élaboré le corpus EIIDA, constitué de 15 articles de recherche et de 15 communications orales transcrites pour chaque discipline dans chaque langue (Carter-Thomas & Jacques : 2017).

⁴ URL: https://repository.ortolang.fr/api/content/cefc-orfeo/head/documentation/guide_segmentation_orfeo.pdf. Dernière consultation le 24 août 2021.

3.2. Méthodes d'analyse

À partir de ces deux corpus oraux, nous avons repéré, à l'aide de concordances intégrant le lemme *dire* (à partir de l'outil *AnaText*⁵), les candidats de formule parenthétique avec le verbe *dire*, et leur contexte. En appliquant notre définition des formules parenthétiques, nous avons ainsi retenu les formules phrastiques, qui sont soit assertives avec le pronom personnel de la première personne *je* et *on*, soit injonctives (*disons*) ou interrogatives (*comment dire*, *comment dirais-je*, etc.), mais sans rection, ni subordonnant. Ces critères nous ont conduites à exclure des formules comme celles qui incluent un pronom anaphorique *le* (par exemple, *on l'a dit*) et celles qui intègrent un subordonnant en *comme* ou *si* (*comme on dit*, *comme j'ai dit*, etc.), s'apparentant à une proposition subordonnée. Par ailleurs, le repérage des formules n'a pas été trivial parce que les deux corpus n'ont pas été transcrits ni segmentés selon les mêmes conventions.

Afin de procéder à l'analyse des formules, nous avons effectué une annotation des fonctions rhétoriques sur l'ensemble des formules repérées dans le CComSci et, pour le CFPP2000, les 3 formules les plus fréquentes en commun dans les deux corpus (*disons*, *on va dire*, *je veux dire*). Pour cela, nous avons pris en compte de différents éléments comme la position syntaxique par rapport à la portée, la prosodie, l'accentuation, les pauses, les balbutiements, etc. Inspiré des analyses sémantico-pragmatiques de Steuckardt (2005, 2016, 2018) sur les marqueurs discursifs en *dire*, ce travail est également parti de l'hypothèse que la fonction principale de certaines formules parenthétiques avec le verbe *dire* est de reformuler une formulation précédente en y ajoutant une valeur supplémentaire comme la précision, l'exemple, la correction, etc., comme le propose Saunier (2012) pour la formule injonctive *disons*.

À côté de cette fonction principale de Reformulation, nous pouvons aussi observer certaines formules qui indiquent la recherche d'une formulation adéquate ou qui commentent l'approximation de la formulation réalisée. D'autres contribuent plutôt à la structuration discursive (Introduction, Rappel) ou à mettre l'accent sur des éléments importants (Insistance). Le tableau 3 présente les différents cas de figure rencontrés dans les deux corpus.

Catégorie des fonctions	Description	Exemples
Reformulation	Reformuler une formulation précédente	ensuite le secteur puis l'environnement disons <u>les conditions de travail</u> <u>les les</u> interlocuteurs euh de de du candidat puis les missions puis le contrat (1A_07 – CComSci)
Recherche	Indiquer la recherche d'une formulation adéquate	comment ces effets d'orthographe et de cohérence se retrouvent euh au niveau de

⁵ Développé par Olivier Kraif dans le laboratoire LIDILEM, *Anatext* est un outil d'analyse lexicométrique basé sur l'étiqueteur Treectagger qui permet différentes extractions (vocabulaire spécifique ou catégorisé par partie du discours) ou recherches (concordance, cooccurrences, patterns morphosyntaxiques, segments répétés) des corpus.

		euh de la comment on va dire de la de la conscience même de l'élève (1B_01 – CComSci)
Approximation	Commenter l'approximation de la formulation réalisée	ça a été importé en France traduit euh et euh et euh <u>médiatisé on va dire</u> publicisé en tout cas par euh Casilli et Cardon (2A_17 – CComSci)
Introduction	Introduire un énoncé	je veux dire on a tous grandi avec du texte et de l'image (3E_03 – CComSci)
Rappel	Renvoyer à ce qui a été dit précédemment	c'est extrêmement contraignant ... parce que je vous ai dit on est euh quatre universités euh quatre sites etcetera (3C_01 – CComSci)
Insistance	Mettre l'accent sur des éléments importants	mais c'est même pas c'est même pas ça en fait c'est <u>il y a des priorités</u> quoi je veux dire enfin ça c'est autre chose (MO-01 – CFPP2000)

Tableau 3. Fonctions des formules parenthétiques avec le verbe *dire*

L'annotation des fonctions est complexe, car certaines formules remplissent deux fonctions en même temps, par exemple celle de recherche et celle d'approximation, comme dans l'exemple suivant :

donc même dans le les occurrences euh on va dire de défigement du syntagme on a cet élément qui est retranscrit qui est un un discours de presse plus stéréotypé (1A_08 – CComSci)

Pour décider les cas ambigus de façon cohérente, nous avons appliqué certains critères syntaxiques et prosodiques comme ci-dessous.

- 7) Pour la fonction de Reformulation, deux formulations sont exigées, l'une qui est reformulée et l'autre qui la reformule.
- 8) Les formules de Recherche sont antéposées par rapport à la portée du verbe *dire* et souvent accompagnées par une petite pause, *euh*, *hmm*, ou les balbutiements avec la répétition des mots grammaticaux (*de*, *à*, *le*, *la*).
- 9) Les formules d'Approximation sont dans la plupart du temps postposées par rapport à la portée du verbe *dire*, mais peuvent être antéposées quand la portée participe à une sorte de dénomination ou de qualification.
- 10) Quant à la fonction de l'Introduction, les formules sont toujours antéposées par une phrase à introduire.
- 11) Les formules d'Insistance sont souvent postposées par rapport à la portée du verbe *dire*, c'est-à-dire un énoncé à mettre l'accent, et parfois prononcées rapidement avec le mot *enfin* situé devant elles.

Malgré la taille réduite des corpus, cette annotation fonctionnelle nous permet d'entrevoir globalement les différents rôles rhétoriques des formules parenthétiques avec le verbe *dire* et de comparer les deux sous-registres oraux, communications scientifiques et interactions en

entretiens, dans cet aspect.

4. Résultats de l'analyse

L'examen des deux corpus révèle que les communications scientifiques comportent beaucoup moins de formules parenthétiques avec le verbe *dire* (142 occurrences) que les interactions en entretiens (356 occurrences), ce qui pourrait probablement confirmer les résultats de Biber (2006) sur la proximité de ce genre oral avec celui des écrits scientifiques comme articles de recherche, mémoires ou thèses. Ce relevé nous a permis de mettre en évidence quelques points intéressants que nous présentons par la suite à l'aide des tableaux et schémas descriptifs et/ou comparatifs.

Tout d'abord, le tableau 4 ci-dessous présente toutes les formules repérées, dont certaines sont regroupées en une seule catégorie selon leur proximité syntaxique et/ou sémantique, avec le nombre d'occurrences et la proportion sur l'ensemble des formules dans chaque corpus.

	CComSci		CFPP2000	
	Nombre d'occurrences	Fréquence	Nombre d'occurrences	Fréquence
disons	53	37.32%	20	5.62%
on va dire	31	21.83%	95	26.69%
on dira	1	0.70%	0	0.00%
je veux dire	6	4.23%	102	28.65%
S + avoir envie de + DIRE (j'ai envie de dire, j'aurais envie de dire)	1	0.70%	3	0.84%
S + pouvoir + DIRE (je peux/pourrais dire, on peut/pourrait dire)	6	4.23%	11	3.09%
je (vous) dirais	14	9.86%	19	5.34%
j'allais dire	5	3.52%	11	3.09%
je dois dire	0	0.00%	1	0.28%
S + DIRE présent (je (te/vous) dis, on dit)	0	0.00%	35	9.83%
je vais (te/vous) dire	2	1.41%	6	1.69%
S + DIRE passé (j'ai dit, je vous ai/avais dit, je disais, on a dit, etc.)	7	4.93%	13	3.65%
comment (S + AUX) + DIRE (comment dire, comment on pourrait dire, comment dirais-je)	16	11.27%	40	11.24%
Total	142	100%	356	100%

Tableau 4. Formules parenthétiques avec le verbe *dire* dans les deux corpus

La comparaison des 5 formules les plus productives (Schéma 1) permet de contraster assez clairement les deux sous-registres.

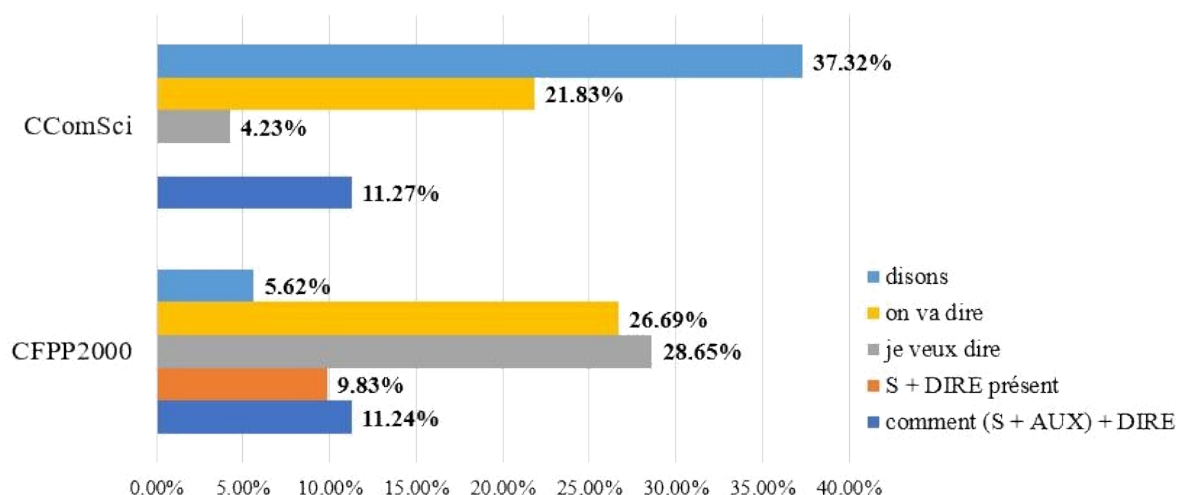


Schéma 1. Les formules parenthétiques avec le verbe *dire* les plus productives dans les deux corpus

Dans le schéma 1, nous constatons des différences significatives. Un des résultats remarquables est que la catégorie des formules [S + DIRE présent] (*je te dis, je vous dis, on dit*), qui jouent une fonction de l'insistance subjective de la part du locuteur sur ce qui est dit dans l'interaction, est absente dans le CComSci, alors qu'elle est présente à 9.83% dans le CFPP2000. Cela nous montre que les communications scientifiques orales sont un registre plus formel par rapport aux interactions en entretiens, dans la mesure où, pour mettre l'accent sur certains éléments du discours, le locuteur utilise d'autres moyens comme des routines (*ce qui est important c'est que..., il est essentiel de noter que..., etc.*) plutôt que des formules parenthétiques avec le verbe *dire* qui relèvent de l'usage familier.

Ensuite, nous relevons également que *disons* est beaucoup plus fréquent dans le CComSci (37.32%) que dans le CFPP2000 (5.62%). L'annotation fonctionnelle nous montre que *disons* a souvent une fonction de Reformulation (15 sur 53 occurrences = 28.3%) davantage que d'autres marqueurs comme *on va dire* (une seule occurrence sur 31 cas de reformulation). Selon Steuckardt (2016 : 300), *disons* crée une voix du locuteur « à la fois plus guidée et plus autoritaire ». Cela nous conduit à conclure que, dans le discours scientifique oral, la reformulation en *disons* contribue à renforcer l'éthos du locuteur-chercheur, voire son crédit scientifique, en affirmant ses choix terminologiques et théoriques.

On observe également que *je veux dire* est peu fréquent dans le CComSci (4.23%) par rapport au CFPP2000 (28.65%). Cette faible fréquence d'un marqueur de glose « auto-correctif » est probablement liée au fait que les communications scientifiques orales sont un genre discursif préconstruit plutôt que spontané, où le locuteur s'auto-corrige moins. Steuckardt (2005) montre que *je veux dire* s'inscrit dans le cadre d'un « auto-dialogisme », contrairement à d'autres marqueurs de glose, comme *autrement dit* ou *c'est-à-dire*, qui peuvent aussi intervenir dans un « hétéro-dialogisme » (Steuckardt : 2005). Cela est aussi confirmé par notre annotation fonctionnelle qui montre que, parmi les trois formules de reformulation (*disons, on va dire, je veux dire*), *je veux dire* est la plus dominante (49 sur 69 occurrences = 81.67%) dans le CFPP2000, alors que cette formule présente seulement 2 occurrences sur 18 au total dans le CComSci.

À partir de cette analyse globale, nous avons procédé à l'annotation fonctionnelle sur les trois formules les plus productives dans les deux corpus, qui sont *disons*, *on va dire*, *je veux dire*, dont on peut résumer le résultat comme dans les schémas 2 et 3 suivants.

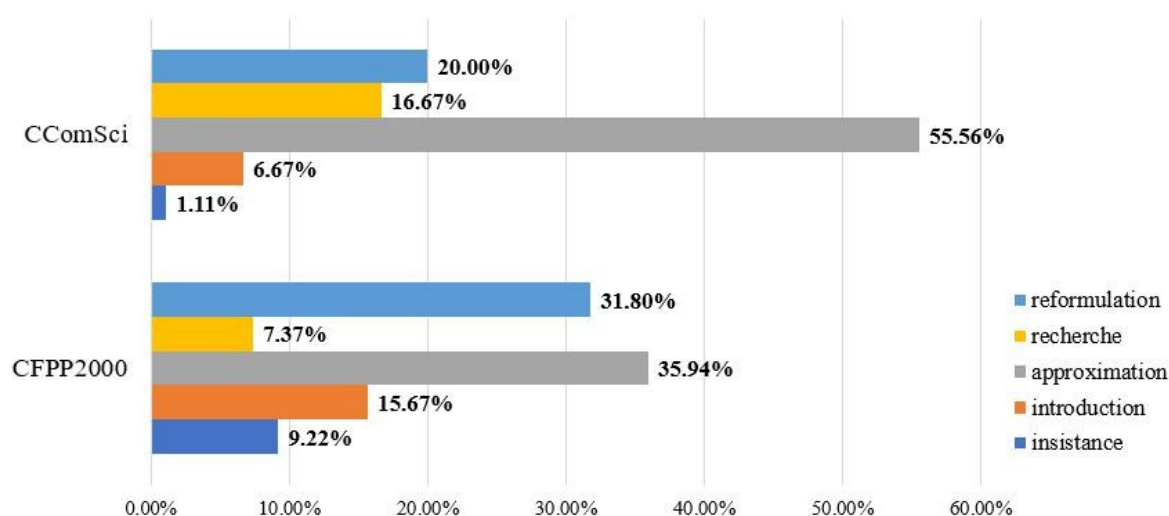


Schéma 2. Fonctions des formules parenthétiques avec le verbe *dire* dans les deux corpus

Tout d'abord, nous constatons plus de proximité entre les deux sous-registres dans les fonctions réalisées par les formules (schéma 2) que dans les formules elles-mêmes (schéma 1). Toutefois, dans le CComSci, la fonction d'Approximation (55.56%) joue un rôle plus important que celle de Reformulation (20%) ou de Recherche (16.67%), alors que dans le CFPP2000 les fonctions d'Approximation (35.94%) et de Reformulation (31.80%) présentent à peu près la même proportion. Cela nous amène à conclure que l'approximation de la formulation réalisée, c'est-à-dire la modulation de la prise en charge de la formulation, est un enjeu stratégique essentiel dans le discours scientifique oral. De plus, parmi les 50 occurrences d'approximation dans le CComSci, 28 occurrences représentant plus que la moitié sont situées avant la portée du verbe *dire*, ce qui crée en même temps un effet de recherche de la formulation adéquate. Nous pouvons en conclure que, dans les communications scientifiques, le locuteur se montre plus prudent vis-à-vis sa formulation que dans les interactions en entretiens parce que dans le discours scientifique il est essentiel de communiquer les termes ou les significations appropriés pour une meilleure compréhension du discours chez les interlocuteurs.

Ensuite, la fonction de Recherche est plus significative dans le CComSci (16.67%) que dans le CFPP2000 (7.37%). Et comme nous l'avons dit par rapport à la haute fréquence de *disons* dans le CComSci, dans la plupart des cas c'est *disons* (14 sur 15 occurrences) qui prend en charge de cette fonction. Il nous semble qu'avec la formule injonctive *disons*, le locuteur engage les interlocuteurs dans sa recherche de formulation adéquate. Autrement dit, il construit une forme de coopération dans cette construction de la formulation avec les interlocuteurs.

Nous pouvons également constater qu'il y a plus de fonction de Reformulation dans le CFPP2000 (31.08%) que dans le CComSci (20%), ce qui pourrait apparaître surprenant parce que la reformulation est effectivement une des opérations métalinguistiques primordiales dans le discours scientifique (Ji & Tutin : 2019). En effet, comme nous l'avons dit précédemment,

c'est *je veux dire* qui occupe la plupart des cas de la reformulation (81.67%) dans les interactions en entretiens, en contribuant à illustrer plus précisément la formulation précédente comme dans l'exemple suivant :

et est-ce que ça continue à changer je veux dire est-ce que [+] de la petite ou moyenne bourgeoisie tu le vois [+] basculer vers une gentrification encore plus importante ou pas (15-02_CFPP2000)

Cet « auto-dialogisme » (Steuckardt : 2005) de reformulation créé par *je veux dire* est beaucoup moins présent dans le CComSci et c'est plutôt *disons* qui y domine parmi les formules de reformulation. De ce fait, nous pouvons en conclure que, dans les communications scientifiques, la reformulation a plutôt pour objectif de fournir une précision pour les termes utilisés en engageant les interlocuteurs dans la construction du discours. L'hétéro-dialogisme engendré par la participation des interlocuteurs dans le discours et la co-construction des savoirs est une des clés du guidage des interlocuteurs vers une meilleure compréhension de la communication dans le discours scientifique oral.

Enfin, nous avons effectué l'annotation fonctionnelle sur le reste des formules repérées du corpus CComSci, dont le résultat est décrit comme dans le schéma 4 ci-dessous.

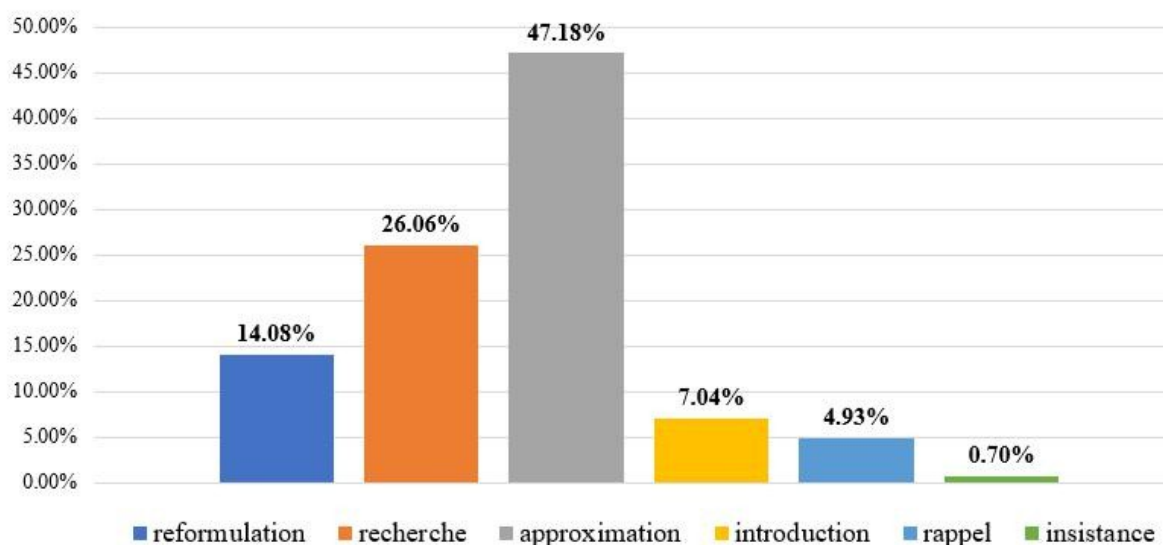


Schéma 4. Fonctions des formules parenthétiques avec le verbe *dire* dans le CComSci

À côté des principales fonctions métalinguistiques ou métaénonciatives évoquées ci-dessus, dans le CComSci nous repérons également des fonctions qui contribuent à la structuration discursive comme l'Introduction ou le Rappel. Toutefois, leur faible fréquence ne veut pas dire qu'elles ne jouent pas un rôle important dans les communications scientifiques orales. En effet, ces fonctions de la catégorie de Structuration Discursive se manifestent sous des structures préfabriquées qui ne sont pas intégrées dans notre définition des formules parenthétiques (par exemple, *comme je viens de dire*, *comme je disais*, *comme j'ai dit tout à l'heure* pour le Rappel), et surtout avec d'autres verbes comme *parler* (... *dont j'ai parlé tout à l'heure*) ou *montrer* (*comme je vous ai montré dans l'introduction*), que nous ne pourrions pas détailler dans cette communication.

5. Conclusions et perspectives

Malgré un faible nombre d'occurrences, les formules parenthétiques avec le verbe *dire* sont bien présentes dans les interactions orales. Leur étude fonctionnelle permet de comprendre les différentes valeurs rhétoriques qui peuvent apparaître à travers la structure parenthétique ainsi que le verbe *dire*. Plus particulièrement, nous avons effectué une analyse comparative entre un corpus de communications scientifiques et un corpus d'entretiens semi-directifs. Cette comparaison a pour objectif de mettre en évidence la particularité des communications scientifiques orales en tant que genre discursif spécifique, où en raison de la spécificité des informations scientifiques il est important de commenter sur « ce qui est dit » ou « ce qui est à dire » pour aider les interlocuteurs à ne pas perdre le fil du discours.

À partir de l'annotation fonctionnelle des formules, nous avons pu remarquer que les fonctions métalinguistiques ou métaénonciatives réalisées par les formules parenthétiques avec le verbe *dire* restent assez comparables dans les deux registres oraux. C'est davantage dans la proportion des formules utilisées que l'on observe des différences significatives. Cela valide la principale hypothèse de notre étude, qui vise à démontrer que les communications scientifiques orales constituent un genre distinct d'autres registres oraux. Pour confirmer cette hypothèse, il nous reste toujours à développer ce travail préparatoire en affinant la typologie des fonctions proposées et aussi en analysant d'autres phénomènes phraséologiques comme par exemple des formules pragmatialisées en *comme* ou *si* (*comme on dit, si je puis/peux dire, etc.*) qui ne sont pas étudiées dans cette communication, ou des formules préfabriquées avec d'autres verbes comme *sembler* (*semble-t-il, me semble-t-il*), *voir* (*vous voyez ?*) ou *savoir* (*vous savez ?*).

Dans la perspective d'une étude plus étendue sur les routines sémantico-rhétoriques dans le discours scientifique oral, ce travail est une des premières exploitations outillées du corpus des communications scientifiques (CComSci). Cette étude sera complétée par l'exploration d'autres dimensions phraséologiques comme les formules parenthétiques avec d'autres verbes ou les routines sémantico-rhétoriques qui remplissent certaines fonctions spécifiques au discours scientifique.

Corpus et outils d'analyse

- [CFPP2000] Discours sur la ville. Corpus de Français Parlé Parisien des années 2000, CLESTHIA (EA 7345 & Université Paris 3). URL : <http://cfpp2000.univ-paris3.fr/>
- [Corpus EIIDA] Études Interdisciplinaires et Interlinguistiques du Discours Académique, LATTICE (CNRS/ENS/Université Paris 3). URL : <http://scientext.msh-alpes.fr>
- [AnaText] URL : <http://phraseotext.univ-grenoble-alpes.fr/anaText/>

Références bibliographiques

- Andersen, H.L. (2007). Marqueurs discursifs propositionnels. *Langue française* 154. pp. 13-28.
- Biber, D. (2006). *University Language: A corpus-based study of spoken and written registers*. Amsterdam-Philadelphie: John Benjamins.
- Blanche-Benveniste, C. (1989). Constructions verbales « en incise » et rectification faible des verbes.

- Recherches sur le français parlé* 9. pp. 53-73.
- Branca-Rosoff, S. et al. (2012). Discours sur la ville. Présentation du Corpus de Français Parlé Parisien des années 2000 (CFPP2000). Consulté le 24 août 2021.
<http://cfpp2000.univ-paris3.fr/CFPP2000.pdf>
- Carter-Thomas, S. & Jacques, M.P. (2017). Interdisciplinary and interlinguistic perspectives on Academic Discourse: the mode variable. Introduction to the special issue on the French EIIDA project. *CHIMERA Romance Corpora and Linguistic Studies* 4(1). pp. 1-11.
- Delais-Roussarie, E. (2006). La prosodie des incidentes en français. *Cahiers de Grammaire* 30 « Spécial Anniversaire ». pp. 129-138.
- Gachet, F. & Avanzi, M. (2008). La prosodie des parenthèses en français spontané. *Verbum* 30(1). pp. 53-84.
- Grossmann, F. (2014). De quelques routines phraséologiques liées aux verbes parenthétiques dans les genres scientifiques. *IV Congrès Mondial de Linguistique Française*, 19-23 Juillet 2014, Berlin. pp. 759-770.
- Grossmann, F. & Tutin, A. (2010). Evidential Markers in French Scientific Writing: The Case of the French Verb *voir*. In E. Smirnova & G. Diewald (éds), *Evidentiality in European Languages. Empirical Approaches to Language Typology*. Berlin-New York: Mouton de Gruyter. pp. 279-308.
- Grossmann, F. & Tutin, A. (2014). *L'écrit scientifique : du lexique au discours. Autour de Scientext*. Rennes: Presses universitaires de Rennes.
- Jacques, M.P. (2017). La structuration textuelle en discours scientifique: comparaison oral/écrit. *CHIMERA Romance Corpora and Linguistic Studies* 4(1). pp. 89-115.
- Jacques, M.P. & Tutin, A. (2018). *Lexique transversal et formes discursives des sciences humaines*. Londres: ISTE Editions.
- Ji, Y. & Tutin, A. (2018). Les routines de topicalisation dans les écrits scientifiques en français. *11es journées du réseau Lexicologie, Terminologie, Traduction « Lexique(s) et genre(s) textuel(s) : approches sur corpus »*, Septembre 2018, Grenoble.
- Ji, Y. & Tutin, A. (2019). Les routines métalinguistiques dans les écrits scientifiques en français. *Studii de Lingvistica* 9(2). pp. 177-200.
- Kraif, O. & Tutin, A. (2017). Des motifs séquentiels aux motifs hiérarchiques : l'apport des arbres lexico-syntaxiques récurrents pour le repérage des routines discursives. *Corpus* 2. Consulté le 24 août 2021. URL : <http://journals.openedition.org/corpus/2889>
- Legallois, D. & Tutin, A. (2013). Présentation : vers une extension du domaine de la phraséologie. *Langages* 189. pp. 3-25.
- Saunier, E. (2012). Disons : un impératif de dire ? Remarques sur les propriétés du marqueur et son comportement dans les reformulations. *L'information Grammaticale* 132. Paris: Peeters Publishers. pp.25-35.
- Schneider, S. (2007). *Reduced Parenthetical Clauses as Mitigators. A corpus study of spoken French, Italian and Spanish*. Amsterdam-Philadelphia: John Benjamins.
- Steuckardt, A. (2005). Les marqueurs formés sur dire. In A. Steuckardt & A. Niklas-Salminen (éds), *Les marqueurs de glose*. Aix-en-Provence: Publication de l'Université de Provence. pp. 51-65.

- Steuckardt, A. (2016). À la recherche du consensus : on va dire, on va dire ça, on va dire ça comme ça. In L. Rosier & J.-C. Anscombe (éds), *Histoire de dire. Petit glossaire des marqueurs formés sur le verbe dire*. Berne: Peter Lang. pp. 293-314.
- Steuckardt, A. (2018). Les marqueurs de reformulation formés sur dire : exploration outillée. *Langages* 212. pp. 17-34.
- Tutin, A. (2018). Les expressions polylexicales transdisciplinaires dans les articles de recherche en sciences humaines : retour d'expérience. In M.P. Jacques & A. Tutin (éds), *Lexique transversal et formules discursives des sciences humaines*, Londres: ISTE Editions, p.91-112.
- Urmson, J.O. (1952). Parenthetical verbs. *Mind* 61. pp. 480-496.

La syntaxe au prisme des propriétés lexicales

Dominique Legallois, LATTICE Université Sorbonne Nouvelle

L'objectif de cette contribution s'inscrit dans le volet "pratiques et méthodes de description" du colloque. Nous partons du principe que la syntaxe peut être parfois avantageusement étudiée et décrite à partir de propriétés qui incombent généralement au seul lexique. Une analyse de Claire Blanche-Benveniste (1978) nous éclaire sur ce point ; examinant le comportement de clitiques, notamment les alternances telles que :

1- mon frère lui a appartenu, à cette équipe vs mon frère y a appartenu, à cette équipe

2- ta belle-mère, tu lui survivras bien vs ta belle-mère, tu y survivras bien

C. Blanche-Benveniste a montré que la séquence « *lui* + verbe » exerce un effet individualisant sur le lexique placé en redondance (*cette équipe, ta belle-mère*), alors que « *y* + verbe » exerce un effet désindividualisant sur ce même lexique. Il s'ensuit que, quel que soit le lexique engagé (marqué + humain ou non), c'est la sémantique de la syntaxe qui prime sur celle du lexique. D'où parfois, selon l'auteur, des effets métaphoriques ou de personnification :

3- la tapisserie, va-t-il penser à elle ?

La sémantique de la syntaxe réinterprète, selon le terme de l'auteure, la sémantique lexicale. Aujourd'hui, ce phénomène est connu et travaillé sous le terme de *coercition* (Lauwers et Willems, 2011). Mais la coercition a d'abord été identifiée comme un phénomène lexical (Pustejovsky 1995)¹, comme dans cet exemple repris à Jackendoff (1995) et à Audring et Booij (2016) :

4- a wooden turtle

où on voit que l'adjectif est en conflit avec le nom qui dénote en principe un être animé. Ce conflit est résolu par un opérateur noté par Jackendoff "model/ replica of N" qui permet d'imposer une lecture cohérente et pertinente (la tortue est un jouet). Ce mécanisme de coercition, identifié surtout par les études sur le lexique, est donc également à l'œuvre en syntaxe, comme le montrent les exemples de C. Blanche – Benveniste.

Il existe d'autres phénomènes, qui, de la même façon, concernent notoirement le lexique, mais semblent s'appliquer également à la syntaxe. Citons-en trois :

La métaphore grammaticale (Halliday 1984). Il y a métaphore grammaticale lorsque le locuteur fait le choix d'employer une forme X au lieu d'une forme Y pourtant généralement plus congruente. Par exemple :

5-Beaucoup de truites peuplent cette rivière, au mois de juin

au lieu de :

6-il y a beaucoup de truites dans cette rivière au mois de juin.

Il y a ici « métaphore » pour Halliday puisque la construction existentielle congruente ou "naturelle" n'a pas été employée, alors qu'une forme transitive – a priori non pertinente pour

¹ Non sans rapport bien sûr avec la syntaxe.

exprimer ce type d'expérience – est exploitée pour exprimer une localisation : à l'image d'une métaphore lexicale, lorsqu'un mot est employé "à la place" d'un mot "littéral". Il s'ensuit dans 5 une sorte de déviance grammaticale par lequel un GN objet a une valeur de locatif. Dans le même ordre d'idée, certaines sous-constructions transitives pourraient être des « condensations » devenues conventionnelles de procédés destinés à des effets rhétoriques (mise en présence ou participation fictive de l'interlocuteur à une scène) ; c'est le cas du mouvement fictif dont certaines occurrences (par ex. *le chemin longe la rivière*) pourraient être rapprochées de procédés rhétoriques tels que l'hypotypose.

La tradition discursive. La notion de tradition discursive (Koch 1997) est issue de la romanistique allemande. Il s'agit de concevoir que beaucoup de formes lexico-grammaticales sont déterminées par un ensemble de contraintes comme le genre discursif, des sociolectes particuliers, des manières de dire propres à une communauté, des formations discursives, etc., bref, des règles et des normes discursives. Le lexique est assurément "pris" dans des traditions discursives ; par exemple, la locution *suite à* (Legallois et François à par.) est fortement liée au discours administratif ou commercial comme dans

7- Bonjour, suite à notre conversation téléphonique je souhaiterais réserver pour la nuit du 1 au 2 août. (Tripadvisor site)

La notion de tradition discursive intéresse aussi l'étude syntaxique ; par exemple, Condamines considère comme émergentes des formes employées par des passionnés de pêches :

8- Difficile de pêcher une rivière que l'on connaît mal (Condamines, 2013)

Or, il n'en est rien ; la forme *pêcher* + Objet Locatif est présente au 19^e siècle dans les textes législatifs sur la pêche ou encore les livres techniques sur le même sujet :

9- On ouvre la bonde A, pour la laisser se remplir de nouveau ; ainsi on pêche l'étang par éclusées (*Pisciculture ou L'Art de la pêche à la ligne*, Jean Cussac, 1816).

La tradition maintient ainsi des emplois que le syntacticien aurait tort de considérer comme une évolution ou une émergence.

La greffe. Les collocations ou unités phraséologiques peuvent faire l'objet, en production, de télescopage ; par exemple,

10- chaque chose prend son temps (entendu à l'oral)

est construit involontairement sur une *greffe* (Legallois 2012) de deux expressions plus ou moins figées – donc enregistrées par la norme d'usage : *chaque chose en son temps* / *toute chose prend du temps*. Le lexique connaît également ce type de greffe, mais il s'agit alors d'un processus général de formation lexicale (le plus souvent intentionnel cette fois-ci) : *photocopillage*, *emmenotter*, *bureaucratie*, etc. Qu'en est-il pour la syntaxe ? Une phrase comme

11- La journée s'annonce pour être la pire de l'histoire.

dont la forme "fautive", pourtant très répandue, est le résultat de ce mécanisme de greffe :

la journée s'annonce la pire de l'histoire et la journée est partie pour être la pire de l'histoire.

Tous ces exemples montrent que la syntaxe peut être affectée par des processus qui concernent également le lexique. Mais il faut certainement le dire autrement : il existe des

processus (coercition, métaphore, greffe, etc.) qui sont autant de mécanismes généraux portant aussi bien sur le lexique que sur la syntaxe. La communication a donc pour objectif de discuter de ces phénomènes, de réfléchir à la façon dont ils pourraient être généralisés, et d'examiner la place qu'ils pourraient prendre dans les travaux de syntaxe.

Références

- Audring J. and Booij G. (2016) Cooperation and coercion, in *Linguistics*, 54(4): 617–637.
- Condamines Anne (2013). "Quand la passion autorise la transitivisation d'un circonstanciel de lieu". *Journal of French Language Studies*, 23, 3: 335-356.
- Blanche-Benveniste C. (1978) "À propos des traits sémantiques utilisés en syntaxe : critique du trait '+/- Humain'", *Cahiers de linguistique 8 : Syntaxe et sémantique du français*, pp. 1-15.
- Blanche-Benveniste C., Deulofeu J., Stéfanini J., Van den Eynde K. (1984) *Pronom et syntaxe. L'approche pronominale et son application au français*, Paris, SELAF.
- Corminboeuf G. (2012), 2012 « Des apo koinou aux constructions louches », *Le verbe en verve. Réflexions sur la syntaxe et la sémantique verbales*, Van Peteghem M, Lauwers P., Tobback E., Demol A. & De Wilde L. (éds), Gent, Academia Press, 215-231
- François J. et Legallois D. (à par.) "Construction Grammar and Discourse Tradition", in *Manual of Discourse Traditions in Romance*, in E. Winter-Froemel & A.O. Toledo y Huerta, De Gruyter.
- Gadet F., Leon J, Pêcheux M. (1984), Remarques sur la stabilité d'une construction linguistique, *Linx*, n°10, 23-50
- Halliday, M.A.K. 1985. *Introduction to Functional Grammar*. London: Arnold.
- Jackendoff R. 1997. *The architecture of the language faculty*. Cambridge, MA: MIT Press
- Koch P., 1997, « Diskurstraditionen: zu ihrem sprachtheoretischen Status und ihrer Dynamik », pp. 43-79, in : Frank B., Haye Th., Tophinke D., eds, *Gattungen mittelalterlicher Schriftlichkeit*, Tübingue, G. Narr.
- Lauwers P. & Willems D. (2011), "Coercion : Definition and challenges, current approaches, and new trends", *Linguistics* 49 (6), 1219-1235.
- Legallois D. (2015) «Comment appréhender la déviance argumentale ? Quelques propositions pour une grille d'analyse», in *Pratiques*, 167-168.
- Legallois D. (2013) «Les greffes phraséologiques : quand la syntaxe se compromet» in *Langages* 189, p. 103-120
- Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, MA: MIT Press.

Description et analyses de l'oral : le cas de régulateurs discursifs en français et en japonais

Chiara Manno, MODYCO Université Paris Nanterre

Notre recherche s'inscrit dans le domaine de l'analyse des interactions verbales. Par « interaction verbale », on entend « toute forme de discours produit collectivement par l'action ordonnée et coordonnée de plusieurs interactants » (Kerbrat-Orecchioni, 1998 : 55). L'entrée d'analyse est le comportement de l'allocutaire et plus précisément les signaux verbaux et non verbaux qu'il emploie pour montrer sa participation active dans l'interaction, autrement appelés régulateurs discursifs.

La notion de régulateur discursif a été étudiée par plusieurs linguistes (Allwood et al., 1993 ; Bevacqua, 2011 ; De Gaulmyn, 1987 ; Duncan et Fiske, 1985 ; Edmondson, 1981 ; Kerbrat-Orecchioni, 2006 ; Poggi, 2005 ; Tannen, 1984 ; Traverso, 1999 ; Yngve, 1970). Toutefois, une définition univoque n'a pas été donnée. Puisque dans la société japonaise le niveau de conscience linguistique relatif à ce phénomène est très élevé, Allen (2019) a interviewé 24 locuteurs natifs du japonais dans le but de mieux cerner la nature de ces signaux. En nous appuyant sur cette étude récente et sur les travaux antérieurs, nous utilisons la définition suivante : il s'agit de formes verbales courtes ou de hochements de tête utilisés par l'allocutaire afin de satisfaire des fonctions communicatives de perception, de contact, de compréhension, d'approbation, de politesse et d'empathie.

En tant que spécialistes du français et du japonais, nous partons de l'hypothèse qu'il y a des différences dans la fréquence, les fonctions et les variétés des régulateurs discursifs de ces deux langues. Nous avons formulé cette hypothèse lors de plusieurs séjours de longue durée au Japon et en France. Des chercheurs comme Clancy et al. (1986), Ohama (2006) ou encore Maynard (1987) ont effectué des recherches contrastives entre des langues telles que le japonais et l'anglais et ont montré que les locuteurs japonophones emploient les régulateurs discursifs plus fréquemment que les locuteurs anglophones. Toutefois, aucun travail contrastif entre le français et le japonais n'a été effectué.

À travers l'analyse de deux corpus de conversation spontanée issus de ces deux langues, nous comptons combler cette lacune et dégager les caractéristiques des régulateurs discursifs. Le type d'interaction que nous avons choisi d'étudier est la conversation entre amis, souvent caractérisée par des récits et des séquences explicatives où l'un des interactants raconte un événement et l'autre joue le rôle de récepteur du discours (Traverso, 2013 : 83).

Nous avons enregistré le corpus français auprès de l'université Paris Nanterre : dix étudiants ayant entre 19 et 24 ans ont conversé librement en dyades pendant environ vingt minutes. Pour le recueil du corpus japonais, nous avons utilisé la même méthodologie, en enregistrant dix conversations spontanées lors d'un séjour de recherche à l'Université de Kobe (Japon) en

2019. Les enregistrements ont été effectués à l'aide d'une caméra, dans le but de tenir compte des hochements de tête et des regards des interactants.

Nous avons examiné la fréquence, les fonctions et les variétés des régulateurs discursifs à travers des analyses quantitatives et qualitatives. L'analyse quantitative nous a permis de comprendre que les locuteurs français emploient un nombre de hochements de tête majeur que les locuteurs japonais, lesquels privilégient les régulateurs verbaux. Nous croyons que cette différence est due à une conception différente de la politesse : utiliser un nombre de régulateurs verbaux élevé pourrait être considéré par un locuteur français comme « envahissant » pour le territoire de l'autre. L'analyse qualitative, basée notamment sur l'approche méthodologique de l'Analyse conversationnelle, est actuellement en cours : nous avons saisi les variétés morphosyntaxiques des régulateurs français et japonais, leur place dans l'organisation séquentielle du discours et sommes maintenant en train d'interpréter leurs fonctions dans l'interaction.

Dans cette intervention, nous fournirons des éléments de réflexion sur la structure conversationnelle de la langue française en confrontant cette dernière avec une langue éloignée : nous mettrons à jour l'état de l'art relatif aux régulateurs discursifs du français et expliquerons leurs fonctions à l'aide d'une comparaison interculturelle. Les résultats de ces analyses pourront constituer des pistes pour l'élaboration de supports didactiques dans l'enseignement des stratégies conversationnelles orales en français langue étrangère.

Bibliographie

Allen, T.J. (2019). « “It has the ability to make the other person feel comfortable”: L1 Japanese speakers' folk descriptions of aizuchi », *Lingua*, 230, 1-20.

Allwood J., Nivre J., Ahlsen E. (1993). « On the semantics and pragmatics of the linguistic feedbacks », *Semantics*, 9(1), 1-26.

Clancy P., Thompson S., Suzuki R., Tao H. (1996). « The conversational use of reactive tokens in English, Japanese, and Mandarin », *Journal of Pragmatics*, 26, 355-387.

Kerbrat-Orecchioni, C. (1990). *Les interactions verbales*, Tome I, Paris, Armand Colin.

Kerbrat-Orecchioni, C. (1998). « La notion d'interaction en linguistique : origine, apports, bilan », *Langue Française*, 117, 51-67.

Maynard, S. (1987). « On back-channel behavior in English and Japanese conversations », *Linguistics*, 24 (6), 1079-1108.

Ōhama, R. (2006). *Nihongo kaiwa ni okeru tĕn kōtai to aizuchi ni kan suru kenkyū* (Étude sur les tours de parole et les aizuchi), Hiroshima, Keisuisha.

Traverso, V. (2006). « Repères pour la comparaison d'interactions dans une perspective interculturelle », *Les Carnets du Cediscor*, 9, 19-55.

Traverso, V. (2013). *L'analyse des conversations*. Paris, Armand Collin.

Yngve, V. (1970). « On getting a word in edgewise », Papers from the Sixth regional Meeting of the Chicago Linguistic Society, 567–577.

De la naissance du syntagme au groupe accentuel et à la structure prosodique

Philippe Martin, LLF Université de Paris

0. L'article de Claire Blanche-Benveniste (2003) décrivant la naissance du syntagme à partir des « scories » observées dans le parler spontané, s'est révélé déterminant pour la description phonologique de l'intonation de la phrase. Les « scories » de la parole spontanée, hésitations, répétitions, abandons, reformulations et surtout reprises, constituent autant d'indices observables pour assigner aux groupes accentuels le rôle d'unité non seulement prosodique mais aussi lexicale, complétant le lexique des mots orthographiques.

Le groupe accentuel est alors défini comme une séquence de mots comprenant une seule syllabe accentuée habituellement en position finale de groupe en français. Cette accentuation se différencie de l'accent d'insistance, qui résulte d'un choix du locuteur, alors que l'accent du groupe accentuel résulte d'une contrainte rythmique incontournable. De cette définition du groupe accentuel comme unité minimale prosodique découlent les quelques observations qui suivent.

1. Lorsque la prononciation d'un groupe accentuel est interrompue avant son achèvement, c'est-à-dire avant sa syllabe accentuée finale, le locuteur reprend (presque) toujours l'entièreté du groupe accentuel depuis sa première syllabe : *c'est pas logi c'est pas du logiciel*. Du reste, en l'absence de syllabe accentuée finale, le groupe accentuel interrompu et donc incomplet n'est généralement pas détecté en tant qu'unité lexicale ou syntaxique par l'auditeur (effet de la « gomme prosodique », Martin, 2018). La reprise ou la reformulation par un groupe accentuel complet suggère que ces unités sont mémorisées telles quelles et constituent de véritables unités semblables aux mots du lexique. Ce n'est que dans des conditions très particulières qu'on observe la complétion d'un groupe accentuel interrompu, pour simuler une hésitation. C'est du reste l'absence quasi systématique de productions de scories avec reprises des groupes accentuels qui rend peu crédible l'interprétation de dialogues de parole censée être spontanée dans les représentations théâtrales ou filmées.

2. Les mots appartenant à un groupe accentuel ne sont pas nécessairement des mots dits « du contenu » ou mots lexicaux (verbes, noms, adjectifs et adverbes, par opposition aux mots grammaticaux tels que prépositions, pronoms, conjonctions), le français étant une langue dépourvue d'accent lexical. On peut aussi bien trouver des séquences comme *mais pas quE* dépourvues de mots lexicaux, que *l'ère glaciEre* avec deux mots lexicaux. On peut aussi rencontrer des mots comportant un grand nombre de syllabes comme *intergouvernementalisation* apparemment impossibles à prononcer avec une seule syllabe accentuée finale, et donc formant deux ou trois groupes accentuels, par exemple : *intergouvernemENTalisatiON*, *intErgouvernementalisatiON* ou *intErgouvernemENTalisatiON* selon le débit syllabique. Les syllabes accentuées supplémentaires, traditionnellement qualifiées d'accents secondaires, sont généralement positionnées sur des frontières morphologiques.

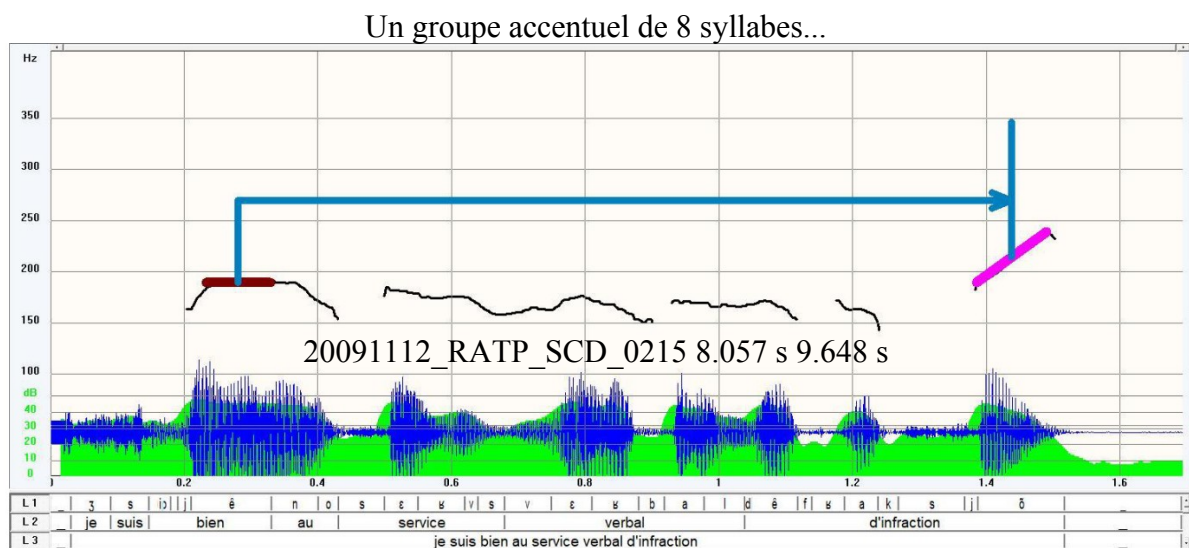
3. Arc accentuel : la première syllabe du premier d'un groupe de deux mots lexicaux est accentuée, comme dans *la politique étrangère*, *un artiste engagé* ou *le beaujolais nouveau*, sans que cette configuration résulte nécessairement d'une collision d'accent. Cette configuration est à distinguer de l'accent dit d'insistance placé également sur la première syllabe accentuée des mots lexicaux *l'exceptionnelle grandeur de Versailles* ou grammaticaux *ni Paul ni Pierre ni Marie n'ont compris ce problème*. L'accent d'insistance de ces deux derniers exemples ne résulte pas d'une collision puisque le deuxième mot impliqué contient plus d'une syllabe.

L'arc accentuel et l'accent d'insistance déterminent par définition un seul groupe accentuel à partir de deux groupes accentuels potentiels, ce qui implique que les deux accents soient

de nature différente, le second devant assurer l'indication partielle de la structure prosodique par un mouvement mélodique déterminé, alors que l'accent initial présente des caractéristiques différentes de l'accent d'insistance, soi-disant toujours réalisé par une montée mélodique.

4. Les réalisations des syllabes accentuées, et donc le phrasé en groupes accentuels, dépend du débit de parole ou de la vitesse de lecture, même silencieuse. Prononcée à débit moyen, *la ville de Paris* est segmentée en deux groupes accentuels : [*la ville*] [*de ParIs*], alors qu'un débit rapide détermine un seul groupe : [*la ville de ParIs*]. Un débit très lent entraîne une segmentation en syllabes détachées, et donc en cinq groupes accentuels : [*lA*] [*vIlle*] [*dE*] [*PA*] [*rIs*]. Ceci interpelle quant à la dépendance de la vitesse de lecture des annotateurs de syllabes accentuées) (cf. les difficultés mentionnées par Avanzi, 2013).

5. On peut montrer qu'un des deux critères de constitution d'un groupe accentuel est en fait sa durée de prononciation, d'où le nom de groupe rythmique. L'autre contrainte, phonologique, porte sur la syllabe finale qui doit être accentuée et qui définit la limite « droite » des groupes accentuels. Les réalisations du « parler jeune » comportent souvent des groupes accentuels contenant jusqu'à 10 syllabes par seconde alors que le débit moyen est de 4 à 5 secondes. On peut montrer que l'intervalle entre deux syllabes accentuées successives (emphatiques ou non) se place dans un intervalle de 250 ms à 1250-1350 ms environ, ce qui suggère un lien de synchronisation avec les ondes cérébrales delta, oscillant entre 0,8 Hz et 4 Hz (Martin, 2018).



6. La perception de deux syllabes accentuées successives dépend de l'intervalle temporel entre les deux syllabes (parole spontanée) :

Je voudrais attirer votre attention sur le fait que la description qu'on en qu'on en fait

400 ms entre les deux syllabes accentuées

Je voudrais attirer votre attention sur le fait que la description qu'on en qu'on en fait

250 ms entre les deux syllabes accentuées

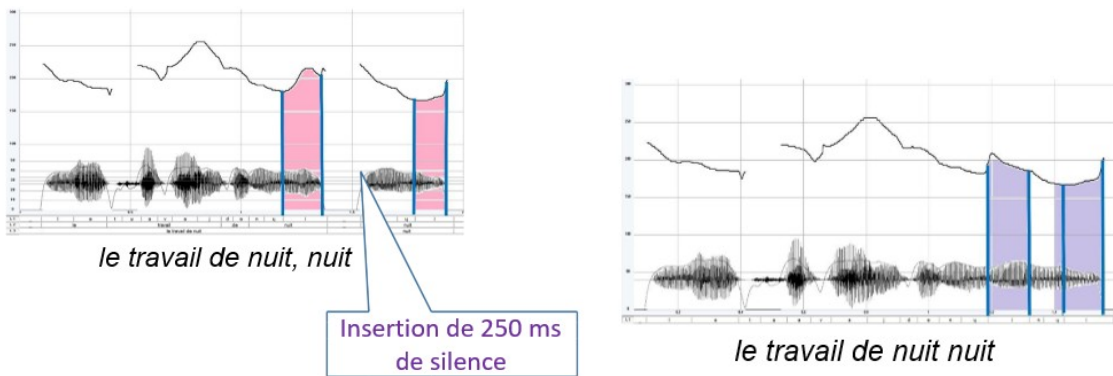
Je voudrais attirer votre attention sur le fait que la description qu'on en qu'on en fait

180 ms entre les deux syllabes : la première n'est plus perçue comme accentuée

Syllabes accentuées successives gap > 250 ms (*dire que*)



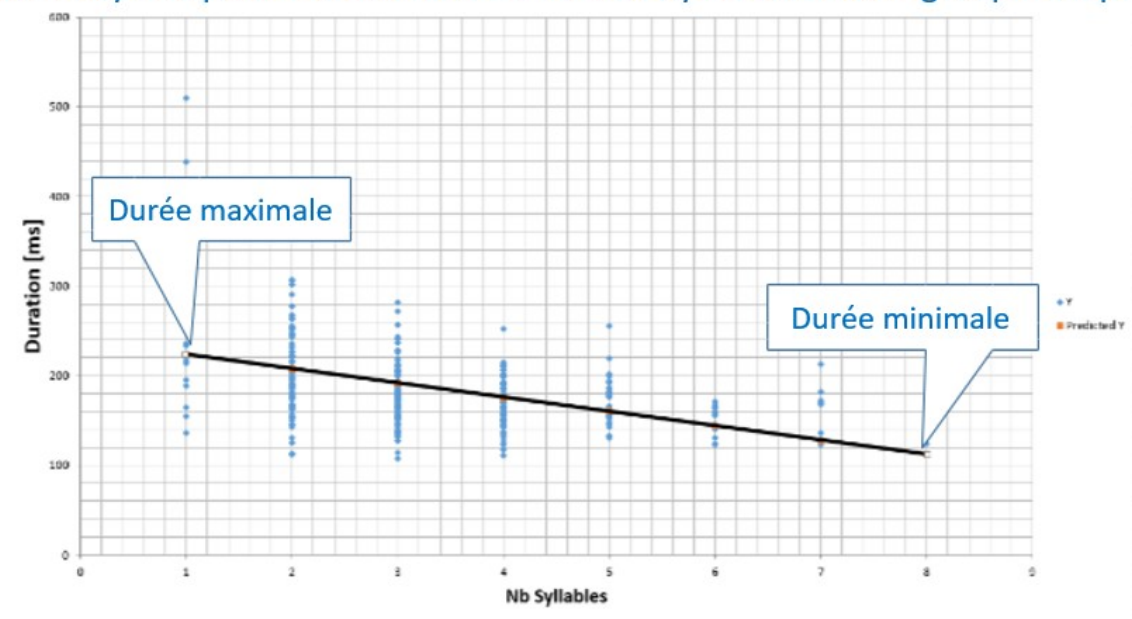
Je suis sur ce sujet en train de vous dire que # j'assume totalement ce qui a été dit...
 Perception des syllabes accentuées successives (parole de synthèse) :



250 ms au minimum sont nécessaires pour percevoir
l'accentuation de deux syllabes accentuées successives

7. Compression syllabique en français selon la vitesse d'élocution

Durée syllabique en fonction du nombre de syllabes dans un groupe temporel



La durée syllabique moyenne décroît avec le nombre de syllabes du groupe accentuel, de manière à préserver la durée maximale de ce groupe.

La durée des groupes accentuels mesurée dans différents styles de parole spontanée varie entre 250 ms et 1250 ms environ.

Un énoncé d'une seule syllabe, nécessairement accentuée, a en réalité une durée d'environ 250 ms, même si la syllabe unique est réduite à une seule voyelle, du fait qu'il faut intégrer la séparation d'avec la syllabe accentuée qui précède (s'il y en a une), nécessairement en position finale du groupe accentuel précédent en français. Cette durée minimale être une limite en dessous laquelle la syllabe cesse d'être perçue comme accentuée.

La durée la plus longue des groupes accentuels est d'environ 1250 ms, ce qui correspond à une moyenne de 7 syllabes. Cela implique que les séquences de plus de 7 syllabes ou plus doivent contenir plus d'une syllabe accentuée, ou doivent comprimer les durées des syllabes (avec une durée minimale de 100 ms).

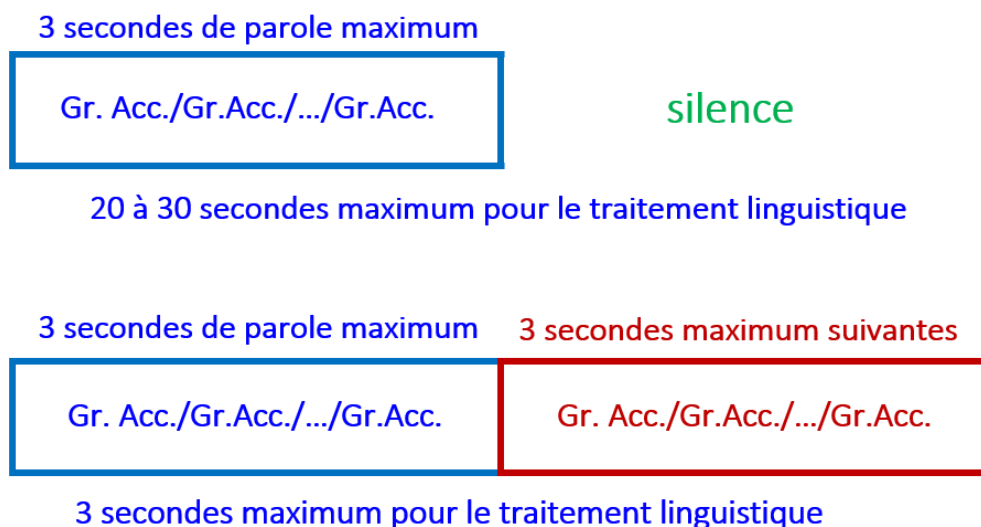
8. La segmentation d'un texte en groupes accentuels est accompagnée de regroupements en syntagmes prosodiques qui ne correspondent pas nécessairement aux unités syntaxiques. Ces regroupements sont indiqués par des marques prosodiques spécifiques placés sur les voyelles des syllabes accentuées. Ces marques, instanciées par des mouvements mélodiques montants et descendants, au-dessus ou en dessous du seuil de glissando, indiquent des relations de dépendance entre les groupes accentuels et déterminent en définitive la structure prosodique de la phrase. Ces contours mélodiques évoquent les continuités majeure et mineure de Delattre (1966).

9. L'écoute ou la lecture silencieuse permet de garder en mémoire à court terme l'image sonore d'une durée maximum de 20 à 30 secondes si la séquence de parole est suivie de silence, mais seulement de 3 secondes en parole continue.

En parole continue, le traitement de l'information linguistique doit donc être effectué au cours de ces 3 secondes.

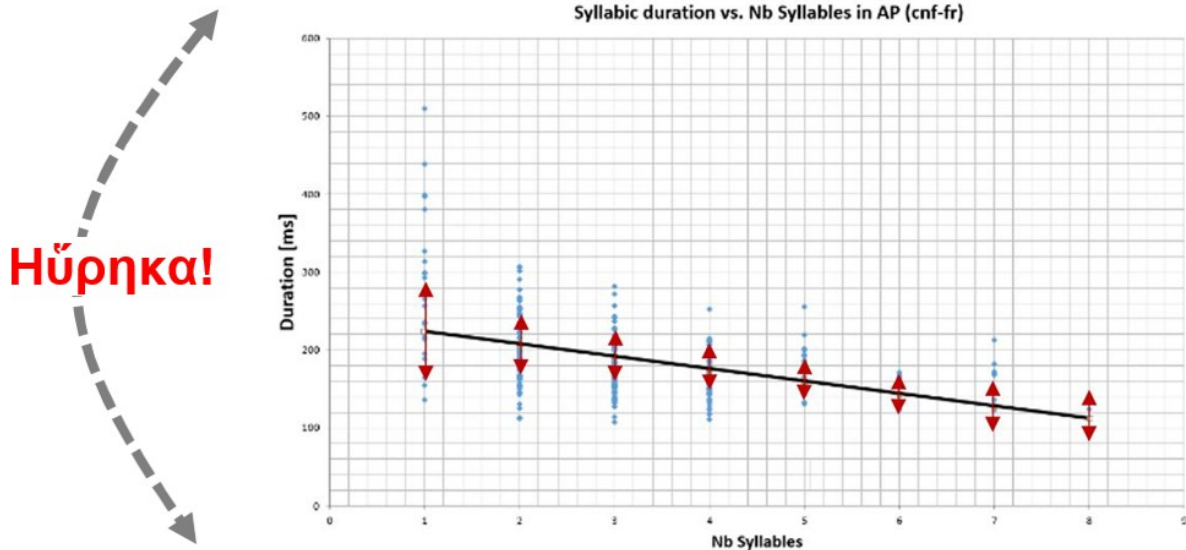
Ce traitement procède par segmentation en blocs de groupes accentuels successifs, dont la durée ne peut dépasser 1250 ms en production orale, même si l'identification des groupes accentuels peut se faire en 250 ms en lecture silencieuse.

Ces regroupements de groupes accentuels sont indiqués par l'intonation, plus précisément par la structure prosodique, qui hiérarchise les séquences de groupes accentuels en plusieurs niveaux.



En l'absence d'une marque prosodique indiquant une fin de séquence de groupes accentuels dans l'espace de 3 secondes, l'image sonore de la séquence sera perdue et le traitement linguistique ne pourra se faire => d'où la nécessité d'avoir une structure prosodique.

10. Les groupes accentuels varient entre 250 ms et 1250 ms

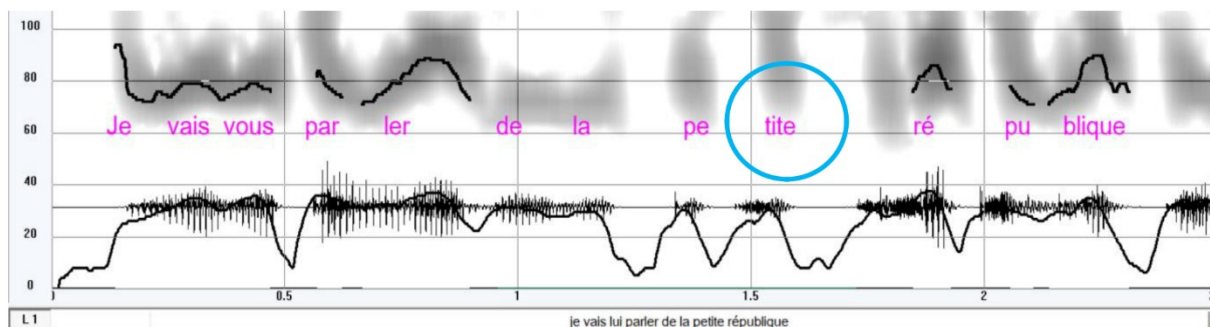


et les ondes cérébrales Delta varient entre 250 ms et 1250 ms !

Écoute à débit variable illustre la synchronisation des ondes cérébrales

20 syllables / s -> moyenne 50 ms	-> 20 Hz	} Θ : 4 Hz à 10 Hz Synchronise thêta
14 syllables / s -> moyenne 70 ms	-> 14 Hz	
6 syllables / s -> moyenne 150 ms	-> 6.6 Hz	
2 syllables / s -> moyenne 300 ms	-> 3.3 Hz	
1 syllabe / s -> moyenne 740 ms	-> 1.3 Hz	} Δ : 1 Hz à 4 Hz Synchronise delta

11. Perception par l'auditeur d'une syllabe accentuée acoustiquement absente



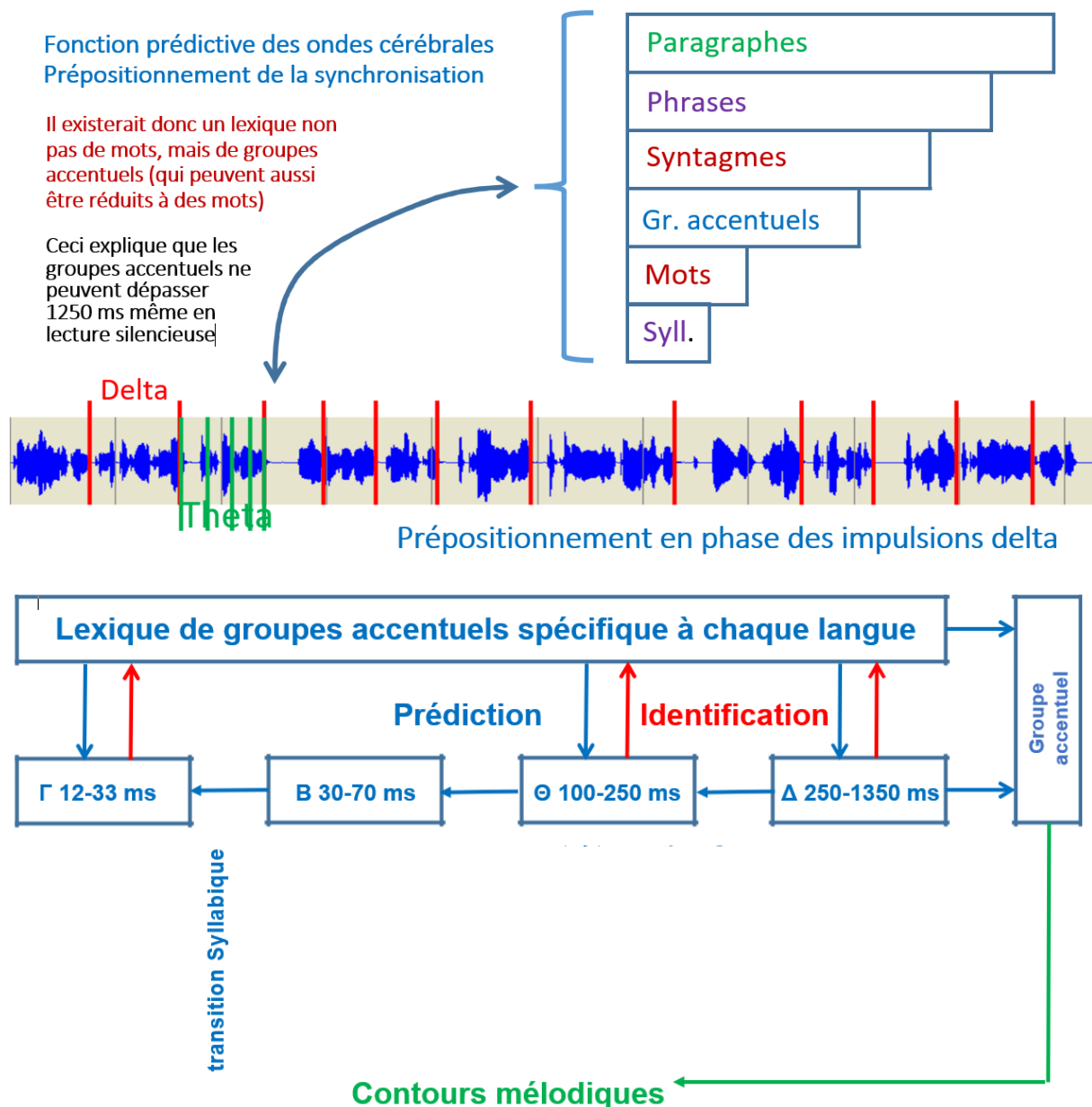
ce qui suggère un processus top-down pour le phrasé.

Ce qui pose question sur les transcriptions des syllables accentuées en français

12. Les unités de perception et de production ne sont pas les mots du dictionnaire, mais les groupes accentuels.

Les locuteurs et les auditeurs disposent d'un lexique de groupes accentuels.

Et qui suggère aussi qu'il existe une micro-syntaxe des groupes accentuels.



Codage de la structure prosodique spécifique à chaque langue

Conclusion

Le fait qu'en lecture orale ou silencieuse on ne peut se passer ni d'une segmentation en groupes accentuels ni de leurs regroupements en structure prosodique, conduit à considérer, et c'est une évidence, qu'en fait le texte écrit n'existe pas comme objet linguistique sans être oralisé, pas plus qu'une partition musicale qui demande un interprète pour exister, même en lecture silencieuse de la partition. C'est la création incontournable d'une intonation de phrase par le locuteur ou le lecteur qui en définitive guide l'auditeur dans le décodage du texte et l'accès au sens. Ceci pose la question de la pertinence des analyses (macro)syntaxiques basées seulement sur la transcription écrite de parole spontanée. La structure prosodique, nécessairement recrée par le lecteur ne peut simplement être réduite à des « coupures prosodiques » insérées dans la transcription orthographique.

Du coup :

1. L'unité minimale phonologique est la syllabe (et non le phonème, concept élaboré sous l'influence de l'écriture alphabétique, et non syllabique).

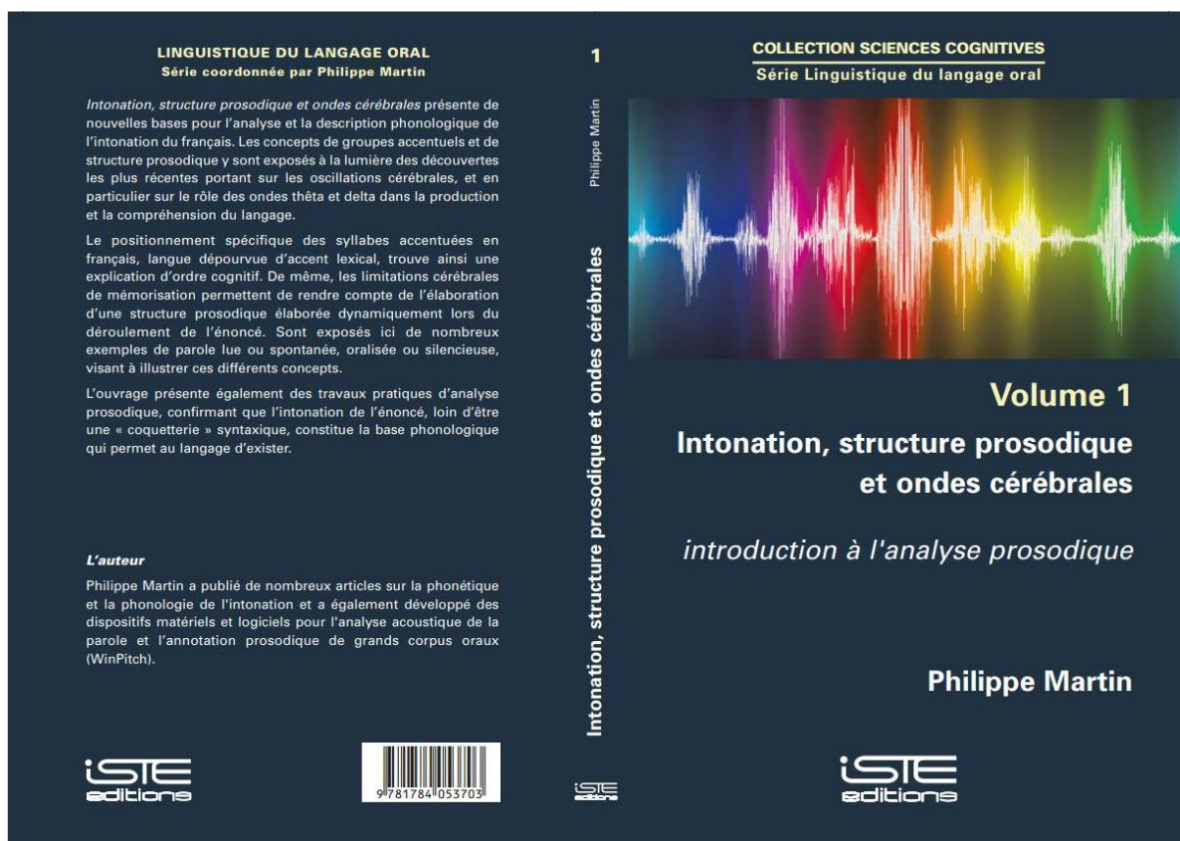
2. L'unité minimale « textuelle » est le groupe accentuel (et non le mot, concept élaboré sous l'influence de l'écriture vers les Xème siècle en Irlande ? Cf. P. Saeger, 1997).

(Expérimentation récente : Hamilton et al., 2021).

Références

- Avanzi, M. (2013). Note de recherche sur l'accentuation et le phrasé à la lumière des corpus du français. *Tranel*, vol. 58, 5-24.
- Bally Charles (1941) Intonation et syntaxe, *Cahiers de Ferdinand de Saussure* 1, 33-42.
- Blanche-Benveniste Claire (2003) La naissance des syntagmes dans les hésitations et répétitions du parler, in Araoui J.L. ed.
- Le sens et la mesure. Hommages à Benoît de Cornulier*, Honoré Champion, Paris, 40-55.
- Debaisieux, J-M., et Ph. Martin (2010) Les parenthèses : étude macrosyntaxique et prosodique sur corpus in *La Parataxe. Entre dépendance et intégration*, M-J. Béguelin, M. Avanzi and G. Corminboeuf ed., Berne : Peter Lang.
- Delattre, Pierre (1966) Les dix intonations de base du français, *French Review* 40, 1-14.
- Fónagy Ivan (1980) L'accent en français : accent probabilitaire, in *L'accent en français contemporain*, Studia Phonetica, 15, Ivan Fónagy et Pierre Léon (eds.), Didier, Paris, 123-233.
- Hamilton et al. (2021) Parallel and distributed encoding of speech across human auditory cortex, *Cell* 184, 1-14.
- Martin, Philippe (2018) *Intonation, structure prosodique et ondes cérébrales*, London : ISTE.
- Saeger Paul (1997) *The Space Between Words: The origins of silent reading*, Stanford University Press.

Si ça vous intéresse...



Syntaxe en n dimensions : choisir et représenter les espaces d'analyse

Nicolas Mazziotta, Université de Liège

Plan

- 1 Diagrammes, réifications et configurations
- 2 Espace syntaxique S et diagrammes fréquents
 - 2.1 Ordre des mots O
 - 2.2 Dépendances D
 - 2.3 Entassements E
 - 2.4 Macrosyntaxe M
- 3 Ce que montrent les diagrammes
 - 3.1 Polysémotique, saillance et efficacité
 - 3.2 Utilisation heuristique des inscriptions conjointes
 - 3.3 Incrire et montrer
- 4 Conclusion

Exemples

- (1) tout chez lui est fait de compassion et non de violence ou de châtement (Rhap-M2003, *apud* Kahane et al. 2019 : 83)
- (2) J'en connais déjà la fin (Mel'čuk, 1988 : 37)
- (3) tout chez lui est fait { de compassion | { et non de violence | ou de châtement } } (Rhap-M2003, *apud* Kahane et al. 2019 : 83)
- (4) donc alors ça date de quand à peu près ce fauteuil-là (Rhap-M2006, *apud* Pietrandrea & Kahane 2019 : 122)
- (5) donc < alors < ça date de quand à peu près > ce fauteuil-là
- (6) – J'irai !
– Et moi avec vous !

Figures

tout chez lui est fait	de compassion
et non	de violence
	ou de châtement

Fig. 1. – Analyse en grille de 194

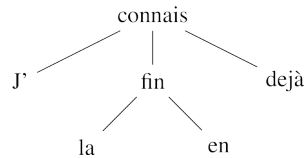


Fig. 2. – Diagramme inscrivant une analyse dépendancielle

$J' \rightarrow en \rightarrow connais \rightarrow déjà \rightarrow la \rightarrow fin$

Fig. 3. – Inscription de l'ordre des mots *O* sous forme de chaîne

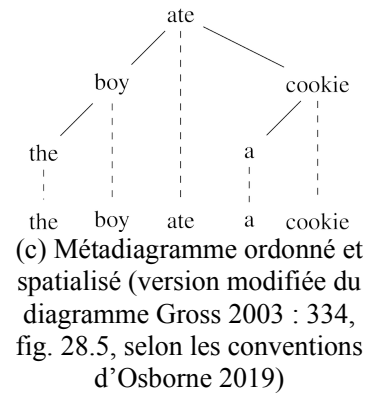
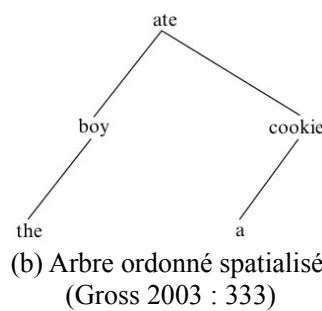
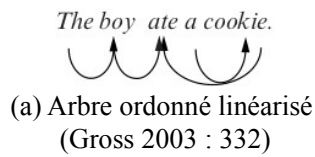


Fig. 4 Incriptions conjointes de *OD*

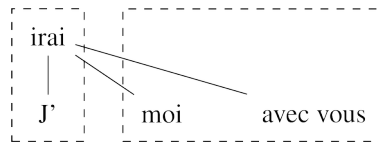


Fig. 5 : Emploi heuristique d'inscriptions conjointes de *DM*

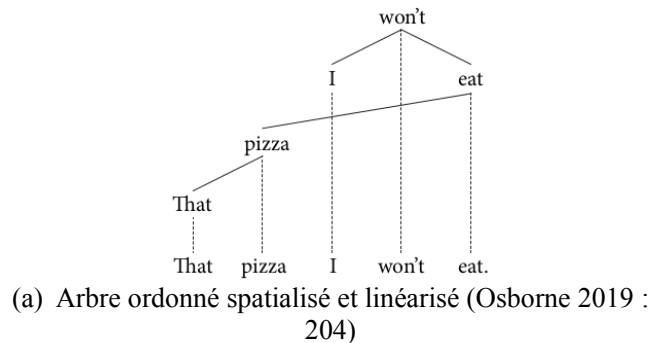
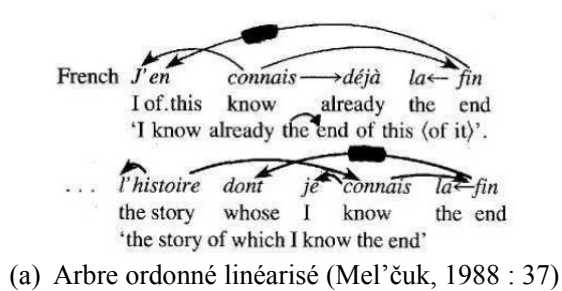


Fig. 6 : Emploi heuristique d'inscriptions conjointes de *OD*

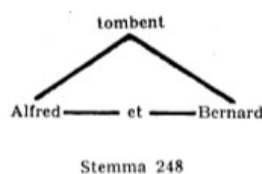


Fig. 7 : Inscription de *DE*
(Tesnière 1959)

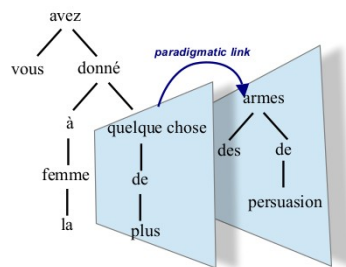


Fig. 8 : Inscription pseudo-tridimensionnelle de *DE*
(Kahane et al. 2019 : 94, dessin original fourni par S. Kahane)

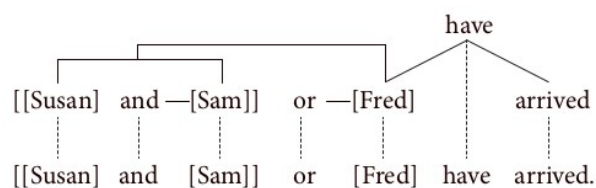


Fig. 9 : Inscription de *ODE*
(Osborne 2019 : 323)

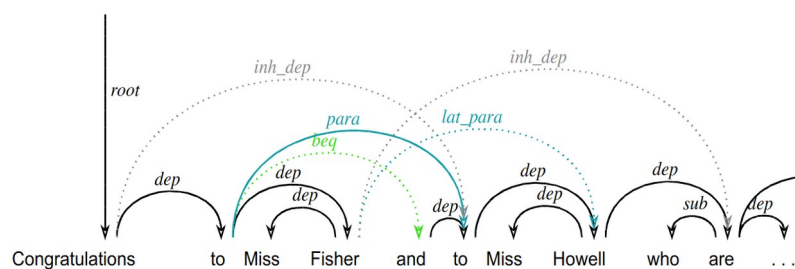


Fig. 10 : Inscription de *ODE* (Gerdes & Kahane 2015)

Citations

Notre premier souci a été d'ordre pratique : fournir une représentation visuelle de ces transcriptions qui permette d'en suivre le développement de façon plus aisée. (Blanche-Benveniste et al. 1979 : 164)

La ligne horizontale que l'on suit dans la lecture habituelle correspond à deux sortes de dispositions : – l'une nous donne la suite concrète des mots [... ;] – l'autre nous donne une suite ordonnée de syntagmes, analysables par référence à une grammaire de langue, [qu'on analyse] comme deux suites de mots appartenant au même emplacement syntaxique. [...] À titre d'élucidation, le GARS a proposé des représentations graphiques qui permettent de visualiser ces phénomènes en exploitant un axe vertical. (Blanche-Benveniste et Jeanjean 1986 : 166-167)

Références

- Bachimont, B. 2007. *Ingénierie des connaissances et des contenus. Le numérique entre ontologies et documents*. Hermès.
Bachimont, B. 2010. *Le sens de la technique : le numérique et le calcul*. Les belles lettres.

- Benzitoun, C. et al. 2010. *tu veux couper là faut dire pourquoi*. Propositions pour une segmentation syntaxique du français parlé, in Neveu, F. et al. (éds). *2^e Congrès Mondial de Linguistique Française*, Institut de linguistique française, 2075-2090.
- Bertin, J. 1967. *Sémiologie graphique. Les diagrammes – les réseaux – les cartes*. EHESS.
- Blanche-Benveniste, C. & Jeanjean, C. 1986. *Le français parlé. Transcription et édition*. Didier.
- Blanche-Benveniste, C. 1997. *Approches de la langue parlée en français*. Ophrys.
- Blanche-Benveniste, C. et al. 1979. Des grilles pour le français parlé, *RSFP*, 2, 163-204.
- Brittain R.C. 1973. *A Critical History of Systems of Sentence Diagramming in English*. University of Texas [Thèse inédite].
- Debaisieux, J.-M. 2007. La distinction entre dépendance grammaticale et dépendance macrosyntaxique comme moyen de résoudre les paradoxes de la subordination, *Faits de Langue*, 28, 119-132.
- Gerdes, K. & Kahane, S. 2009. Speaking in Piles : Paradigmatic Annotation of a French Spoken Corpus, in *Proceedings of the Corpus Linguistics Conference CL2009*, n. 309.
- Gerdes, K. & Kahane. 2015. Non-constituent coordination and other coordinative constructions as dependency graphs, in *Proceedings of the 3rd international conference on Dependency Linguistics (Depling 2015)*, 101-110.
- Gerdes, K. 2006. Sur la non-équivalence des représentations syntaxiques : Comment la représentation en X-barre nous amène au concept du mouvement. *Les Cahiers de Grammaire*, 30, 175-192.
- Groß, T. 2003. Dependency grammar's limits – and ways of extending them, in Ágel, V. et al. (éds.). (2003). *Dependency and valency: An international handbook of contemporary research Vol. 1*, de Gruyter, 331-351.
- Groupe μ . 1992. *Traité du signe visuel*. Paris.
- Hébert, L. 2020. *Cours de sémiotique. Pour une sémiotique applicable*. Paris : Garnier.
- Kahane et al. 2019. The annotation of list structures, in Lacheret et al. 2019 : 69-95.
- Kahane, S. et Mazziotta, N. 2015. Syntactic polygraphs. A formalism extending both constituency and dependency, dans Kuhlman M. et al. *Proceedings of the 14th Meeting on Mathematics of Language (MOL 2015)*, 152-164.
- Kahane, S. 2012. De l'analyse en grille à la modélisation des entassements, in Caddeo, S. et al. (éds). *Penser les langues avec Claire Blanche-Benveniste*, Presses de l'Université de Provence, 101-116.
- Lacheret et al. 2019. *Rhapsodie. A prosodic and syntactic treebank for spoken French*. J. Benjamins.
- Leclercq, B. 2019. Iconic virtues of diagrams. Peirce on ampliative reasoning. *Signata* 10. [En ligne: <http://journals.openedition.org/signata/2286>]
- Mazziotta, N. 2016. *Représenter la connaissance en linguistique. Observations sur l'édition de matériaux et sur l'analyse syntaxique*. Habilitation à diriger des recherches. Université Paris Ouest Nanterre – La Défense, Nanterre, France. [En ligne : <http://hdl.handle.net/2268/204408>]
- Mazziotta, N. 2019a. Énoncés non verbaux en ancien français. Contraintes lexicales, contexte syntaxique et structure textuelle, in Capin, D. et al. (éds). 2019. *Le français en diachronie*. ELiPhi, 121-134.
- Mazziotta, N. 2019b. The evolution of spatial rationales in Tesnière's stemmas. *Proceedings of the Fifth International Conference on Dependency Linguistics (Depling, SyntaxFest 2019)*, 69-80.
- Mel'čuk, I. 1988. *Dependency syntax : theory and practice*. State University of New York.
- Mel'čuk, I. & Milićević, J. 2014. *Introduction à la linguistique*. Volume 3. Paris : Hermann.
- Osborne, T. 2019. *A dependency grammar of English. An introduction and beyond*. J. Benjamins.

- Pietrandrea, P. & Kahane, S. 2019. Macrosyntactic annotation, in Lacheret et al. 2019 : 97-125.
- Stewart A.H. 1976 *Graphic representation of models in linguistic theory*. Bloomington.
- Stjernfelt, F. 2007. *Diagrammatology. An investigation on the borderlines of phenomenology, ontology, and semiotics*. Springer.
- Tesnière, L. ²1966, 1959. *Éléments de syntaxe structurale*. Klincksieck.

Syntaxe de l'oral et activité de construction en temps réel : retour sur les grilles

Katja Ploog, LLL Université d'Orléans

1. Les grilles, plus qu'un outil pour la lecture des textes oraux

En 2000, et après avoir travaillé le problème dans plusieurs publications, PA pose un problème théorique central, et, d'après lui, généralement négligé en linguistique : la dimension temporelle inhérente au langage oral. Prendre au sérieux cette dimension dans la théorie et l'analyse linguistique, permettrait de combler une lacune dans les travaux linguistiques d'alors (2000). PA trouve son inspiration chez Gumperz ou Selting, les travaux de Chafe (1979) et de Silverstein (1984) :

We should identify surface forms of what we might call sentence-scope, as these unfold in realtime discourse. (Silverstein 1984:182, in PA_2376-43)

Les travaux de CBB ne figurent pas parmi les références mentionnées ici. Le point de départ de CBB - vouloir saisir le parlé pour ce qu'il est, réellement, foncièrement - ne semble pourtant pas si éloigné de celui de PA.

Dans plusieurs de ses publications, le germaniste Peter Auer réfère cependant aux travaux de Claire Blanche-Benveniste, en mettant au profit de sa réflexion la représentation en grilles des configurations syntaxiques de l'oral. La fragilité de cette jonction (toujours en note de bas de page) s'explique aisément par les traditions scientifiques fort différentes desquelles sont issus CBB et PA - la linguistique française pour la première et la linguistique interactionnelle allemande pour le second, bien que celui-ci inscrit son approche dans la sémiotique de Peirce. Une rencontre épistémologique entre ces deux sommités de la recherche sur la langue parlée semblait pourtant évidente au vu du dénominateur commun majeur de leurs approches respectives : la prise en compte résolue de la linéarité du déploiement de l'oral dans la description syntaxique du parlé.

PA rend explicite la notion de temporalité de la langue parlée du point de vue cognitif et interactif. Le point de départ est le constat que la construction du sens à l'oral est, de manière essentielle, conditionnée par la mémoire : l'évanescence de la substance sonore conditionne tout d'abord l'action verbale synchrone de locuteur et auditeur, auxquels elle pose des défis cognitifs spécifiques dans le processus de production/réception ; la substance sonore conditionne par ailleurs le caractère irréversible de toute production orale ; la conjonction des deux caractéristique a pour conséquence majeure que les efforts de planification sont audibles (*i.e.* visibles).

La démarche heuristique de PA est donc animée par des considérations psycholinguistiques relatives au processus de production, tandis que celle de CBB est guidée par les problèmes pratiques de la description. Or CBB recourt elle aussi aux arguments psycholinguistiques pour argumenter sa démarche d'interprétation structurelle du phénomène, indissociable de l'acte producteur (en discours). La réflexion des deux auteurs converge dans le souci de respecter, rigoureusement, les caractéristiques séquentielles de la production orale en mettant de côté les considérations externes, psychologiques ou sociologiques, bien souvent spéculatives, seul le mécanisme syntaxique intéresse les auteurs. Chemin faisant, autant CBB que PA développent de nouvelles méthodes d'analyse syntaxique, forgées sur mesure pour l'oral.

En toute logique fallait-il commencer par neutraliser les catégories descriptives forgées pour l'écrit, en se basant résolument sur les données elles-mêmes. Les données de corpus collectées devaient être stabilisées par une transcription systématique.

Une fois les enregistrements de français parlé transcrits, linéairement et sans ponctuation, une présentation reste à trouver pour les rendre lisible. (CBB_513-163)

La présentation linéaire, calquée sur l'écrit, dépeint en effet un caractère "désordonné" de la structuration (le cas échéant excusé, presque dans le même souffle, par le direct ou expliqué par la réduction de la structuration à sa dimension morphologique et syntaxique, en l'absence dans la transcription d'informations prosodiques, vocales, multimodales). Présenter les phénomènes oraux dans la modalité graphique était donc une gageure pour "parvenir à une analyse formelle" (CBB_513-165) et ne pas courir le risque que les problèmes de lecture de la transcription induisent des interprétations structurelles erronées :

Les textes ne donnent pas toujours d'un seul tenant les séquences (grammaticales) complètes : ils présentent des fragments de séquences qu'il faut, pour la lisibilité, rapporter aux ensembles auxquels ils appartiennent. (CBB_513-166).

Le principe de la grille consistait à représenter dans deux dimensions distinctes la séquence grammaticale (sur l'axe horizontal) et l'ensemble des fragments effectivement produits (sur l'axe vertical). Les fragments sont repérés et notés dans la grille par rapport à la séquence grammaticale à laquelle ils sont associés. Les fragments appartenant au même type syntaxique sont alignés sur l'axe vertical. La séquence grammaticale maximale n'est typiquement pas réalisée en premier, voire, n'est produite d'un seul tenant à aucun moment.

Dans la grille exemplaire présentée en 1979, les choix interprétatifs sont rendus explicites par les balises suivantes :

- les encadrements désignent des séquences grammaticales ("structures"),
- les tirets entre deux mots marquent la position vide d'un fragment, en symétrie avec un autre fragment similaire,
- les accolades regroupent différents fragments d'une bribe,
- les annotations en majuscules rendent explicite les relations logiques entre structures.

Le formalisme choisi par PA, explicitement inspiré de celui de CBB (1979), recense par les encadrements les assemblages de fragments (concurrents). PA traduit en outre le *processus* de construction syntaxique ("projection" : cf. ci-après section 2) par des flèches et leur complétion (l'accomplissement constructionnel) par des lignes pointillées.

<i>balise graphique</i>	CBB	PA
<i>encadrement</i>	séquence grammaticale	"retraktive syntaktische Operationen" (PA_2376-51) <i>cf. CBB bribe / ensemble de fragments</i>
<i>lignes pointillées</i>		constructions achevées <i>cf. CBB séquence grammaticales</i>
<i>tiret entre deux mots</i>	position d'une séquence grammaticale non occupée dans un fragment donné (en symétrie avec un autre fragment similaire, i.e. rapporté à la même séquence)	
<i>crochets</i>		séquence grammaticale complexe (comportant des projections non appréhendées dans la grille)
<i>accolade</i>	bribe (ensemble de fragments appartenant à la même séquence grammaticale)	syntagme
<i>annotations en majuscules</i>	relations sémantico-logiques entre structures	catégories/types syntaxiques
<i>flèches</i>	-	projection / complétions

Si, quarante ans plus tard, le manque de lisibilité des transcriptions n'est plus à démontrer, les grilles sont le plus souvent réduites à exprimer le grand principe initial dans la pratique actuelle, celui de consigner les deux dimensions séquentielle et structurelle. Or, dès ses débuts, la démarche de notation bi-dimensionnelle soulevait des questions profondes concernant l'interprétation, voire, la perception des structures ainsi exposées. En effet, les grilles ont fait apparaître des régularités jusque là ignorées dans l'organisation en *configurations*. Si les phénomènes découverts ont été attribués au domaine du *discours*, de nombreuses interrogations ont surgi quant à leur articulation avec la *grammaire*. C'est ainsi que la grille a été promu de l'outil de visualisation (de différentes dimensions présentes dans la linéarité) à une méthode puissante pour réfléchir la structuration syntaxique de l'oral.

2. Vecteurs de l'action verbale

La caractéristique majeure du parlé est la progression séquentielle (temporelle) qui défie la linéarité structurelle. CBB explique cette donnée par la superposition dans l'oralité conceptionnelle (*i.e.* en situation de production spontanée) d'énonciation et d'énoncé : l'énoncé est caractérisé par son instanciation. Dès lors, l'explicitation des structures dépend de la compréhension de ce processus d'instanciation. CBB fait le pari que le processus de formulation des séquences par fragments successifs, indissociablement lié à l'activité formulatrice, relève d'une cohérence profonde des choix du locuteur ; concernant l'acte verbal, CBB se contente de constater que la séquence grammaticale maximale n'est pas toujours *donnée* d'un seul tenant dans le texte oral.

Chez PA, l'action verbale se trouve au premier plan. Cette action se réalise, en temps réel, au moyen de l'opération élémentaire de *projection* :

Durch syntaktische Projektionen werden - in der Zeit vorausgreifend - im Rezipienten durch den Sprecher Erwartungen über die weitere Entwicklung syntaktischer Muster hergestellt; es wird eine syntaktische 'Gestalt' eröffnet, die erst durch die Produktion einer mehr oder weniger präzise vorhersagbaren Struktur geschlossen wird. Die Projektion ist dann eingelöst. (PA_2376-47)

Les projections génèrent - en anticipant dans le temps - chez le récipiendaire des attentes concernant le développement de patrons syntaxiques ; une "gestalt" syntaxique est ouverte, qui ne sera close qu'après la production d'une structure plus ou moins précisément prédictible. La projection est alors complétée. (trad. KP)

Sorte de préfiguration mentale d'une trajectoire, la projection est une opération de structuration la plus générale qui soit. La projection syntaxique traduit une hiérarchie (p.ex. les structures réactionnelles du verbe) et se trouve articulée avec des projections d'autres niveaux, prosodique, sémantique. Les contraintes spécifiques de l'oral favorisent les projections à portée limitée, unités procédurales (mémorielles) les plus gérables, qui facilitent la synchronisation des interactants. L'orientation de l'action verbale sur un achèvement (PA : *complétion*) est fortement marquée par l'Analyse conversationnelle, où le concept de *projectivité* (Sacks, Schegloff, Jefferson 1974)⁰, est central pour expliquer la synchronisation de l'action conjointe (Schütz & Luckmann 1975) des participants à l'interaction⁰.

C'est la trajectoire de l'action verbale, orientée, qui est précisément le fil conducteur de l'approche de PA : la dynamique d'élaboration syntaxique en direct procède par les projections successives de constructions plus ou moins ouvertes ou contraintes. Les étapes de complétion successives, qui, dans la perspective CBB, constituent autant de configurations de bribes singulières. Chez PA aussi, la réalisation de plusieurs fragments syntaxiquement redondants est visualisée graphiquement par leur disposition verticale ; mais dans une perspective résolument dynamique, PA les désigne comme *rétractions*, retour à un point d'ancrage en amont de la complétion de la projection, encore accessible dans la mémoire de travail. En conséquence, l'ensemble de l'agencement des fragments peut être considéré comme étant gouverné par le *priming*, mécanisme cognitif destiné à alléger le coût cognitif du processus d'élaboration : l'activité de construction met à profit les données saillantes, dont les projections activées, en les réutilisant.

Guidée par le souci d'adopter la perspective analytique la plus juste (indépendante de l'écrit), CBB cherche à caractériser les configurations du point de vue leur réception ; aussi, son raisonnement ne vise pas à expliciter une grammaire du locuteur (cf. GG), mais de caractériser les structures syntaxiques d'une perspective externe. Pour en cerner le fonctionnement syntaxique⁰ CBB focalise néanmoins explicitement sur ces phénomènes de production et de perception : l'observation des "processus dilatoires" (CBB_514-134) conduit CBB à adopter la perspective du locuteur investi dans la dynamique de production ; la

⁰ La saisie strictement syntaxique de la notion d'*incrémentation*, elle aussi issue de l'Analyse conversationnelle (Schegloff 1996) découle de celle de la projection (+complétion).

⁰ CBB considère la donnée co-constructionnelle comme acquise et n'influençant que marginalement le déroulement syntaxique. Elle la convoque cependant comme argument pour démontrer la cohérence syntaxique des bribes.

⁰ *Vs.* une valeur psychologique incertaine, comme p.ex. la réparation.

terminologie même (*amorces, anticipations*) fait référence à la construction mentale qui domine l'activité de construction en instance :

Les anticipations s'expliquent sans doute par le fait que les locuteurs ont en réserve, avant de la produire, une partie de la forme de l'énoncé. Seleskowitz (1984) explique comment elle a pu vérifier, dans la pratique de la traduction simultanée, que le locuteur avance en ayant en mémoire les sept à huit mots qu'il vient d'énoncer, et, en projet, les sept à huit mots qu'il va énoncer. 'Nous projetons prospectivement, lorsque nous parlons nous-mêmes, les sept à huit mots que nous allons dire, en une sorte de pré-élocution de même durée' (Seleskowitz 1984, p.275). (CBB_514-135)

3. Figures

La méthode de représentation en grilles, désormais standard en syntaxe de l'oral, s'est vite révélée analytique, en conduisant à dégager des caractéristiques irréductibles de la structure orale, saisie en terme de *configurations*⁰. Qualifié de "processus dilatoire" (CBB_514-134), les bribes laissent intacts les structures syntagmatiques. En tant qu'ensemble de fragments textuels appartenant au même "type syntaxique" : CBB_513-168), les bribes suspendent la linéarité (grammaticale) mais ne contreviennent pas à celle-ci. Les *bribes* ont donc un statut syntaxique : ce sont les instances de séquences grammaticales inférieures (*constructions, syntagmes...*), agencées en une construction maximale, *séquence*.

Aussi, les fragments sont alignés selon leur fonction syntaxique, i.e. leur fonction dans un schéma réactionnel : selon le cas, sous un élément de même type ou à la suite lorsque c'"est une séquence grammaticale"... On s'interroge alors sur les critères définitoires, ou, d'un point de vue plus pratique, sur les méthodes d'identification des séquences dans la transcription.

La grille permet de traiter toute suspension de la progression syntaxique, dont, notamment, les listes, dont la relation sémantique avec le premier élément peut être cumulatif, appositif / explicatif, ou encore substitutif. Si l'interprétation précise est parfois délicate, le souci de formulation manifestée par les recherches lexicales peut se trouver explicité, dans certains cas, par des incursions métadiscursives.

CBB note enfin que l'exposition du discours en grille et la constitution d'une séquence maximale est une démarche purement analytique : l'avancement par bribes n'est généralement pas perçu dans le fil de l'échange et la constitution d'une séquence maximale inutile pour les interlocuteurs. Forte des observations au sujet du brouillonnage à l'écrit, CBB avance qu'il s'agit du processus de l'élaboration du discours en temps réel, qui se trouve mis à profit par un dispositif grammatical comme la coordination. Mais CBB ne cherche pas à établir "pourquoi nous cherchons nos mots mais plutôt à la question de savoir comment nous les cherchons" (CBB_514-153) : à bon entendeur.

Les similitudes entre fragments et séquences grammaticales émergents sont descriptibles, dans le modèle des grilles, comme symétries, qui peuvent dès lors s'analyser en différents types. Les symétries sont observées sur le plan des unités supérieures (séquences grammaticales complexes) tout autant que sur celui des bribes.

⁰ En 1979, apparaissent avec des régularités incontournables les bribes, tenues en mémoire, symétries, et différents types de rythmes.

Il est amusant que c'est précisément le fait d'observer une mémoire de construction interne au discours à l'oeuvre - et donc la perspective *rétrospective* - qui est désignée par CBB comme "projection" :

"Sans qu'il s'agisse de coordination au sens classique du terme, on trouve des cas où des syntagmes apparents isolés se laissent ramener à un statut similaire, et s'interprètent par **projection** sur le canevas des relations syntaxiques établi ailleurs dans le texte."
(CBB_513-174)

CBB cite en exemple :

tia pas beaucoup de jeunes à Marseille	comme nous crois-moi	hé
		non non
	aussi corrects	hein

où le fragment *aussi corrects* ne se substitue pas à la séquence grammaticale précédente, ni ne constitue un fragment possible de celle-ci qui serait réalisé à postériori ; c'est le rapprochement global d'une séquence à une séquence précédente qui est qualifié de *projection* ; en terminologie PA, l'on pourrait décrire cette configuration comme *rétraction à ancrage floue*...

Les rétractions relèvent, d'après PA, d'opérations syntaxiques non projectives ("nicht-projektionsrelevante Syntaxoperationen", PA_2376-49), alors que leur relation avec la mémoire structurelle du locuteur est patente. Il resterait alors à déterminer leur statut dans le processus de structuration.

4. La grammaire est dans le discours

Catégories descriptives

La représentation en grilles oblige à raisonner chaque fragment au regard des règles syntaxiques régissant la composition des bribes, i.e. leur association à des constructions ou familles de construction :

"On dira qu'il existe dans le texte une construction grammaticale maximale (celle d'un verbe ou d'une autre catégorie) et qu'il est possible de rencontrer des fragments de cette construction, sans qu'on puisse prévoir ni l'ordre d'apparition ni la distribution des fragments" (CBB_513-167)

a b c d

a c

a c d

a b d

L'approche reconnaît donc deux niveaux de structuration, celui, abstrait, de l'organisation syntaxique et celui, attesté, de la réalisation phonétique.

Le dédoublement de la dimension linéaire soulève la question des catégories descriptives par lesquelles les phénomènes rencontrés sont nommés. Tout comme pour la transcription (qui anéantit la dimension de structuration "verticale"), les catégories descriptives sont conçues par/pour une grammaire externe (seconde), qui contrôle non seulement les discours métalinguistiques bien au-delà du cadre scolaire mais exerce aussi une pression certaine sur les pratiques... Par exemple, au-delà de la grammaire même et parfois plus encore que celle-ci,

la linguistique utilise des expressions spatiales pour désigner le comportement relationnel (p.ex. left/right dislocation).

PA opère principalement avec les catégories standard de la linguistique allemande ; sa perspective est ancrée dans la grammaire des valences, courant puissant chez les Germanistes allemands, en particulier dans les années 70/80. Son opposition naturelle au courant générativiste peut s'expliquer dans la perspective des premières problématiques développées, orientées sur des phénomènes d'usage (*code-switching*), intégrant les domaines d'analyse morphosyntaxique et pragmatique. Si PA utilise finalement très peu les grilles dans ses publications, il n'hésite pas à y reporter, le cas échéant, les catégories standard.

Pour CBB, l'un des avantages des grilles était précisément de pouvoir se passer d'un métalangage élaboré pour caractériser les phénomènes à un stade où cela lui semblait prématuré (càd, lors de l'édition du corpus : avant l'analyse). Elle revendique cependant l'intégration de deux approches linguistiques dans sa démarche, l'une orientée sur l'analyse syntaxique, l'autre sur la description du français parlé :

Il se trouve que j'ai vu converger, au cours de ces dernières années, deux directions de recherche auxquelles je m'étais intéressée au départ comme à deux directions de recherche séparées : une recherche en théorie syntaxique, menée avec Karel van den Eynde, à Louvain [note], et une grande enquête sur le français parlé et les phénomènes de production orale, menée à Aix-en-Provence essentiellement avec Colette Jeanjean. (CBB_514-124)

D'un côté, l'approche pronominale lui permet de décrire les séquences grammaticales en constructions régies et non régies. De l'autre, l'étude des corpus oraux - et les grilles - servent l'argumentation de l'approche pronominale par la mise à jour d'un éventail de dispositifs jusque là jamais décrits.

Les conséquences pour l'étude syntaxique sont importantes. La prétendue prévalence de parataxe est pondérée par l'organisation même de la structure orale en projections successives et imbriquées : les relations à l'oral sont bien plus diverses que l'opposition parataxe/hypotaxe le laisse penser. Chacun de son côté, PA et CBB, ont toutes les raisons d'abandonner la notion de phrase comme unité d'analyse de base, jugée inadaptée à la description des données de français parlé.

Précisément, aux limites de la syntaxe relationnelle, les dispositifs ouvrent vers un domaine d'étude nouveau, la macro-syntaxe. Ce niveau de structuration (en noyau, affixes, etc. : CBB_1990) ne repose plus sur des relations relationnelles mais sur des contraintes prosodiques et discursives.

Si l'idée que les bribes ne sont pas toujours produites en une seule fois (instanciées par un seul fragment) s'est avérée extrêmement fertile pour les recherches, l'articulation théorique de ce concept avec son homologue structurel, la séquence grammaticale, paraît plus problématique. En visant à mettre en exergue deux dimensions de cohésion différentes - verticale, entre fragments, et horizontale, entre bribes - la méthode des grilles reconnaît un rôle central à la grammaire dans la saisie du déploiement discursif des ressources syntaxiques. CBB observe aussi que les configurations (CBB_514-132 : "lieux de bafouillage") concernent essentiellement les structures régies (vs. associées), en particulier sous forme d'amorce de

construction⁰, ce qui laisse penser que le phénomène est étroitement lié aux contraintes structurelles. En "langage PA", les caractéristiques des projections sont conditionnées par les caractéristiques grammaticales spécifiques ; en cela la notion de projection ici promue rejoint celle de la grammaire générative, tout en la dépassant très largement par ailleurs. PA note que les contraintes relatives au traitement *online* de la projection sont d'autant plus fortes que la force intégrative d'une construction est élevée. Dès lors, PA s'engage dans une perspective typologique visant à remettre en perspective les configurations relevées par rapport aux caractéristiques syntaxiques des grammaires sous-jacentes, ce qui conduit à constater précisément des "dépassements" notoires des contraintes positionnelles caractéristiques de la syntaxe standard (écrite), comme p. ex. celle de la règle de la *Verbklammer* de l'allemand.

Ces observations incitent alors à chercher à interpréter le processus de production au-delà de la temporalité discursive immédiate : quel est le statut grammatical des configurations discursives, quel est leur lien avec les séquences grammaticales ?

5. Au-delà de la temporalité discursive

La maxime de n'écarter aucune occurrence ("no token left behind") met en exergue les limites d'une notion du *système* qui exclut la variation *de facto* et relègue au rang des scories les structures non compatibles avec les catégories descriptives canoniques... Dès lors, que faut-il considérer comme *type* syntaxique ? Dans quel mesure les configurations (i.e. la géométrie dans laquelle les fragments se déploient) révèlent-elles des structures ou des dynamiques de structuration ?

Autant CBB que PA utilisent les arguments quantitatifs pour contrer les représentations courantes relatives au fonctionnement de l'oral. La validité structurelle est prouvée par le caractère récurrent (i.e. régulier) de certains phénomènes. A demi-mot, cela revient à reconnaître que la temporalité façonne les structures à l'oral en générant des empreintes structurelles de plus en plus profondes au gré de répétitions (fréquences d'occurrence). Cette saisie gestaltiste des phénomènes est assez explicite chez PA, beaucoup plus discrète chez CBB. Reste que l'approche quantitative des données de corpus permet parfois de comprendre les constructions comme phénomène émergent, ne serait-ce que pour distinguer grammaire première et seconde.

La manipulation des bribes et fragments aux airs de famille et fonctionnellement congruents pose le problème de la sous-détermination structurelle (plus ou moins marquée selon les cas) notoire des constructions jusqu'en contexte. Du point de vue observateur, cette problématique est rendue explicite chez CBB dès la présentation de la méthode des grilles, où les auteur/es qualifient comme fait relativement banal la possibilité d'adopter plusieurs dispositions différentes pour une configuration donnée :

Plusieurs grilles sont possibles pour un même texte, selon que l'on choisit de privilégier telle ou telle régularité. Cela implique plusieurs interprétations, entre lesquelles il n'est pas aisé de trancher. Nous avons été arrêtés à plusieurs reprises par cette difficulté, et nous avons en ce cas produit plus d'une grille, quitte à en adopter une arbitrairement comme la plus satisfaisante." (CBB_513-200)

Dans la perspective dynamique, la sous-détermination est inhérente à la projection dans les phases précoces de l'élaboration :

Wenn wir die These von der Synchronisierung zwischen Sprech- und Rezeptionszeit ernst nehmen, müssen wir annehmen, dass Rezipienten nicht bis zum Ende einer syntaktischen Struktur warten, bevor sie beginnen, sie zu verarbeiten; sie werden also zuerst eine Phase

⁰ Le constat a été corroboré par des études comparatives récentes (Birkner ##).

relativer Prozessierungsunsicherheit durchlaufen, in der mehrere syntaktische Projektionen (und entsprechend mehrere syntaktische Kategorisierungen) der ersten geäußerten Struktur (z.B. des ersten Wortes oder der ersten Phrase) 'im Spiel' sind und gleichzeitig verfolgt werden. (PA_2376-48)

Si nous prenons au sérieux l'idée de la synchronisation entre les temporalités de production et de réception, nous devons supposer que les récepteurs n'attendent pas la fin d'une structure syntaxique pour commencer à la traiter ; ils subiront alors une phase d'insécurité procédurale initiale, pendant laquelle plusieurs projections syntaxiques (et, en corollaire, plusieurs catégorisations) de la première structure énoncée (p.ex. du premier mot ou du premier syntagme) sont en jeu et poursuivies en parallèle. (trad.KP)

PA décrit ce fonctionnement inhérent de l'oral comme étant une problématique syntaxique majeure, peu problématisée dans l'ensemble, et encore moins au-delà de la description de l'oral. De nombreuses observations de phénomènes locaux (de constructions spécifiques, non standard, etc.) prennent appui sur la relation floue entre projection et complétion, comme les contaminations (greffes, coalescences) grammaticales entre différentes constructions (séquences), réalisées dans la même configuration (dans l'exemple qui suit, entre *savoir la raison / déterminer la cause*), les constructions apokoinou (c. à pivot, c. miroir) (ici, *ils ont des appareils / il y a des appareils*, respectivement anté- et postposés) :

ils ont des appareils pour pour vraiment pour savoir d'où ça pour déterminer vraiment la cause il y a des appareils

Les coalescences grammaticales représentent des changements de la construction en cours, phénomènes de grande banalité selon PA :

Eine konsequent zuende gedachte Behandlung der Zeitlichkeit Gesprochener Sprache muss natürlich auch das strukturalistisch gesehen unangenehme, jedoch alltägliche Faktum berücksichtigen, dass sich der syntaktische Status einer bestimmten Konstituente im Verlauf der Erzeugung der syntaktischen Struktur verändern kann. (PA_2376-51)

Un traitement cohérent de la temporalité de la langue parlée doit évidemment prendre en compte le fait, certes désagréable du point de vue structuraliste mais habituel, que le statut syntaxique d'un constituant peut changer au fil de la genèse de la structure syntaxique. (trad. KP)

PA cite pour exemple :

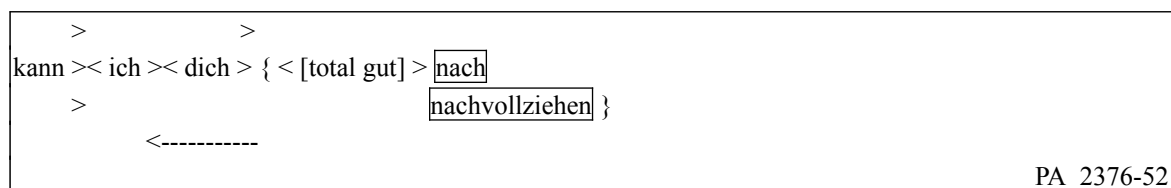
x kann ich dich total gut nachvollziehen

Bien que complétant la projection par *ich* (sujet) d'un verbe fini, *nachvollziehen* était exclu en effet suite à *dich*>, qui projette un prédicat (en position finale) avec la restriction +animé pour le 2nd actant, impossible dans le schéma actancier de *nachvollziehen*. *Verstehen* ne projette pas de telles contraintes.

- ✓ kann ich dich total gut verstehen
- ✓ kann ich total gut nachvollziehen
- ✓ kann ich total gut verstehen

Un verbe trivalent comme *verstehen* était donc attendu, ce qui est exprimé dans la grille suivante par des flèches, qui représentent les projections. La présence de *nachvollziehen*

(bivalent), s'explique comme rétraction (<-----) jusqu'au point d'ancrage *ich*> où la projection est encore sous-déterminée quant à la structure valencielle.



Dans les constructions chevauchées (c. *pivot*, *apokoinou*), Spiegelkonstruktionen chez PA (en référence à Scheutz 1992), l'un des constituants (*koinon*) projette en même temps qu'il complète une projection déjà ouverte :

[ich möchte auch] >> [ein kleines bisschen
ein kleines Stückchen] Daseinsberechtigung] >> möchte ich schon noch haben]

Le rôle joué par les configurations est certainement variable, mais l'omniprésence de ces configurations et leur "structure" interne présentent indéniablement un intérêt pour toute théorie grammaticale qui se voudrait réaliste, basée sur l'usage.

6. Bilan : deux dimensions et deux temporalités

Tout porte à croire que les manifestations typiquement orales méritent à être prises en considération dans l'analyse grammaticale. La complémentarité des approches de PA et de CBB est patente à ce sujet : la force descriptive des grilles pour manifester les bribes fragmentées en référence à la linéarité séquentielle trouve son soubassement théorique dans les opérations de projection et de complétion. La "projection maximale" d'un réseau cohésif de séquences grammaticales serait une configuration discursive, qui pourra être explicitée avec les outils de la macrosyntaxe. Le degré de cohésion d'une séquence grammaticale peut être appréhendé par sa force de projection. Celle-ci peut être appréhendée :

- dans la temporalité courte par les configurations, géométries de fragments constitutifs d'une séquence grammaticale (i.e. la complétion première et les incrémentations rétractives),
- dans la temporalité médiane, par l'observation de configurations discursives complexes, des "dispositifs" (i.e. ensemble de projections de portée variable, intriquées et déployées par séquences successives), dont le caractère gestaltiste peut être argumenté par des fréquences de co-occurrence.

Ces acquis nous incitent à penser une version intégrée dans la perspective d'une *basic linguistic theory* de l'oral, porteuse dans le cadre de la UBCxG ou pour la modélisation syntaxique en TAL.

Références

- Auer, Peter. 2000. Online-Syntax. Oder: Was es bedeuten könnte, die Zeitlichkeit der mündlichen Sprache ernst zu nehmen. *Sprache und Literatur* 85, 43-56
- 2009. Online-Syntax: thoughts on the temporality of spoken language. *Language sciences* 31/1, 1-13.
 - 2013. The temporality of language in interaction: projection and latency. *InLiSt* no. 54.
 - 2014. Syntactic structures and their symbiotic guests. Notes on analepsis from the perspective of on-line syntax. *Pragmatics* 24/3, 533-560

- Biber, Douglas, 1986. Spoken and written textual dimensions in English. resolving the contradictory findings. *Language* 2, 384-416.
- Blanche-Benveniste, Claire. 1983. Examen de la notion de subordination. *Recherches sur le français parlé* 4, 71-115.
- 1987. Syntaxe, choix de lexique et lieux de bafouillage. *DRLAV* 36-37, 123-157.
 - 1990. *Le français parlé. Etudes grammaticales*. Paris : CNRS.
- Blanche-Benveniste, Claire *et al.* 1979. Des grilles pour le français parlé. *Recherches sur le français parlé* 2, 169-206.
- Sinclair, John M. 1991. *Corpus, Concordance, Collocation*. Oxford University Press, Oxford.

Les relatives en *où* en français parlé

Marie-Noëlle Roubaud, LPL Aix-Marseille Université

Frédéric Sabio, LPL Aix-Marseille Université

Comme le relèvent Deulofeu et Valli (2007), les emplois de relatives obliques sont en français parlé d'un usage très limité par rapport aux formes directes exprimées par le complémenteur *qui* ou *que*. Parmi ces formes obliques marginales, celles en *où* apparaissent les mieux attestées : elles sont plus fréquentes que *dont* et, surtout, plus que les formes prépositionnelles du type *auquel* ou *pour qui*.

Nous proposons de décrire les relatives en *où*, en relevant que, si elles ont donné lieu à de minutieuses études de syntaxe et de sémantique (Hadermann, 1993 ; Le Querler, 2000), il manque à ce jour un bilan sur les usages attestés en français parlé contemporain.

Une description de données orales nous paraît d'autant plus nécessaire que le domaine des relatives est, selon divers auteurs, particulièrement sujet à variation selon les genres textuels (Gadet, 1989).

Pour cette étude, nous avons puisé dans la base Orféo (Benzitoun, Debaisieux, Deulofeu, 2016) et sélectionné plusieurs corpus oraux différenciés en genre : réunions de travail, conversations informelles, entretiens...

Notre première tâche a été de séparer les emplois de *où* relatif de ses autres emplois (*où* interrogatifs, emplois locutionnels divers), ce qui permet de constater que les *où* relatifs occupent approximativement les trois quarts des occurrences de *où* dans les corpus examinés.

Notre étude s'organisera en trois points :

1) La forme de l'élément relatif : elle peut apparaître sous les formes non standard « *où que* » (Creissels, 2006, 235) et « *où c'est que* » (Blanche-Benveniste, 2010a, 71) :

(1) mais il y a des choses où que c'est inadmissible parce que les grandes surfaces en plus ils ont eu le droit d'ouvrir à l'heure actuelle les jours de fête les jours où que nous on pourrait travailler un petit peu plus euh les grandes surfaces ils ont droit d'ouvrir aussi (449-450, CRFP)

(2) on a trop de fleurs ça c'est clair mais euh quand on voit que c'est une semaine où c'est que ça bouge beaucoup qu'il y a beaucoup de commandes on va prendre plus que d'habitude (433, CRFP)

Mais ces réalisations ne sont peut-être que des disfluences, comme pour les exemples suivants, dans lesquels les locuteurs produisent une suite d'éléments relatifs différents :

(3) on a évoqué euh une partie qui concerne euh les euh les constructions euh NNAAMMEE euh et et autres dossiers où euh les pour lequel interviennent NNAAMMEE sur les les opérations comme l'ANRU (275, RT)

(4) je suis né en dix-neuf cent treize dans un petit village dans lequel euh où mes parents euh étaient agriculteurs de sorte que j'ai passé ma première jeunesse à la campagne (345, CRFP)

2) La fonction syntaxique relativisée : même si les emplois régis spatiaux et temporels dominent dans le corpus, d'autres usages peuvent être dégagés, dont nous proposerons une analyse dans le cadre méthodologique de l'Approche Pronominale (Blanche-Benveniste et al. 1984). Par exemple un type de rection verbale en « *y – à ça* » doit être identifié pour rendre compte de l'exemple suivant :

(5) on trouve très peu de jeux où on gagne tous hm (61, CLAPI)

(5') on y gagne tous, à ça

D'autres types de constructions sont plus difficiles à caractériser. Nous relevons en particulier de nombreux énoncés qui introduisent un « antécédent » sémantiquement vague, suivi d'une

séquence relative qui apporte une spécification sémantique à cet antécédent. Par exemple, dans :

(6) on évite de faire le système qu'il y a eu il y a il y a six mois où les gens sont partis (66, RT)

Ce qu'il faut comprendre par le mot système est explicité dans la séquence « les gens sont partis » ; on est proche ici de l'effet deux-points qu'avait décrit C. Blanche-Benveniste (2010b) :

(6') Il y a eu un système : (c'est que) les gens sont partis.

On retrouve le même schème syntactico-sémantique avec des lexèmes tels que solutions et processus :

(7) excuse-moi parce que dans les solutions que j'avais envisagées où on rentrait là la salle de bain avait son vitrage je trouve que le lit n'a pas sa place (138, CLAPI)

(8) les accords de Lizarra-Garazi euh avaient été euh étaient la conséquence de euh d'un processus à l'irlandaise où plusieurs partis politiques plusieurs syndicats se sont regroupés autour d'une table (719, CRFP)

Pour ce type d'emploi, il sera particulièrement intéressant d'étudier le type de lexème sous spécifié servant de support à la relative, et d'en évaluer la fréquence.

3) Structures particulières d'énoncés : les relatives en où apparaissent fréquemment dans des énoncés en c'est et il y a :

(9) est-ce qu'il y a des endroits en France où tu aimerais vivre (487, CRFP)

(10) par contre je passe jamais du Véronique Sanson puisque puisque c'est peut-être pas l'endroit où il faut passer du Véronique Sanson je sais pas (702, CRFP)

Il reste à confirmer si la fréquence remarquable de ces structures est spécifique aux relatives en où.

Bibliographie

- Benzitoun C., Debaisieux J.-M. & Deulofeu H.-J. (2016). « Le projet ORFÉO : un corpus d'étude pour le français contemporain », *Corpus*, 15 [https://doi.org/10.4000/corpus.2936]
- Blanche-Benveniste C. (2010a). *Le français. Usages de la langue parlée*. Leuven-Paris : Peters.
- Blanche-Benveniste C. (2010b). « Les pseudo-clivées et l'effet deux points », in M.-J. Béguelin, M. Avanzi & G. Corminboeuf (éds.), *La parataxe*, tome 2, 185-217.
- Blanche-Benveniste C., Deulofeu J., Stéfanini J. & van den Eynde K., (1984). *L'Approche pronominale et son application au français*. Paris : SELAF.
- Creissels D. (2006). *Syntaxe générale, une introduction typologique 2. La phrase*. Paris : Hermès.
- Deulofeu J. (1981). « Perspective linguistique et sociolinguistique dans l'études des relatives en français », *Recherches sur le français parlé*, 3 : 135-193.
- Deulofeu J. & Valli A. (2007). « Sur l'aspect normatif des descriptions linguistiques en français », in G. Siouffi et A. Steuckardt (éds), *Les linguistes et la norme. Aspects normatifs du discours linguistique*, Berne : Peter Lang, 87-110.
- Gadet F. (1989). « La relative non standard saisie par les grammaires », *LINX*, 20-1 : 37-49.
- Hadermann, P. (1993). *Etude morphosyntaxique du mot « où »*. Paris-Louvain-La-Neuve : Duculot.
- Le Querler N. (2000). *Le connecteur « où »*, *Syntaxe et sémantique*, 1 : 113-132.

Grammaire de l'oral vs grammaire de l'écrit ? Le cas des pronoms dits « indéfinis » /+hum/

Catherine Schnedecker, Université de Strasbourg, LiLPa/FRLC

La prise en charge des données de la langue parlée est en train de changer notre perspective générale sur la langue et sur la pratique de description linguistique. (Blanche Benveniste, 2005, 73)

Les grammaires usuelles n'aident pas toujours leur lecteur à se construire une idée précise des usages de certaines formes. C'est en particulier le cas de celui des pronoms dits « indéfinis » dont on sait qu'ils posent des problèmes de classification aussi notoires que persistants et qui sont l'objet de commentaires parfois déroutants ou surprenants dès lors que ceux-ci portent sur la question des usages de l'oral et/ou de l'écrit. Témoin ces quelques jugements, qu'on pourrait multiplier à l'envi :

Autru est un nominal masculin qui appartient **surtout à la langue écrite** (religieuse, juridique, littéraire). La langue courante dit *les autres, un autre*. Il s'agit toujours de personnes (Grevisse, § 713 ; les gras sont de nous)

Nul (caractéristique de **la langue littéraire et juridique**) ... (*Grammaire Méthodique du français*, éd. 2009, 382 ; ; les gras sont de nous)

***Nul* appartient à la langue écrite** ; en parlant, on le remplace par *personne* ou par *aucun* (Grevisse, éd. 1993, § 711 ; ; les gras sont de nous)

Aucun « Au pluriel, précédé de *d'*, dans la langue écrite et soignée, mais aussi dans la langue parlée de certaines régions notamment l'Orléanais ». (Grevisse, § 710 a1)

Un chacun/tout un chacun « Ces locutions ont fait l'objet de jugements contradictoires : « très vulgaires » (M. Schöne, dans le *Fr.mod.*, janv 1947, p. 71) ; langue familière (Martinon, p. 167) ; « ont disparu de la langue écrite » (Dupré) : -de *un chacun* : « encore parfois par plaisanterie » (Foulet, §273) ; - de *tout un chacun* : « archaïque » (Ac. 1987 [qui ignore un chacun] ; « style soutenu » (Dic. Contemp.). A vrai dire, elles apparaissent par écrit dans tous les styles (Grevisse, § 721a)

La forme renforcée *tout un chacun* se rencontre surtout dans l'usage familier (*GMF*, 379)

Or, ces commentaires sont d'autant moins fondés qu'ils

- restent le plus souvent « ponctuels » et partiels, se limitant à quelques items auxquels est prêtée une forme d'« exotisme » totalement dénuée de fondement,
- émanent de la subjectivité de descripteurs qui ne recourent pas toujours aux vérifications empiriques nécessaires, malgré les nombreuses ressources dont nous disposons actuellement. C'est du reste ce que souligne à juste titre Cappeau (2005, 68) :

L'un des reproches que l'on peut adresser à la plupart des grammaires tient au fait que, pour de nombreuses classes grammaticales, les unités sont présentées sous forme d'inventaire sans que soient fournies aux lecteurs des indications sur les faits fréquents ou rares » (Cappeau, 2005, 68)

À quoi s'ajoute le fait qu'ils ne suffisent pas à combler

Le caractère grandement partiel et même lacunaire d'une grammaire du Français, qui ne retient pas des productions orales. (Cappeau, 2021, 11)

À ce point de vue, la description des pronoms dits « indéfinis » par la plupart des grammaires usuelles du français contourne la question de leurs usages à l'oral. Or, certains travaux s'inscrivant dans la description du français parlé et portant sur les variations diamesiques ont mis au jour certaines particularités du fonctionnement de ces formes à l'oral, notamment Cappeau (2018) qui, dans une étude stimulante suggère que :

[...] le système des indéfinis à l'oral n'est pas exactement celui que présentent les grammaires de l'écrit. Un certain remodelage est nécessaire : certaines formes sont délaissées (et elles ne

peuvent donc plus prendre place dans le système) et d'autres formes doivent peut-être être intégrées (Cappeau, 2018, 62)

Partant, notre objectif sera double. Circonscrivant notre propos aux unités et locutions exclusivement pronominales renvoyant à l'humain, pour des raisons de cohérence (cf. Schnedecker, soum.), il s'agira, dans la lignée des travaux de spécialistes de l'oral sur les indéfinis et sur la base d'un double corpus (le CEFC⁰, CERF⁰), de compléter le tableau.

Pour ce faire, nous montrerons les variations de fréquence que subissent ces formes selon leur médium et selon leurs genres d'occurrence. Vu le faible nombre de formes usitées à l'oral dans ce corpus, nous serons amenés à questionner les modalités utilisées à l'oral pour référer à la personne « indéfinie ».

Outre les dispositifs syntaxiques propres à l'oral suppléant ou accompagnant certains indéfinis, que nous rappellerons, nous montrerons également que l'oral i) fait émerger de nouvelles formes ; ii) est responsable de l'extension d'emplois de formes au départ non indéfinies, iii) renouvelle l'usage de formes bien attestées par ce qu'on peut nommer une « régénération » de pronoms nominaux.

Il résultera de ce travail un panorama des indéfinis intégrant les usages oraux qu'occultent généralement les grammaires usuelles, et qui, nous l'espérons, amènera à une meilleure appréhension, plus complète et globale, de cette « catégorie » grammaticale, ainsi qu'à celle du changement linguistique, suivant les principes défendus par Claire Blanche Benveniste :

Je voudrais défendre l'idée que la connaissance de la langue parlée fait progresser la description grammaticale de l'ensemble de la langue. (Blanche Benveniste, 2007, 129)

BIBLIOGRAPHIE

- Bath D.N. S. (2004) *Pronouns*, Oxford, Oxford U.P.
- Benzitoun C., Cappeau P., Corminboeuf G. (2017), « Réflexions sur les exploitations différenciées de la grammaire », *Revue de Sémantique et Pragmatique*, 41-42, 135-153.
- BLANCHE BENVENISTE C. (2003) « *Quelqu'un, quelque chose, quelque part, quelquefois* », *Verbum* XXV/3, 277-288.
- BLANCHE-BENVENISTE C., (1997) *Approches de la langue parlée en français*, Paris, Ophrys.
- BLANCHE-BENVENISTE C., (2007) *Corpus de langue parlée et description grammaticale de la langue, Langage & société* 121-122, 129-141.
- BLANCHE-BENVENISTE C. ET JEANJEAN, C. (1987) *Le français parlé. Transcription & édition*, Paris, Didier.
- BLANCHE-BENVENISTE, Claire (2005). « Le corpus de français parlé du GARS, Groupe Aixois de Recherches en Syntaxe », *Tradizione & Innovazione. Il parlato: teoria - corpora - linguistica dei corpora. Atti del VI Convegno Internazionale della SILFI, Duisburg 28.06.-02.07.2000 (Quaderni della Rassegna 43)*. Firenze: Cesati, 57-75.
- CAPPEAU, P. (2005) : « Les indéfinis à l'aune de l'oral ». *Scolia*, 20, 67-82.
- CAPPEAU, P. (2007) *De la constitution à l'exploitation des données. Réflexions et illustrations*, mémoire en vue de l'HDR, Paris 10, 164p.
- CAPPEAU P. (2018) : (Pro)Noms d'humains à l'oral, *Linx*, 76, 57-83.
- CAPPEAU P. (éd.) (2021) *Une grammaire à l'aune de l'oral ?* Rennes, P.U. Rennes.
- CAPPEAU, P & DEULOFEU, J. (2002), « Partition et topicalisation : il y en a "stabilisateur" de sujets et de topiques indéfinis », *Cahiers de praxématique* n° 37, 45-82.
- CAPPEAU, P & DEULOFEU, J. (2006) Les « indéfinis » en relation avec la position sujet dans trois types de constructions prédicatives, in Corblin *et al.* (éds) *Indéfini et prédication*, Paris, PUS, 125-138.
- CAPPEAU P. & SCHNEDECKER C. 2014 : Des gens bien différents à l'oral et à l'écrit, (Actes du Colloque *Réanalyses*, Neuchâtel, 05-07/09/2012, M.-J.. Béguelin, G. Corminboeuf, L.A. Johnsen org.), en coll. avec P. Cappeau, *Verbum* XXXVI, 1, 55-74.

⁰ Base de données « mixte » 6 millions de mots pour l'écrit/1M de mots pour l'oral
<https://repository.ortolang.fr/api/content/cefc-orfeo/11/documentation/site-orfeo/index.html>

⁰ Le *Corpus Evolutif de Référence du Français* (CERF) a été constitué par Jean Véronis. Constitué de 10 tranches de 1M de mots, il donne accès à l'oral mais aussi à divers genres/domaines (presse, discours institutionnels, littérature, pratique, sciences, forums). Nous remercions très chaleureusement pour son aide aussi précieuse que conséquente Paul Cappeau qui nous a mis à disposition les chiffres et fichiers concernant les formes analysées dans cet article.

- CAPPEAU P. & SCHNEDECKER (2015) *(Les/des) gens vs (les/des) personnes* : évolution diachronique et comparaison diamésique. Des SN en voie de pronominalisation, (Colloque DIA II - Copenhague - 19-21 novembre 2012), in K. Jeppesen Kragh & J. Lindschouw (éds), *Les variations diasystématiques et leurs interdépendances dans les langues romanes*, Strasbourg, Travaux de linguistique romane, pp. 449-463.
- COVENEY AB (2003). 'Anything you can do, tu can do better': tu and vous as substitutes for indefinite on in French. *Journal of Sociolinguistics*, 7(2), 164-191. [DOI](#).
- COVENEY AB (2009). On pour tous, et tu pour on. In Peeters B, Ramière N (Eds.) *Tu et Vous: l'embaras du choix*, Limoges: Editions Lambert-Lucas.
- CREISSELS D. (1995) *Eléments de syntaxe générale*, Paris, PUF.
- FLAUX N. (2008) Les pronoms indéfinis en français : une classe à (re)définir, *Travaux de Linguistique* 56, 7-46.
- GADET F. (dir.) 2017, *Les parlers jeunes dans l'Île-de-France multiculturelle*, Paris et Gap, Ophrys.
- HASPELMATH M. (1997) *Indefinite Pronouns*, Oxford, Clarendon Press.
- JOHNSEN L.A. (2019) [La sous-détermination référentielle et les désignateurs vagues en français contemporain](#). Bern: Peter Lang.
- PREVOST S. & SCHNEDECKER C. (2004) Aucun(e)(s)/d'aucun(e)(s)/ les aucun(e)(s) : évolution du français médiéval au français moderne, *Scolia* 18 : 39-73. En coll avec S. Prévost.
- SCHNEDECKER C. (1999) *Autru* : un pronom vraiment pas comme les autres, *Journal of French Language Studies* 1999/9 : 97-118.
- SCHNEDECKER C. (2002) *Quelqu'un* : la bonne à tout faire des pronoms indéfinis ?, *Verbum* XXIV, 4, 375-398.
- SCHNEDECKER C. (2003) *D'aucuns*, un pronom indéfini « fini » ?, *Scolia* 16, 233-249.
- SCHNEDECKER C. 2003b : *Les uns* : une pluralité « singulière », *Langages* 151, 87-104.
- SCHNEDECKER C. 2003c : *Quelques-uns* « partitif » : approche sémantico-référentielle, *BSLP* XCVIII, 1, 197-227.
- SCHNEDECKER C. 2005a : *Certain* et ses « avatars » (*certain/un certain N ; certains N/de certains N*) : approche diachronique, (Actes du Colloque « Indéfinis et quantificateurs faibles », Bruxelles, 6-8 janvier 2005). *Travaux de linguistique* 50, 131-150.
- SCHNEDECKER C. (2005b) Hypothèses sur l'évolution de certains pronoms en *-un(s)*, (Journée d'étude « La quantification côté déterminants et côté préfixes », Université d'Artois, Arras, 18 mars 2005), *Verbum* XXVII, 4, 331-359.
- SCHNEDECKER C. (2006) Que vise *quelqu'un* ?, in F. Corblin, S. Ferrando, L. Kupferman (éds), *Indéfini et prédication*, Paris, PUPS, 403-415.
- SCHNEDECKER C. (2007) Évolution des pronoms en *-uns* : les destins croisés de *quelqu'un* et *quelques-uns*, in B. Combettes & C. Marchello-Nizia (éds), *Études sur le changement linguistique en français* (actes colloque *DIACHRO* 2, Paris 15-17 janvier 2004), Nancy, PUN, 247-259.
- SCHNEDECKER C. (2011) « Pourquoi *tout le monde*, il ...n'est pas « beau » ou étude sémantico-référentielle de la locution pronominale *tout le monde* », in G. CORMINBOEUF & M.-J BÉGUELIN (éds), *Du système linguistique aux actions langagières, Mélanges en l'honneur d'Alain Berrendonner*, Duculot/De Boeck, pp.507-521.
- SCHNEDECKER C. 2012a : *Tout le monde, tous, (tous) les gens* : Relations sémantiques entre les expressions dénotant la totalité /+hum/ ?, in N. Le Querler, F. Neveu et E. Roussel (éds), *Relations, connexions et dépendances*. Hommage au professeur Claude Guimier. 2012, Rouen, PUR, 123-150.
- SCHNEDECKER C. 2015 : L'enrichissement du paradigme de pronoms indéfinis humains du français ? Étude du processus d'évolution des SN en « gens » du 18^{ième} au 19^{ième} siècle, in A. Carlier, M. Goyens, B. Lamiroy (éds), *Le français en diachronie. Nouveaux objets et méthodes*, (Actes du Colloque *Diachro* 6, Leuven 17-19 octobre 2012), Berne, P. Lang, 247-268.
- SCHNEDECKER C. 2016 Les pronoms dits « indéfinis » référant à la personne humaine : des grammaires à l'usage et réciproquement, *Bulletin de la Société de Linguistique de Paris*, CXI, 1, 341-389.
- SCHNEDECKER C. 2017 « L'expression « indéfinie » de la personne par le SN *dét+quidam* : un cas de transcatégorisation original ? » in B. Fagard, & S. Prévost (éds) *Le français en diachronie*, Berne, P. Lang, 269-290.

SCHNEDECKER C. 2018 »La locution pronominale « tout un chacun » dans le système des pronoms dits indéfinis du français », *Revue de Sémantique et pragmatique* 41-42, 207-221.

SCHNEDECKER C. (2020) « Les pronoms indéfinis /+hum/ du français et les innovations : le cas de *degun(s)* », *Le français innovant*, F. Diemoz, G. Dostie, P. Hadermann, F. Leveuvre, org., actes du Colloque DIA III, 15-16 mars 2018, pp.57-82.

WILLEMS D., MEULLEMAN M. (2010) « Il y a des gens ils viennent acheter des aspirines pour faire de l'eau gazeuse ». Sur les raisons d'être des structures parataxiques en *il y a*, in M-J. Béguelin, M. Avanzi et G. Corminbœuf (éds). *La parataxe. Structures, marquages et exploitations discursives*. Bern. Peter Lang. 167-184.

L'apprenant de FLE, bleu de méthylène de l'état de la description du français parlé ?

Christian Surcouf, Université de Lausanne

Alain Ausoni, Université de Lausanne

1. De certains « rendez-vous manqués » de la linguistique avec les corpus de français parlé

Comme le relève BILGER (2021, 86) « le domaine anglo-saxon est riche de deux descriptions grammaticales de la langue anglaise qui font référence, l'une selon l'approche *corpus based* [...], l'autre selon l'approche *corpus driven* [...], le domaine français apparaît encore déficitaire par l'absence même de cette nécessaire complémentarité descriptive ». CAPPEAU (2021, 14) constate effectivement que « les descriptions grammaticales centrées sur le français parlé restent, pour l'instant, relativement rares ». Un tel constat paraît d'autant plus étonnant que déjà « en 1961, lors de journées d'études organisées à Paris par le BELC⁰, pour discuter des relations entre la "Recherche Universitaire et l'enseignement du français langue étrangère", plusieurs linguistes français de renom poussent des cris d'alarme : on manque d'études sérieuses sur le français parlé et il serait urgent d'en faire » (BLANCHE-BENVENISTE & JEANJEAN 1987, 38).

Soixante ans se sont écoulés, et l'on peut s'interroger sur l'origine des « rendez-vous manqués » par la linguistique à cet égard. En effet, en 1961, les linguistes auraient déjà pu s'intéresser au premier corpus de français parlé, le *Français Élémentaire*, plus tard rebaptisé *Français fondamental (1^{er} degré)* (GOUGENHEIM, MICHÉA, RIVENC & SAUVAGEOT 1964, 13), dont la compilation avait commencé dix ans auparavant (GOUGENHEIM 1955, 407) et à propos duquel ses concepteurs précisait : « le français élémentaire devant être aussi proche que possible des conditions réelles d'utilisation, on a pris pour base le français parlé » (INSTITUT PÉDAGOGIQUE NATIONAL & MARIE 1954, 9). Ainsi rappelaient-ils à l'occasion l'aspect novateur de leur projet, tirant le maximum de profit des dernières avancées technologiques⁰ de l'époque :

Des études de fréquences des mots avaient déjà été effectuées sur les langues écrites. C'est la première fois qu'on a fait une étude analogue sur la langue parlée. Cette innovation a été rendue possible par les perfectionnements du magnétophone. Cent soixante-trois conversations ont été recueillies à l'aide de magnétophones, elles ont été ensuite transcrites et dépouillées intégralement. [...] Pour l'établissement de la grammaire, il a été procédé, sur les conversations enregistrées et transcrites, à un dépouillement des moyens d'expression analogue au dépouillement du vocabulaire. Le total des mots dépouillés s'élève à 312 000. (INSTITUT PÉDAGOGIQUE NATIONAL & MARIE 1954, 9 & 10-11)

Si les préoccupations premières du *Français fondamental* étaient certes didactiques (l'enseignement-apprentissage du français langue étrangère (FLE)), rétrospectivement, il est clair qu'un corpus de 163 conservations intégralement transcrites présentait en soi un intérêt linguistique manifeste. Pourtant, en 1983, SAUVAGEOT – initiateur du projet – rapportait avec amertume : « en fait, le *Français Fondamental* n'a pas eu d'influence immédiate ; les linguistes français ont totalement ignoré l'expérience. [...] L'ouvrage que j'ai publié chez Hachette, *Analyse du français parlé*, en 1972, personne n'en a parlé » (propos cités dans BLANCHE-BENVENISTE & JEANJEAN 1987, 80)⁰. Un tel désintérêt serait-il symptomatique d'une attitude des linguistes vis-à-vis d'initiatives didactiques dont les préoccupations rejoignent pourtant celles de la linguistique ? Il y a près de trente-cinq ans, BLANCHE-BENVENISTE & JEANJEAN (1987, 77) relevaient pourtant que :

⁰ Bureau pour l'enseignement de la langue et de la civilisation françaises à l'étranger.

⁰ « Le magnétophone permet des enregistrements fidèles et commodes (bien qu'en 1951, date à laquelle le travail a été commencé, il ne fût pas aussi perfectionné qu'il l'est à présent) » (GOUGENHEIM 1955, 407).

⁰ Rappelons qu'Émile Benveniste faisait partie du projet et « a apporté à toutes les séances le précieux concours de sa haute autorité en matière linguistique » (GOUGENHEIM *et al.* 1964, 15), même si, comme le remarque CHEVALIER (2006, §4) : « on ne trouve guère trace de ses actions et interventions ».

L'enseignement de la langue aux étrangers a toujours joué un rôle moteur : c'est en prévoyant l'enseignement que l'on perfectionne la description, ou que l'on réclame des perfectionnements à ceux qui sont supposés la décrire. Cela vaut pour le français parlé comme pour les autres langues. (BLANCHE-BENVENISTE & JEANJEAN 1987, 77)

2. De l'apport du prisme didactique pour stimuler le perfectionnement des descriptions linguistiques

Des entreprises menées dans le domaine de la didactique des langues ont souvent contribué à l'avancée des connaissances linguistiques en proposant de nouveaux éclairages théoriques et descriptifs (MCLELLAND 2020). Pour le français, ainsi en est-il de « la première grande grammaire » (CHEVALIER 1996, 13), *Lesclarcissement de la langue francayse* (PALSGRAVE, 1530), dont la finalité était pédagogique : l'enseignement du français à la princesse Mary (STEIN 1997, 4). À la fin du XIX^e siècle, une préoccupation identique était à l'origine de l'élaboration de l'alphabet phonétique international, qui, s'il est devenu un outil descriptif précieux en linguistique, ambitionnait au départ de « représenter toutes les articulations possibles des sons des langues humaines et d'en faciliter l'enseignement et l'apprentissage » (SCHWEITZER, DODANE & LAZAR 2018, 362). Au cours du XX^e siècle, outre *Le français élémentaire* (1954) déjà mentionné, l'enseignement du FLE est également à l'origine d'une autre grande collecte de données langagières orales dès la fin des années 60 : *L'Enquête Socio-Linguistique à Orléans* (ESLO) (BLANC & BIGGS 1971). Déplorant le « retentissement [...] trop maigre » de ces deux initiatives, BLANCHE-BENVENISTE & JEANJEAN (1987, 78) en ont cependant souligné l'intérêt pour la linguistique : « Ces deux entreprises ont fourni des données importantes : enregistrements et transcriptions ; et les gens qui ont observé de près ces données ont montré comment cela permettait de dissiper les vieux mythes sur le français parlé ». En effet, même s'ils poursuivaient des objectifs didactiques, les concepteurs du *Français Fondamental* reconnaissent déjà dès les années 1950 « l'intérêt scientifique que présente l'étude de la langue parlée » (GOUGENHEIM *et al.* 1964, 61). Pour leur part, après la compilation du corpus d'Orléans, BLANC & BIGGS (1971, 24) prédisent que « plusieurs années de recherches linguistiques et sociolinguistiques devront s'écouler avant de pouvoir répondre à toutes les questions que nous nous sommes posées ». Mais malgré l'aspect très novateur de leurs entreprises, ces concepteurs « ne sont pas accueillis comme des pionniers » (BLANCHE-BENVENISTE 2000, 1) et il faudra attendre des années pour voir surgir en France un intérêt *de la part des linguistes* pour les corpus de français parlé.

Si la didactique du français langue étrangère avait très tôt ouvert la voie, la linguistique contemporaine a-t-elle vraiment fini par véritablement prendre en compte la langue parlée ? Dans le préambule de l'une de leurs anecdotes sur l'importance de la dimension parlée dans l'enseignement-apprentissage des langues vivantes, PASSY & RAMBEAU (1901, 17) avertissent déjà qu'« il y a, entre la langue qu'on apprend dans les livres et celle du peuple, des différences parfois très grandes dont nous autres Français nous ne doutons pas parce qu'elles nous sont familières ». En 2010, plus d'un siècle après PASSY & RAMBEAU, BLANCHE-BENVENISTE (2010, 1) y fait écho dès la première ligne de son dernier ouvrage en remarquant que « le français parlé, partagé par des millions de personnes, est encore mal connu par le public et par une partie des linguistes français ». Si l'auteure précise immédiatement que cette méconnaissance résulte du fait que « la grammaire n'en a été pas été décrite dans son ensemble » (2010, 1), une telle situation interpelle cependant. Comment expliquer que le français parlé, bien que « partagé par des millions de personnes » puisse être considéré comme « mal connu par le public », et, de manière plus préoccupante d'un point de vue scientifique, « par une partie des linguistes français », dont la tâche est précisément de décrire la langue et ses fonctionnements ?

Rappelons quelques faits fondamentaux susceptibles d'éclairer certains des rouages du paradoxe de ce constat préoccupant à propos d'« une partie des linguistes français ». Parmi les sciences – et même les sciences humaines –, la linguistique occupe une place singulière en raison du statut même de son objet. Alors qu'un astrophysicien étudie des entités qui lui sont

extérieures, et dont l'existence ne dépend pas de la sienne propre⁰, la langue n'existe que par le fait qu'elle est parlée par des individus. En d'autres termes, le linguiste travaillant sur sa propre langue étudie un objet qu'il a lui-même acquis naturellement et dont il contribue à l'existence sociale. En effet, avant⁰ d'être un scientifique, le linguiste est le locuteur d'au moins une langue, qu'il a acquise naturellement en tant qu'enfant *en la parlant* durant plusieurs années avant d'en apprendre et d'en maîtriser la forme écrite dans le cas où elle comporte un système d'écriture, comme c'est le cas du français. En somme, même s'il s'agit là d'une évidence, tout linguiste est conjointement un scientifique, un être social et un individu. Or tous ces niveaux s'enchevêtrent. Quel que soit l'idéal scientifique auquel aspire le linguiste, ses représentations sociales et individuelles sur la langue peuvent toujours interférer dans l'étude *scientifique* de l'objet-langue. En effet, par la nature même de son parcours scolaire et universitaire, le linguiste incarne par excellence un individu fortement alphabétisé dont les représentations et les réflexions sur la langue ont été conditionnées par l'École et le rôle fondamental qu'y jouent l'écrit et son apprentissage (voir COULMAS 2003, 14) :

l'école exige la mise en œuvre d'un rapport réflexif au langage qui suppose une mise à distance d'un langage-objet, étudié en lui-même et pour lui-même à partir de divers savoirs scripturaux constitués sur la langue (alphabétique-phonologique, lexical, orthographique, grammatical, textuel). (LAHIRE 2008, 12)

Ainsi BLANCHE-BENVENISTE (2010, 1) déplore-t-elle qu'en morphologie, « la description rencontre [...] des obstacles idéologiques » : « les analyses grammaticales ont été pendant si longtemps si fortement ancrées dans l'orthographe qu'il est difficile de s'en émanciper ». Par exemple, à l'issue d'une analyse minutieuse de 11 664 adjectifs du français, SÉGUIN (1973, 54) constate que « 67% sont invariables en genre oral » et confesse qu'un tel résultat « surprend assez par rapport à la définition habituelle de l'adjectif » (voir également CATACH, GRUAZ & DUPREZ 1995/2010, 207). Il est probable que l'origine de cet étonnement provienne de la règle traditionnelle d'accord enseignée dans la grammaire scolaire sur des bases purement orthographiques : « l'adjectif s'accorde en genre et en nombre avec le nom auquel il se rapporte et qu'il qualifie » (BESCHERELLE 1835-1836, 216). Épistémologiquement, que des linguistes, locuteurs natifs de surcroît, en viennent à s'étonner de découvrir des mécanismes langagiers aussi élémentaires et courants que l'accord en genre de l'adjectif semble illustrer la présence effective d'« obstacles idéologiques » (BLANCHE-BENVENISTE 2010, 1) susceptibles d'interférer dans la description objective des fonctionnements langagiers. Le *Cours de linguistique générale* nous avertissait pourtant déjà que « l'écriture voile la vue de la langue : elle n'est pas un vêtement, mais un travestissement » (SAUSSURE (DE) 1916/1994, 51-52). S'il est parfaitement légitime que la linguistique consacre une partie de ses activités à la description des fonctionnements de l'écrit, il est en revanche plus problématique que l'attention qu'on leur accorde en vienne à occulter l'importance, déterminante, de la langue parlée. Pourtant, malgré les efforts constants de BLANCHE-BENVENISTE pour attirer l'attention sur le français parlé pour la description du « français tout court » (BLANCHE-BENVENISTE 1983)⁰, les ouvrages de référence actuels continuent de proposer des descriptions biaisées :

Les descriptions qui se présentent comme générales et que l'on trouve notamment dans les manuels (dictionnaires et grammaires) sont encore celles qui correspondent essentiellement à l'usage de l'écrit littéraire. D'autres usages, comme celui de la Presse écrite, sont rarement signalés dans ces ouvrages, et en ce qui concerne l'oral, toujours dans ces mêmes ouvrages, ce dernier n'est généralement mentionné qu'à l'occasion de commentaires normatifs et de tournures dites « familières », voire « fautives » (BILGER 2021, 85).

Dans le domaine du français langue étrangère, de tels biais peuvent être dommageables. En situation d'apprentissage guidé, les apprenants de FLE puisent naturellement dans les

⁰ La lune, le soleil, Jupiter existent, que l'humain puisse les observer ou non.

⁰ Chronologiquement et socialement.

⁰ Voir notamment ses deux monographies majeures : *Approches de la langue parlée en français* (2000) et *Le français. Usages de la langue parlée* (2010).

descriptions et les documents sonores (pour la plupart fabriqués⁰) qu'ils trouvent dans les ouvrages qu'ils consultent (dictionnaires, grammaires, manuels). En découlent des connaissances parcellaires sur le français, au détriment notamment de la langue parlée. Prenons un exemple provenant d'étudiants étrangers de niveau B2 résidant en Suisse romande. L'exercice consistait à transcrire en orthographe⁰ l'énoncé [2], passé quatre fois avec une pause conséquente entre chaque écoute, après découverte du contexte amont :

[1] Le contexte amont :

je fabriquais des câbles euh aéronautiques que pour les avions mais c'est diversifié tu fais plusieurs choses c'est euh c'est l'usine sans être l'usine aussi tu vois (France Culture, *Les pieds sur Terre*, 14/02/2017)

[2] L'énoncé à transcrire :

[setesɛpaðbyvedekafeəfɛtɛlibɚəkɔwastepadlaʃœ̃] (*c'était sympa on buvait des cafés (euh) enfin tu étais libre quoi c'était pas de la chaîne euh*)

Remarquons que cet énoncé de 22 syllabes est articulé en 4,5'', soit un débit de 4,9 syllabes/seconde, c'est-à-dire plus lentement que la moyenne de 5,9 (pour les femmes) calculée par LÉON (1996, 104). Toutefois, même pour des étudiants de FLE de niveau B2 résidant en Suisse romande, un tel énoncé continue de poser des problèmes de compréhension, engendrés par certaines spécificités peu familières – voire inconnues – du français parlé. Ainsi en est-il de l'utilisation de *enfin* dans son sens « réparateur » (BEECHING 2007, 81), ici plutôt articulé [fɛ̃], l'usage générique de *tu*, élidé de surcroît devant *étais* [tɛtɛ], la présence du « ponctuant » *quoi*, et la réduction de *c'était* à une seule syllabe [stɛ].

Si de telles incompréhensions – récurrentes – peuvent dans le parcours de tout apprenant avoir des origines variées (notamment en partie sous l'influence de la langue maternelle ou du type de données sonores présentées dans l'enseignement), nous voudrions ici les envisager comme indices possibles des lacunes dans les descriptions sur le français parlé à la disposition de l'apprenant. Or, comme l'illustre la schématisation de la Figure 1, ces descriptions résultent d'un cheminement en paliers dans l'élaboration des savoirs sur la langue. Ainsi les savoirs produits en ❶ tendent à influencer ceux des niveaux inférieurs en ❷ et ❸, avec pour conséquence la reproduction des mêmes biais, lacunes ou approximations, susceptibles d'handicaper en aval les apprenants de FLE, qui, contrairement aux francophones natifs, n'ont pas les connaissances intuitives suffisantes pour pallier ces déficiences descriptives.

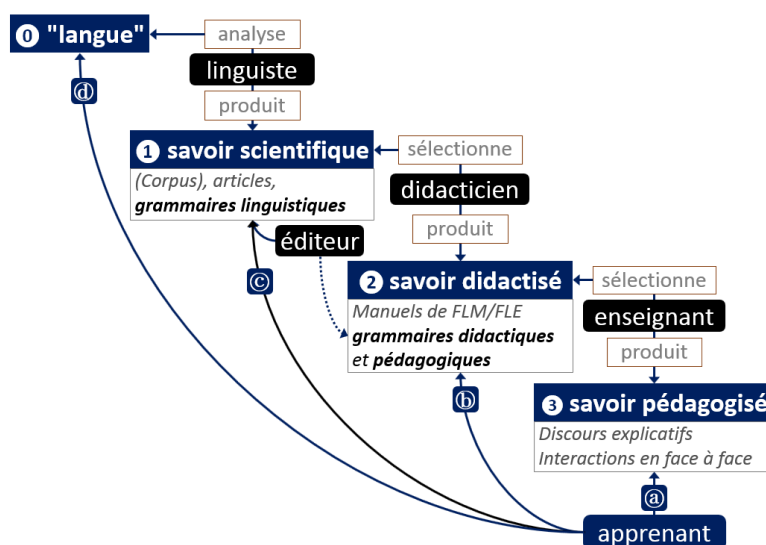


Figure 1 – La diffusion des savoirs sur la langue (d'après SURCOUF & GIROUD 2017)

Dans la première partie de notre communication, c'est sur l'origine possible des biais, lacunes et approximations dans les descriptions de la langue parlée que nous voudrions nous pencher en prenant notre expérience d'enseignement du FLE comme point de départ de notre réflexion

⁰ Voir par exemple SURCOUF & GIROUD (2016), VIALLETON & LEWIS (2014).

⁰ La transcription avait pour finalité d'indiquer le niveau de compréhension fine de l'énoncé par l'étudiant, indépendamment de la correction orthographique.

épistémologique et linguistique. Ce sera l'occasion notamment d'évoquer l'expérience de l'écrivaine hongroise d'expression française Katalin MOLNÁR qui relate ses déboires d'apprenante quand elle découvre avec stupéfaction, à son arrivée en France, l'écart considérable entre le français qu'elle avait appris et celui effectivement parlé par les Français. Ainsi, recourra-t-elle des années plus tard à une écriture phonographique personnelle pour donner à lire certaines spécificités du français parlé, tout en invitant à réfléchir aux raisons pour lesquelles, pour reprendre la citation de BLANCHE-BENVENISTE (2010, 1), « le français parlé, partagé par des millions de personnes, est encore mal connu par le public et par une partie des linguistes français » :

lé fransè saf pa èkspliké koman il parl. Chui dézolé d vout dir mé a lékol, on na apsoluman pa apri ô fransè le fonksionman du fransè parlé é du kou, il saf pa koman il parl. Voila ske chepeû pa suporté. Parlé fransè, sa na rien a voir avèk ékri an fransè, sé tunn lang a par ke tu apran a par, sa pran du tan mé se tanla, sa kont pa parske sèt lang, èl kont pa, le fransè parlé, tu rigol ? Sé môvè, sé vulguèr, sé bouréd fôt. (MOLNÁR 1999, 121)

Partant du principe que les difficultés de nos étudiants en aval révèlent certains des problèmes contenus dans les descriptions du français parlé en amont (voir Figure 1), nous avons entrepris de compiler FLORALE⁰, un corpus de français parlé à visée pédagogique pour sensibiliser aux fonctionnements du français parlé à l'aide d'exemples sonores authentiques consultables et écoutables via une interface. S'est toutefois posée la redoutable question de savoir quels traits langagiers annoter (manuellement), question qui nous occupera dans la deuxième partie de la communication. Où trouver les informations ? Des listes sont-elles disponibles ? Quelles sont-elles ? Comment expliquer la disparité d'un linguiste à l'autre⁰ ? Et quel serait l'étalon sous-jacent à l'élaboration de ces listes ? Nous nous demanderons en somme dans quelle mesure la didactique du FLE pourrait, à nouveau, jeter un éclairage sur les descriptions de la langue parlée et donner aux linguistes un aperçu de certaines des lacunes à combler ou des perfectionnements à apporter (voir BLANCHE-BENVENISTE & JEANJEAN 1987, 77).

RÉFÉRENCES

- AYRES-BENNETT, WENDY (2020), Syntaxe de l'oral, in *Grande grammaire historique du français (volume 2)*, MARCHELLO-NIZIA, CHRISTIANE, *et al.* (Eds), Berlin, De Gruyter Mouton, 1466-1479.
- BEECHING, KATE (2007), La co-variation des marqueurs discursifs *bon, c'est-à-dire, enfin, hein, quand même, quoi et si vous voulez*: une question d'identité ?, *Langue Française* 154, 78-93.
- BESCHERELLE, LOUIS-NICOLAS (1835-1836), *Grammaire nationale*, Paris, L. Bourgeois-Maze, 678 p.
- BILGER, MIREILLE (2021), Sur quelques usages spécifiques à l'oral et à l'écrit?, in *Une grammaire à l'aune de l'oral?*, CAPPEAU, PAUL (Ed.), Rennes, Presses Universitaires de Rennes, 83-95.
- BLANC, MICHEL & BIGGS, PATRICIA (1971), L'enquête socio-linguistique sur le français parlé à Orléans, *Le français dans le monde* 85, 16-25.
- BLANCHE-BENVENISTE, CLAIRE (1983), L'importance du "français parlé" pour la description du "français tout court", *Recherches sur le français parlé* 5, 23-45.
- BLANCHE-BENVENISTE, CLAIRE & JEANJEAN, COLETTE (1987), *Le français parlé. Transcription et édition*, Paris, Didier, 264 p.
- BLANCHE-BENVENISTE, CLAIRE (2000), *Approches de la langue parlée en français*, Gap/Paris, Ophrys, 164 p.
- BLANCHE-BENVENISTE, CLAIRE (2010), *Le français. Usages de la langue parlée*, Leuven/Paris, Peeters, 241 p.
- CAPPEAU, PAUL (2021), Présentation, in *Une grammaire à l'aune de l'oral?*, CAPPEAU, PAUL (Ed.), Rennes, Presses Universitaires de Rennes, 9-19.
- CATACH, NINA ; GRUAZ, CLAUDE & DUPREZ, DANIEL (1995/2010), *L'orthographe française. Traité théorique et pratique avec des travaux d'application et leurs corrigés*, Paris, Nathan, 328 p.
- CHEVALIER, JEAN-CLAUDE (1996), *Histoire de la grammaire française*, Paris, Presses Universitaires de France (Que sais-je?), 128 p.

⁰ L'interface est accessible ici : <https://florale.unil.ch/> (pour plus d'informations sur la constitution du corpus, voir SURCOUF & AUSONI 2018 ; SURCOUF 2020; 2021). Signalons également l'existence d'autres corpus de français parlé à finalité pédagogique qui visent eux aussi à familiariser les apprenants avec certaines caractéristiques de la langue parlée : CLAPI-FLE (<http://clapi.icar.cnrs.fr/FLE/>) et FLEURON (<https://fleuron.atilf.fr/>).

⁰ Récemment AYRES-BENNETT (2020, 1467) constate effectivement que « les listes qui caractérisent la syntaxe orale du français moderne proposées par les chercheurs ne sont pas identiques et peuvent même être parfois assez différentes ».

- CHEVALIER, JEAN-CLAUDE (2006), Le pamphlet : *Français élémentaire? Non*. 1955. L'affrontement Georges Gougenheim – Marcel Cohen, *Documents pour l'histoire du français langue étrangère ou seconde* 36, en ligne, doi: 10.4000/dhfles.1187.
- COULMAS, FLORIAN (2003), *Writing systems: an introduction to their linguistic analysis*, Cambridge/New York, Cambridge University Press, 270 p.
- GOUGENHEIM, GEORGES (1955), Le Français élémentaire. Étude sur une langue de base, *International Review of Education* 1-4, 401-412.
- GOUGENHEIM, GEORGES ; MICHÉA, RENÉ ; RIVENC, PAUL & SAUVAGEOT, AURÉLIEN (1964), *L'élaboration du français fondamental : étude sur l'établissement d'un vocabulaire et d'une grammaire de base*, Paris, Didier, 302 p.
- INSTITUT PÉDAGOGIQUE NATIONAL & MARIE, ANDRÉ (1954), *Le français élémentaire*, Paris, Ministère de l'Éducation Nationale, 67 p.
- LAHIRE, BERNARD (2008), *La raison scolaire : école et pratiques d'écriture, entre savoir et pouvoir*, Rennes, Presses Universitaires de Rennes, 190 p.
- LÉON, PIERRE (1996), *Phonétisme et prononciations du français*, Paris, Armand Colin, 269 p.
- MCLELLAND, NICOLA (2020), The history of language learning and teaching (HoLLT) in the history of linguistics, in *History of linguistics 2017*, AUSSANT, ÉMILIE & FORTIS, JEAN-MICHEL (Eds), Amsterdam, John Benjamins, 189-204.
- MOLNÁR, KATALIN (1999), *Konférans pour lé zilétre*, Romainville, Al Dante, 42 p.
- PASSY, JEAN & RAMBEAU, ADOLPHE (1901), *Chrestomathie française, morceaux choisis de prose et de poésie, avec prononciation figurée à l'usage des étrangers*, New York, Henry Holt, 253 p.
- SAUSSURE (DE), FERDINAND (1916/1994), *Cours de linguistique générale*, Paris, Payot, 520 p.
- SCHWEITZER, CLAUDIA ; DODANE, CHRISTELLE & LAZAR, JAN (2018), L'histoire des alphabets phonétiques du XVIII^e jusqu'à l'API, *XXXIIe Journées d'Études sur la Parole*, Aix-en-Provence, 356-364, doi: 10.21437/jep.2018-41.
- SÉGUIN, HUBERT (1973), Le genre des adjectifs en français, *Langue Française*, 52-74.
- STEIN, GABRIELE (1997), *John Palsgrave as Renaissance linguist : a pioneer in vernacular language description*, Oxford, Clarendon Press, 511 p.
- SURCOUF, CHRISTIAN & GIROUD, ANICK (2016), À quelle langue accède l'apprenant ? Examen critique du traitement de l'oral dans les premières leçons de manuels de français langue étrangère pour débutants, *Linguistik Online* 78-4, 11-27, doi: 10.13092/lo.78.2947.
- SURCOUF, CHRISTIAN & GIROUD, ANICK (2017), L'usage exclusif de l'API dans l'enseignement et l'apprentissage de la morphologie verbale en FLE: un défi insurmontable?, *Bulletin Vals-Asla* 105, 59-76.
- SURCOUF, CHRISTIAN & AUSONI, ALAIN (2018), Création d'un corpus de français parlé à des fins pédagogiques en FLE: la genèse du projet FLORALE, *EDL (Études en didactique des langues)* 31, 71-91.
- SURCOUF, CHRISTIAN (2020), Les enjeux de la compréhension du français oral quotidien en FLE: atouts possibles d'un corpus de français parlé annoté à des fins pédagogiques, *Études de Linguistique Appliquée* 198 – *Les corpus oraux et leur didactisation*, 241-256.
- SURCOUF, CHRISTIAN (2021), Le français oral quotidien, un objectif spécifique en FLE? Retour sur les défis de la création d'un corpus de français parlé annoté à visée pédagogique, in *Des corpus numériques à l'analyse linguistique en langues de spécialité*, FRÉROT, CÉCILE & PECMAN, MOJCA (Eds), Grenoble, UGA Éditions, 107-133.
- VIALLETON, ÉLODIE & LEWIS, TIM (2014), Reconsidering the authenticity of speech in French language teaching: theory, data, methodology, and practice, in *French through Corpora: Ecological and Data-Driven Perspectives in French Language Studies* TYNE, HENRY, et al. (Eds), Newcastle upon Tyne, Cambridge Scholars Publishing, 293-316.

Corpus bilingue multimodal du parler des jeunes étudiants : enjeux méthodologiques et avantages pour l'étude des particularités syntaxico-discursives

Polina Ukhova, LPL Aix-Marseille Université

Introduction

La remarquable avancée technologique de ces dernières années se trouve à l'origine de la fréquence du recours au corpus dans les études menées par les linguistes ainsi que de la diversification des corpus constitués. Aujourd'hui, des dispositifs d'enregistrement et des logiciels de traitement sont accessibles et facilement manipulables. De ce fait, la constitution de corpus autres que textuels préoccupe les chercheurs qui ont pour objectif d'étendre les perspectives d'analyse.

En outre, les nouvelles technologies numériques ont fait émerger de nouvelles pratiques, de nouveaux contextes de communication qui élargissent le champ de l'analyse linguistique. Notons, que si les années 2000 ont vu le triomphe du corpus en linguistique, ce sont les sémanticiens, les lexicologues et les dialectologues qui ont été les premiers à se revendiquer du corpus. Or, aujourd'hui, on peut observer un intérêt croissant pour les phénomènes d'ordre syntaxico-pragmatique. Citons le numéro 9 de la revue dédiée entièrement au sujet : *Corpus* [cf. *Corpus n°9*, 2010, « La syntaxe en corpus » (M. Oliviéri)], sans oublier les études menées par Dekhissi (2013), Moreno (2016), Vigneron-Bosbach (2016), Secova (2017), Cappeau (2017), Savreux (2017) et d'autres.

Les linguistiques de corpus prônent une démarche de recherche à partir des usages discursifs en constante évolution. Outre cet intérêt pour les usages, on remarque également une attention portée aux usagers de la langue. De ce fait, comme le souligne Teubert (2009), un certain nombre de travaux de recherche effectués à partir de corpus se situent à la frontière de la linguistique et de la sociologie. Cela semble d'autant plus justifié avec le développement de corpus vidéo et multimodaux qui permettent de « rendre compte d'un certain nombre de phénomènes co-verbaux analysables en termes sociolinguistiques que les corpus textuels ne sont pas à même de retranscrire » (Azzopardi, 2010, p.209).

Notre étude réalisée entre 2017 et 2019 que nous souhaiterions présenter se situe au croisement de toutes les problématiques que nous venons d'exposer en ce qu'elle est menée à partir de données bilingues multimodales, porte entre autres sur les caractéristiques syntaxico-pragmatiques du parler non-conventionnel et s'intéresse aux pratiques langagières d'un public, à notre connaissance, peu étudiée. Il s'agit d'une analyse comparée de « parlars jeunes » que l'on peut observer en France et en Russie. Les études de ces variantes de l'une et l'autre langue se caractérisent en ceci qu'elles posent souvent la question sociale en arrière-plan à la collecte comme à l'analyse des données : en France par exemple, depuis les années 80, elles associent largement la variation aux trois facteurs que sont la dimension exclusivement orale des productions observées, l'impact du contact des langues sur des productions langagières collectées auprès de populations socialement fragiles, issues de l'immigration et souvent marquées par l'exclusion à la fois du système scolaire et du système économique en général.

Dans notre étude, nous avons fait le choix de baser la collecte des corpus sur d'autres paramètres sociaux : nous nous intéressons à une population de jeunes de 18 à 23 ans inscrits dans des facultés de « lettres et langues » à l'Université de Poitiers et à l'Université de Iaroslavl. Si l'on peut supposer que la variante linguistique utilisée par les étudiants repose sur le fait que cette population *adulcescente* se positionne encore par rapport au « monde des adultes » sur un mode de rébellion identitaire, nous sommes dans un contexte où le manque de maîtrise du code linguistique standard ne devrait pas jouer un rôle déterminant. Un tel projet a bien entendu soulevé les questions interdépendantes de la constitution du corpus et des méthodes d'analyse retenues.

I. Cadre méthodologique : collecte de données & présentation du corpus

Il nous a fallu pouvoir être certaine que le corpus constitué nous fournissait les métadonnées nécessaires à l'identification du profil du locuteur (étudiant de 18 à 23 ans). Nous avons résolument opté pour une approche qualitative portant sur un corpus de travail (cf. Rastier) relativement important, constitué de 3 temps de collecte en interaction les uns avec les autres :

1. **Enquête sociolinguistique (méthode introspective)** auprès de 100 étudiants de Poitiers + 100 étudiants de Iaroslavl, menée afin de constituer un glossaire thématique servant de point d'entrée « filtrant » dans les données recueillies.

On estime que la querelle entre le linguiste introspectif et le linguiste de corpus, n'a pas de raison d'être : il est tout à fait possible de faire collaborer deux méthodologies, elles sont compatibles (cf. Plénat). Si l'introspection peut

repérer certaines variations dans les pratiques langagières, elle est impuissante à décrire leur distribution dans la population : par définition, le social y est impossible à prendre en compte. Inversement, la linguistique de terrain, qui vise la prise en charge des aspects sociaux des variations langagières, privilégie par force certains secteurs de la description linguistique. Les corpus permettent de mettre à jour les divers emplois d'un élément et de préciser pour chaque usage les contraintes en présence. La combinatoire de ces contraintes est cependant complexe et peut dépasser le cadre des seuls corpus. L'introspection peut alors prendre le relais en tentant d'appréhender l'acceptabilité des cas ou situations non présents dans les textes.

Ainsi, afin d'élaborer un premier point d'entrée dans les données, nous nous sommes basée sur la thèse de Tatiana Rétsinskaya (2004), linguiste russe ayant travaillé sur l'argot des jeunes français ; son étude débouche sur une organisation des champs sémantiques liés au monde de la jeunesse. Les voici :

<ul style="list-style-type: none"> - « processus scolaire » (« disciplines / matières », « nom d'un établissement scolaire », « chef d'établissement », « notes », « examen ») ; - « amour / sexe » ; - « loisirs / musique » ; - « argent » ; - « beauté/force physique » ; - « querelle / bagarre » ; - « loyer » ; - « toxicomanie » (alcool / drogue) ; - « fatigue » ; - « dépit/dégoût/colère » -« matériels/équipement informatiques » 	<ul style="list-style-type: none"> - « transport » ; - « inquiétude / stress » ; - « gens » (sexe / milieu social / statut social / nationalité) ; - « crime » (arme / fraude / vol / police) ; - « bêtise / folie » ; - « amitié / discussion entre amis » ; - « transport » ; - « malchance » ; - « nourriture /boissons » ; - « vêtements /chaussures » ; - « maladie/défaut du corps/mort ».
--	---

Tableau 1. : Champs sémantiques (étude de T.Rétsinskaya, 2004)

Nous avons repris ces champs sémantiques afin de procéder à la collecte et à l'élaboration de listes de termes du parler jeune qui ont servi de filtre et de fondement pour l'ultérieure observation de leur réalisation dans les discours spontanés à l'oral et à l'écrit. La consigne était donc de noter des termes du parler des jeunes se rapportant à un champ donné, à condition de n'indiquer que ceux qu'ils emploient eux-mêmes. Une case supplémentaire offrait la possibilité d'ajouter d'autres champs sémantiques avec des items correspondants.

Par la suite, nous avons procédé à la collecte des données spontanées attestées en situation d'interaction. Pour chaque item indiqué dans le questionnaire nous avons trouvé sa réalisation dans les discours authentiques à l'oral et à l'écrit. Cette partie du travail a été ensuite synthétisée sous forme de glossaire.

Nous avons donc constitué **4 sous-corpus** :

1. **2. Corpus oral** : 105h d'enregistrement en français et en russe : conversation en milieu naturel (observation participante dans les universités) + conversation publique lors d'émission de radio identitaires « jeunes » mentionnant l'âge des interlocuteurs.

La principale méthode utilisée est celle de l'*observation participante*. Nous avons effectué plusieurs enregistrements par nous-mêmes, ayant participé à des discussions entre les étudiants de l'UP. C'était des moments de la vie étudiante de tous les jours (conversation devant la salle, au restaurant universitaire et à la bibliothèque, à l'arrêt de bus). Etant donné la proximité de l'âge des interlocuteurs, l'introduction dans la conversation se faisait naturellement, d'autant plus que nous suivions quelques cours de Master 1 liés au sujet de notre recherche, à savoir « Linguistique contrastive », « Méthodologie de corpus » et « Contacts de langue ».

Pour enregistrer les conversations entre étudiants, il faut obtenir une autorisation. Pour respecter des principes déontologiques, nous avons préféré demander leur autorisation à l'avance, mais les enregistrements n'ont pas été faits le jour même afin d'obtenir les données les plus authentiques possibles. Cependant, on les a avertis une fois les enregistrements terminés pour redemander leur autorisation de les conserver. Enfin, sachant que le terrain se laisse apprivoiser plus facilement lorsque le chercheur fait partie de la communauté qu'il observe, nous avons fait réaliser quelques enregistrements par des intermédiaires (étudiants de l'Université de Poitiers et l'Université de Iaroslavl qui ont enregistré des discussions avec leurs pairs) en adoptant le même principe (prévenir les interlocuteurs quelques jours à l'avance et à la fin de chaque enregistrement).

En ce qui concerne les données radiophoniques, nous avons choisi l'émission « Guillaume Radio 2.0 » sur NRJ, chaîne radiophonique française, car l'animateur indique systématiquement l'âge de chaque auditeur qui appelle lors

des émissions, ce qui nous a permis de rester fidèle à la tranche d'âge choisie, ainsi que « РадиоАктивноеШоу » (« Show Radioactif ») sur Europa Plus, l'une des principales stations de radio commerciales privées de Russie.

Par rapport à la radio NRJ la collecte des données à partir des podcasts d'Europa Plus a requis davantage d'efforts, puisque le tri des enregistrements s'est avéré problématique, les animateurs n'annonçant que très rarement l'âge des interlocuteurs. Le corpus oral russe s'est avéré donc moins volumineux à la fin de cette étape de collecte à cause de l'insuffisance des métadonnées. Pour des raisons d'équilibre, nous avons pris la décision de consulter le *Corpus National de la Langue Russe* et avons extrait quelques occurrences pour les besoins de notre recherche. Le *Corpus National* est représentatif de la richesse et de l'évolution de la langue russe dans sa diachronie¹ et ses diavariantes. Il est accessible en ligne et contient, entre autres, un sous-corpus oral de plus de 7,5 millions d'entrées. Il comprend des transcriptions d'enregistrements des discours privés, publics et filmiques, et, surtout, son annotation méta-textuelle est riche et nous a permis de sélectionner des données correspondant à nos critères. En effet, les métadonnées renseignent : domaine de fonctionnement (privé, public, filmique), type de texte (conversation, interview, microdialogue, entretiens, thématique (vie privée, médecine et santé, politique et vie sociale etc.), zone d'enquêtes et durée des enregistrements, style (neutre, familier, officiel), profil des locuteurs (nombre d'interactants, âge, sexe, niveau d'études ou métier).

Nous n'avons utilisé que des occurrences correspondant aux critères choisis pour notre étude, à savoir : discours oraux spontanés privés entre étudiants de 18 à 23 ans. 77,6% des données orales spontanées que nous avons trouvées représentent des conversations privées produites au quotidien, 4,7% - des conversations téléphoniques, 17,7% - discours monologiques (souvent sous forme d'un résumé d'une œuvre littéraire, scientifique, filmique ou autre).

3. **Corpus écrit** : constitué d'occurrences relevées dans les réseaux sociaux (Facebook.com, WhatsApp, vk.com), notamment dans des groupes privés, ainsi que sur les forums (Blabla 18-25ans), représentatif de l'écrit oralisé (cf. Dejong, 2002 ; Anis, 2007; E. Stark, 2011, 2014). Notre corpus intègre des dimensions scripturales, puisque l'intense vie relationnelle de cette population en contexte numérique permet de collecter des variantes écrites voire multimodales du parler jeune. Nous avons voulu rompre avec la dichotomie écrit/oral projetée sur des variantes hautes/basses de la langue. Pour être sûre de l'authenticité des données collectées, nous avons demandé et eu l'accès à des groupes fermés créés sur Facebook par des étudiants, dont le groupe de « *L2 LEA Poitiers 2016/2017* », le groupe d'étudiants de *Master 1 en Linguistique* où ils discutent de leur vie étudiante et privée, aussi bien que le groupe « *Иняз, группа 317* » (*Langues Etrangères, groupe 317*) et « *2 группа переводчиков – инглиш* » (*Groupe 2 d'étudiants en traductologie anglaise*) sur le site russe vk.com.

Une fois collectées, les données orales ont été soumises à la transcription, les données écrites – enregistrées telles qu'elles avaient été attestées, l'ensemble des données – codifié et anonymisé. Rendre les données anonymes est très important pour la protection de la vie privée des enquêtés et « pour la légalité des corpus recueillis par les chercheurs » (Baudé (coord.), 2006, p.67). Dans le cadre de notre travail, nous avons utilisé pour nos enregistrements les formes de remplacement suivantes : remplacement par une abréviation (pour les noms de villes et d'institutions) et remplacement par pseudonyme pour les noms (car cette forme permet une bonne intégration de la forme de remplacement dans le fil du discours). Pour les données audio deux formes d'anonymisation ont été utilisées, à savoir : coupure des passages et remplacement par un bruit. Enfin, en ce qui concerne les captures d'écran des messages électroniques, nous avons eu recours au remplacement par un brouillage du signal (pixélisation ou placement d'un bandeau noir sur le texte).

Comme l'analyse linguistique de productions orales est impossible à partir de la seule source sonore, la transcription est une étape essentielle de la recherche. Toute transcription a toujours une finalité qui détermine ses principes. Compte tenu que nos recherches ne concernaient pas directement la forme phonique (phonétique, phonologie, prosodie) et ne visaient pas à marquer des particularités de prononciation, nous avons décidé de transcrire en utilisant l'orthographe traditionnelle sans avoir recours à des modifications graphiques appelées « trucages orthographiques » par Blanche-Benveniste et Jeanjean (1987). Une telle représentation de la phonétique mêlée à l'orthographe est, en effet, très imparfaite : elle est souvent source d'erreurs, c'est également une tâche laborieuse qui ne contribuerait pas par ailleurs à nos analyses. En outre, on ne peut nier l'effet de stigmatisation qu'engendre la lecture d'une transcription ainsi aménagée sur les propos de la personne et, par là, sur la personne elle-même (Gumperz et Berenz, 1993), d'autant plus que le parler des jeunes est a priori un phénomène stigmatisé. Par exemple, lorsque le locuteur fait une élision vocalique, nous avons signalé la non-réalisation de la voyelle par des parenthèses (*j(e) vois, t(u) sais*) et non pas par une apostrophe (*j'vois, t'sais*).

Suite à Blanche-Benveniste & Jeanjean (1987) et à Gadet (1992, p.21) qui plaident pour des transcriptions non ponctuées en estimant que « en ponctuant, le transcripteur suggère une analyse avant de l'avoir faite », nous n'avons conservé dans la transcription que quelques traces de l'élaboration de l'énoncé, justifiées par le fait que nous traitons

¹ Il couvre la période allant du XVIIIème au début du XXIème siècle

le parler jeune comme un phénomène particulièrement expressif ce qui permet de se rendre compte de cette expressivité. Nous avons donc noté : les pauses pleines ou ponctuantes (*euuh, ben, bah, эмм, ну, э.*, etc.) ; les répétitions de mots ou de suite de mots ; les amorces de morphèmes. Pour la même raison nous avons décidé de mettre des marques d'emphase, de cris, d'allongement de voyelles, aussi bien que de marquer quelques faits de prononciation (chute des consonnes /r/ et /l/ comme dans « *fout(r)e la merde* » et « *mett(r)e une droite* » ; « *s'i(l) vous plaît* » et « *le mec i(l) fait* » en français et simplification des groupes de consonnes en russe comme dans « *я то(ль)ко сажусь в поезд* :»² ; des bruits, rires et événements de la situation (en bas de page). A ces quelques remaniements près, nous sommes inspirée des conventions de transcription du GARS³. Le travail de transcription nécessite une grande attention de la part du transcripteur, pour noter des phénomènes qui sont habituellement gommés dans une écoute ordinaire. De ce fait, nous avons pris la décision d'effectuer *une transcription sélective* (Coveney, 2002). Pour les données écrites nous avons respecté l'écriture attestée, avec la ponctuation, les erreurs d'orthographe et les fautes de frappe.

Notre corpus d'étude comprend des données écrites et orales⁴ dans les deux langues :

	Corpus écrit	Corpus oral
<u>français</u>	1050 textes (18258 mots) dont la longueur varie de 1 à 50 phrases.	-30h d'enregistrements d'émissions radiophoniques effectuées sur NRJ 2.0 sur une période de 10 mois ; -18h d'enregistrements effectués en observation participante ; -La durée des extraits soumis à la transcription varie de 5 secondes à 2 minutes.
<u>Russe</u>	930 textes (16097mots). La longueur maximale est de 34 phrases.	57h (dont 35h d'enregistrements réalisés en observation participante, 4h d'enregistrements radiophoniques, auxquelles s'ajoutent des données extraites du Corpus National).

Tableau 2 : Corpus d'étude

Les données disponibles dans ce *corpus d'étude* ont été filtrées par le recours à une liste de mots obtenue par le biais du questionnaire introspectif (Cf. Mellet, 2002) pour constituer notre *corpus de travail*.

Enfin, il nous a donc fallu réaliser le codage de cette « matière brute » qu'est le discours spontané. Dans le cas du corpus oral (dont les exemples sont présentés sous forme d'une liste numérotée) nous avons fait une distinction entre les occurrences collectées par l'observation participante (ce qui est marqué par O) et celles que nous avons retenues lors des enregistrements radiophoniques (marqué par R). Le code contient également un chiffre indiquant la place de telle ou telle occurrence dans l'ordre (par ex. OO12 signifie qu'il s'agit d'une occurrence provenant du corpus oral collecté par l'observation participante dont le numéro d'ordre est 12 ; OR40 indique que l'exemple vient également du corpus oral, qu'il a été glané via radios et figure dans la liste sous le numéro №40).

Le corpus oral russe est organisé de la même façon. Le code comporte des indications suivantes : « Y » (provient de *устный*) - pour *oral*, l'abréviation « НКРЯ » dans le code s'ajoute aux extraits tirés du corpus national russe, « P » (de *радио – radio*) indique qu'il s'agit des données radiophoniques. Toutes les occurrences du corpus oral russe contiennent également un chiffre indiquant leur place dans l'ordre. Par exemple, le code Y30 signifie qu'il s'agit d'une occurrence provenant du corpus oral collectée par l'observation participante dont le numéro d'ordre est 30 ; УНКРЯ40 indique que l'exemple vient également du corpus oral, qu'il a été extrait du corpus national russe et qu'il figure dans la liste sous le numéro №40).

En ce qui concerne le corpus écrit français, il est également constitué d'occurrences collectées sur les forums (par ex. *Blabla 18-25ans*) et les réseaux sociaux (par ex. *Facebook, WhatsApp*) fréquentés par les jeunes. Les occurrences portent donc un code RS quand il s'agit des « réseaux sociaux » ou F qui signifie « forums ». A titre illustratif, citons deux codes : ERS2 qui indique que l'exemple provient du corpus écrit, qu'il est attesté sur un réseau social et il est troisième dans la liste ; tandis que EF10 signifie que l'exemple vient du corpus écrit, qu'il est collecté sur un forum et qu'il s'agit du dixième exemple dans l'ordre d'apparition. Les occurrences russes portent toutes un « П » dans le code (*П* provient de *письменный – écrit*) pour indiquer qu'il s'agit des données écrites, suivi de « CC » quand il s'agit des « социальные сети » (« réseaux sociaux ») ou de « Ф » qui signifie « форум » (« forums »). A titre d'exemple, citons deux codes : ПСС2 qui indique que l'exemple provient du corpus écrit russe, qu'il est attesté sur un réseau social et figure en deuxième position dans la liste. ПФ10 signifie que l'exemple vient du corpus écrit russe, qu'il est collecté sur un forum et qu'il s'agit du dixième exemple dans l'ordre d'apparition.

² Marquage par des parenthèses et non pas par une apostrophe.

³ Groupe Aixois de Recherche en Syntaxe (GARS). Pour les conventions du GARS, cf. Blanche-Benveniste & Jeanjean (1987), Blanche-Benveniste et al (1991), Bilger (2000). Voir également le site suivant : <http://www.up.univ-mrs.fr/delic/corpus/index.html>.

⁴ Nos corpus oraux peuvent être qualifiés d' « écologiques », puisqu'il s'agit d'enregistrements de conversations spontanées entre jeunes étudiants et d'enregistrements radiophoniques.

La méthodologie retenue nous a permis de constituer donc une ressource bilingue de corpus comparables en français et en russe. Et comme il est souvent difficile de faire se correspondre des données collectées dans différents contextes (langues ou type de médium), le glossaire sert de pivot entre les différentes ressources. Il est en quelque sorte une ressource de type « dictionnaire » permettant de relier des items à leur(s) définition(s) et à des emplois attestés.

La dimension expressive et subjective sont fortement caractéristiques des interactions entre jeunes. Il nous a donc semblé logique de nous adresser aux aspects pragmatiques des discours spontanés.

Dans le cadre du colloque, nous souhaiterions en présenter quelques particularités intéressantes. Il s'agira d'un emploi massif des marqueurs discursifs, terme que nous empruntons à Fraser (1999), Dostie (2004), et Andersen (2007). C'est le plus largement répandu bien qu'il entre en concurrence avec « particule énonciative » (Fernandez, 1994), « opérateur discursif » (Redeker, 1990, 1991), « connecteurs métatextuels » (Elo, 1993), « remplisseurs/organisateur » (Raupach 1984). En outre, cette liste n'est pas exhaustive.

II. Cadre théorique : notion de « marqueur discursif »

La notion de « marqueur » renvoie aux unités non référentielles (n'ayant pas un signifié dénotatif mais plutôt instructionnel/conversationnel) agissant sur les représentations cognitives construites par le discours, et dans la construction de ces représentations. Suivant Chanet (2004), précisons que les marqueurs discursifs (désormais MD) donnent des indications sur les opérations cognitives à conduire par les interactants dans l'activité discursive pour optimiser la communication. Les MD démontrent surtout l'attitude du locuteur, les stratégies argumentatives mises en place par celui-ci, ainsi que les rapports qui s'établissent entre le locuteur et l'allocutaire.

Dans la littérature consacrée au sujet (entre autres, Dostie & Pusch, 2007), nous retrouvons quelques caractéristiques fonctionnelles propres aux MD, à savoir :

- Ils sont joints au contenu propositionnel des énoncés mais n'y participent pas (par exemple : *t(u) sais*) ou, encore, ils constituent par eux-mêmes un énoncé et réalisent dans ce cas-là des actes illocutoires (par ex. : *genre ! ; voyons*) ;
- Ils sont extérieurs à la structure phrastique et appartiennent au niveau macro-syntaxique du discours (Blanche-Benveniste 1997 ; 2010) ;
- Ils servent à organiser des textes (par ex. : *genre, t(u) vois, du coup*) et assument diverses fonctions interactionnelles (p. ex. : signaux d'écoute et signaux d'appel à l'écoute) ;
- Ils expriment des sens subjectifs et intersubjectifs en mettant en relief d'un côté une forte implication du locuteur dans son propre discours tout en tenant compte de la présence de l'autre.

L'emploi fréquent des MD est une particularité syntaxico-pragmatique propre, avant tout, aux échanges peu planifiés de face à face, avec leurs répétitions, autocorrections et difficultés de formulation. Le recours aux MD sert alors à organiser le texte, à mettre en relief des informations significatives, à introduire divers éléments explicatifs afin de valider la compréhension de la part de l'interlocuteur. Comme la CMO se déroule dans la quasi-spontanéité, le recours aux MD y est également important et leur emploi à l'écrit prouve davantage qu'ils ont une certaine valeur pragmatique.

Par la suite, nous présenterions brièvement le tissu fonctionnel de quelques MD français et russes qui sont le plus largement actualisés dans les discours spontanés entre jeunes.

III. Genre et *muna*, marqueurs polyvalents

Genre et son équivalent russe *muna* sont les marqueurs les plus utilisés dans le parler jeune, que ce soit à l'oral ou à l'écrit. Leurs principales fonctions sont synthétisées dans le tableau 3 ci-dessous. Lors de notre communication, nous accompagnerons notre présentation par des enregistrements audios. Il est intéressant de souligner que les deux marqueurs ont suivi les mêmes étapes d'évolution. Comme le montre en détail Vigneron-Bosbach (2016), *genre* est d'abord un nom commun qui présente les attributs d'une morphologie nominale. Ensuite, progressivement le nom *genre* apparaît dans des contextes où ses caractéristiques nominales se modifient (il s'agit de « verrues nominales » en termes de Blanche-Benveniste *et al.*, 1990) où il perd en autonomie référentielle. A partir de cette structure déterminative apparaissent des locutions qui se figent, à savoir *dans le genre de* et *du genre de*, et peu à peu *genre* seul. Par la suite, *du genre* et *genre* connaissent des emplois interprédicatifs, et peuvent relier deux segments de discours, le deuxième illustrant alors le premier. Enfin de nouveaux contextes apparaissent dans lesquels seul *genre* est possible : *... eu:h c'est genre vingt-trois heures 'fi:n genre une heure tardive ... (OO40).*

En ce qui concerne « *muna* », il provient de « *mun* », un nom qui reçoit différents déterminants et modificateurs (adjectivaux, relatifs etc.) et qui présente les attributs d'une morphologie nominale (c'est un nom masculin, singulier, il se décline en fonction du rôle qu'il joue dans la phrase : sujet/complément) : *один mun* (« un type/genre), *интересный mun* (un type intéressant), *одного мuna* (du même type/genre), *разных типов* (« de types différents »). Suite à son évolution, ses caractéristiques nominales se sont modifiées : désormais, il fait partie d'une verrue nominale *muna того что* (« du même genre que ; dans le genre de ».) qui fonctionne comme une structure déterminative qualifiante qui n'est pas autonome d'un point de vue référentiel (*что* peut être omis). A partir de cette structure déterminative, nous voyons apparaître *muna* tout seul qui, décatégorisé, se fige dans sa forme déclinée au génitif *muna* (littéralement : « du type, du genre ») que l'on ne retrouve au départ que dans des structures comme *X muna Y*, mais qui se pragmatise par la suite et apparaît dans de nouveaux contextes, employé tout seul. Il connaît donc des emplois interprédicatifs où il fonctionne comme connecteur qui relie des éléments de discours, introduit des séquences illustratives jusqu'à apparaître dans certains environnements contextuels en fonction de ponctuant.

genre	типа
1. Fonction d'opérateur de catégorisation du type « classificateur »	
<p>* <i>elle est genre une reine des gaffes je te jure c pas possible t as envie de l'assommer avec une pelle (ERS 642)</i></p> <p>* <i>[...] e :t le mec mais vraiment huit ans j(e) te jure avec les lunettes et tout genre un gros boloss mais de la cité et [...] un des deux il dit wesh t'es- qu'est-ce t(u)as à me regarder toi tu t(e) prends pour qui alors que : que le-le gars c'était genre un homme d'affaires et tout tu vois [...](OO27)</i></p>	<p>* <i>[...] L1 - а у них в какой форме ++muna курсача L2 muna сутенанс да [...](Y1)</i> (Trad. L1 ils le font sous quel format ++ c'est genre un mémoire L2 c'est genre une soutenance oui).</p> <p>* <i>[...] я ей говорю твоя щас защита это наши четвёртый курс muna техникум вот наши в России, че ты там паникуешь (ERS652).</i> (Trad. L1 je lui dis ta soutenance là c'est comme notre mémoire de la quatrième année, c'est genre notre IUT en Russie, pourquoi tu stresses.)</p>
2A. Introduction du DR	
<p>* <i>du coup elle ma raconter les premieres experiences a M comme moi jetais ko genre elle me dit dun air trop choquée meme pas elle rigole genre : mais je comprend pas le chauffage moi jettoufe, jai demander au garde (mdr) de l'accueil et il ma dit si vous avez chaud sortez faire une pause dehors(ERS264).</i></p> <p>* <i>[...] elle m(e) fait tout u:n cours tu sais genre comme si j(e) savais pas c(e) que c'était un soin genre mai:s vous savez c'est c'est vraiment important le soin pour la pollution [...](OO31).</i></p>	<p>* <i>мы обедали вместе и он мне говорит вдруг muna давай сегодня вечером сходим куданибудь вместе (ПСС571)</i> (Trad. On a déjeuné ensemble, et il me dit d'un coup genre tu veux pas qu'on sorte tous les deux ce soir).</p> <p>* <i>Он говорит + брелок ни о чем muna + но надо заценить то + что он очень издалека [...](УНКРЯ167)</i> (Trad. Il dit que ce porte-clé est trop nul genre, mais faut apprécier qu'il vient de loin).</p>
2B. Profileur d'énonciation (introduction des propos plausibles)	
<p>* <i>[...] donc c'est la zone 5l tu mets juste ça + e:t du coup on en a parlé trop longtemps on a regardé des vidéos et je te jure là j'ai t- je suis en train de rentrer chez moi j(e) suis toute seule dehors mais genre j'ai TROP peur genre t(u) sais c(e) (n)'est pas une histoire de ++genre ouai:s les esprits existent et tout ça n'a rien à voir avec ça [...](OO29).</i></p>	<p>* <i>он понтанулся, muna я крутой, я вчера бухал)) muna мама разрешила (ПСС566).</i> (Trad. et là il commence à se la péter genre je suis trop cool, on a picolé hier genre j'ai eu l'autorisation de ma mère).</p>
3. Introduction d'un élément illustratif	
<p>* <i>Je me lance dans la licence d'anglais par correspondance !!! Je viens de checker le programme, et de lire toute une liste de mots flippants dont j'ignorais l'existence, genre</i></p>	<p>* <i>+ знаешь как воспринимаю русскую речь + я её слушаю и воспринимаю с точки зрения Жюстин знаешь я пытаюсь понять понимает ли Жюстин этот эту лексику или нет знаешь+ muna каждое- каждое слово взвешиваю: даже</i></p>

<p>"morphème, racine, affixe, flexion, allomorphie » tu vois ! dis moi je dois commencer à paniquer =) ??? (ERS186)</p> <p>[...] j(e) (ne) sais (il) y a un truc elle me saoule j(e) sais pas je (ne) sais pas encore j(e) (n') ai pas capté c'est quoi + euh peut-être sa tête euh genre pas d'émotions euh ou : (soupir) ouais j(e) crois que c'est ça genre elle (n')a pas d'expressions du visage donc ça m(e) saoule (OO54)</p>	<p>порой смысл не понимаю потому что думаю поняла ли она слово или нет так смешно++(Y25).</p> <p>(Trad. Tu sais comment je perçois maintenant les discours en russe, j'écoute et je les considère en me mettant à la place de Justine tu sais j'essaie de comprendre si Justine connaît ce vocabulaire ou pas tu sais genre je pèse chaque mot, des fois je perds même le fil parce que je reste concentrée sur les mots en me demandant si Justine comprend ce que ça veut dire ou pas+ c'est marrant).</p>
4. Fonction de ponctuant	
<p>Donc je disais genre j'ai été chez des potes et genre on a parlé mais genre TROP sérieusement pendant peut-être trois heures eu :h de politique euh l'ordre le nouvel ordre mondial et tout + et genre on a parlé trop longtemps de la: zone 51 j(e) sais pas si tu connais c'est en Amérique + t(u)iras t(e) renseigner pa(r)ce que c'est trop long à expliquer+ et du coup on a regardé des vidéos tout et tout genre j'ai trop trop peur e :t ++ ah mais ça monte trop à la tête ce truc mais genre+ ça fait pas peur dans le sens c'est un film d'horreur tu vois mais + 'fin tu te rends compte de de de plein de trucs 'fin moi perso j(e) suis sûre que derrière ce truc il y a tout le complot de : du nouvel ordre mondial dé: - derrière ce truc 'fin tu te renseignera:s [...](OO29).</p>	<p>[...]о :й давно : уже дело было++ ну тупа + обо всём поговорили++ тупа там знаешь про защиту ++ тупа там про всё про это+ что она была тупа ммм++ по ви-конференц-связи тупа+ с Россией тупа (смех) Ирина тупаподтвердила всё что Дафна представила в презентации +ну тупа прикольно было + тупа все поржали [...] (Y2).</p> <p>(Trad. oh c'était il y a longtemps ++ enfin genre ++ on a parlé de tout ++ genre de la soutenance tu vois ++de tout ça quoi+ du fait qu'elle a été genre eu :h+ en vi- en visioconférence genre+ avec la Russie genre (rire) genre Irina a confirmé tout ce que Daphné avait dit lors de sa présentation + genre c'était marrant + genre tout le monde a rigolé).</p>
1.5. Marqueur d'approximation	
<p>Mais en fait genre j'ai attendu trop lgtps et genre t'avais un groupe de genre cinq-six reubeu donc ils arrivent avec leurs joggings et avec plein de gros sacs :D mais tu sais ya pas de bruit au commissariat et ils arrivent genre les six dans le sas [...] (ERS426).</p>	<p>не знаю пока точно во сколько освобожусь, но давай тогда тупа ближе к вечеру встретимся часиков тупа в шесть- семь ок + ты же уже в центре будешь как раз да (ПСС998).</p> <p>(Trad. Je ne sais pas encore exactement à quelle heure je serais dispo mais on se capte plutôt au soir, genre vers six-sept heures d'acc tu seras déjà en centre-ville non).</p>
6. Emploi de manière avec des verbes de dire et de perception	
<p>[...] et le gars genre trop choqué il fait euh pardon c'est une blague là [...] (OO27).</p> <p>[...] il est en train d(e) leur parler mais genre grave mal [...] (OO11).</p>	<p>Она мне ответила тупа жестко так, сухо.. ну я подумала а не пойти бы тебе в жэз овца (ПСС 895)</p> <p>(Trad. Elle m'a répondu en mode clash, d'un ton sec.. Moi, je me suis dit mais vas te faire balader connasse).</p> <p>Я на неё посмотрела короч тупа недовольно, ну дала понять что я ее запалила и что это жесть как бы. (ПСС896)</p> <p>(Trad. Je l'ai dévisagé en mode mécontente, bref je lui ai fait comprendre que je l'avais grillé et que c'était genre nul).</p>

Tableau 3 : Emplois des MD genre et muna

IV. En mode et son équivalent russe – *макоў*

A l'origine *en mode* fait partie d'une locution verbale *être en mode*, probablement calquée sur l'anglais *to be in ... mode* (ou *to get into ... mode*). Employée initialement dans des contextes techniques, notamment en informatique, cette construction s'est répandue en anglais dans la langue courante dans des emplois figurés (*to be in study mode*, *in panic mode*, etc.). Le français admet les emplois techniques de cette expression. Toutefois, on observe également l'usage extrêmement fréquent de la locution *en mode* dans le sens d'« adopter une attitude particulière ». Comme des voitures qui peuvent être utilisées en mode

automatique ou manuel, l'humain se met en mode « stress », « survie » ou « détente ». Le fonctionnement de cette locution nous semble intéressant, d'autant plus qu'elle est polyfonctionnelle.

En ce qui concerne *такой*, à l'origine il est catégorisé par les grammaires russes comme un pronom démonstratif. Il est utilisé afin de mettre en relief un des objets semblables (par exemple : *я уже где-то видел такой знак* - *J'ai déjà vu un signe comme ça quelque part*) ou une des caractéristiques communes (ex. : *такие тучи к плохой погоде* – *De tels nuages sont signe de mauvais temps*). Ayant une forme adjectivale, il est fréquemment catégorisé comme adjectif pronominal. Son emploi accompagne souvent la mimogestualité et indique une qualité sans la nommer, en désignant indirectement une ressemblance ou une comparaison possible des objets ou de leurs paramètres. Employé antéposé à un nom ou à un adjectif, il sert à intensifier la qualité (« *он такой няша !* » - Il est tellement mimi). Dans les dictionnaires bilingues, nous trouvons plusieurs équivalents français pour le pronom *такой*, à savoir : *tel + Nom, tel que, de ce genre, ce type de, pareil que, comme ça*. Lors de son évolution, *такой* développe des emplois pragmatiques et apparaît à l'oral d'abord en coalescence avec un verbe de dire soit antéposé, soit postposé : *он такой говорит, она вдруг сказала такая* (Il dit genre ; d'un coup elle fait genre). Cet ensemble verbal étant par la suite réduit, nous relevons également des cas où le verbe de dire est omis. L'étape suivante d'évolution se manifeste dans l'apparition du marqueur *такой* en coalescence avec des verbes de position (s'asseoir/être assis, se mettre debout). Etant à l'origine un pronom démonstratif et gardant cette visée illustrative, il sert, dans certains contextes, à introduire un élément illustratif, à préciser le fait de dire ou la façon de se tenir, de se comporter, de se placer. *Такой* se trouve donc en voie de pragmatization, car il migre vers un domaine de la langue qui est celui de l'organisation du discours et de la structure informationnelle. Nous l'observons se comporter dans des interactions spontanées comme un élément introducteur du discours rapporté ou fonctionner comme un marqueur de manière. Cependant, ses caractéristiques grammaticales ne sont pas encore modifiées, il continue à s'accorder en genre et en nombre avec un élément par lequel il est régi et garde une affinité sémantique avec son unité.

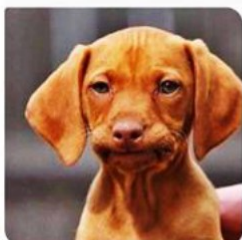
En mode	Такой
1A. Introduction de DR	
<i>[...] et l- la flic mais genre TROP contente genre e-elle est genre tu vois + on faisait que d(e) parler de ce qui s(e) passait et tout elle était là en mode ouai :s mais (il) (n') y a pa :s (il) (n') y a pas assez d€ preuves i(l) faut un truc qui ranime tout [...].</i>	<i>[...] L1 ++ я говорю а чё пото:м я вот через два года приеду будем жить вместе он такой все возможно++ я такая бл* + потом я не помню как мы это на-как на это мы вышли он такой у нас никогда не будет общего семейного бюджета (VI).</i> (Trad. L1 je fais et plus tard ca va se passer comment je rentrerais dans deux ans et on emménagera ensemble et lui il fait genre tout est possible ++ moi j'étais là genre purée et après je ne me souviens plus comment on a abordé ce sujet mais il dit genre on aura jamais un budget familial commun).
1B. Profileur d'énonciation (introduction des propos plausibles)	
<i>Du coup je tirais agueule je répondais pas Il me donne une tape sur l'épaule en mode bah alors qu'est-ce qui se passe Et là c'était le truc a pas faire (ERS76).</i> <i>[...] et du coup (il) y a un mec à côté qui le :s regarde en mode euh oh tu tu t(e) crois où tu vois mais il a rien dit [...]. (OO27).</i>	Nous n'avons pas relevé d'emploi de <i>такой</i> en fonction de profileur du discours. Cela serait dû à son affinité avec le sens prototypique (celui de démonstration, d'illustration : il demeure donc dans la perspective du réel et non pas dans le fictif, contrairement à <i>мина</i> , par exemple, qui peut jouer le rôle de modalisateur).
2. Emploi adverbial de manière	
<i>[...] et la flic elle nous dit+ elle nous dit de venir mais en mode assez sévère t(u) vois+ c'est normal on (n') est pas chez le médecin wesh toujours la même référence [...]. (OO53)</i>	<i>[...] потом я ее поздравила подожди++ с днём рожде:ния и она а и вот она мне послала с днём рождения+ такая (смех) сухарь типа+ ну вот [...](VI) .</i> (Trad. Puis je lui ai envoyé un message attends euh un message d'anniversaire et elle aussi oui voilà elle m'a envoyé aussi genre joyeux anniversaire en mode (rire) trop sec et puis voilà quoi).

Dans ce type d'emploi, *en mode* porte une nuance connotative particulière en faisant ressortir une valeur intentionnelle. La personne délocutive qui *regarde/parle/vient en mode* adopte délibérément ce type de comportement, autrement dit, elle « se met en mode ». Ainsi, le marqueur en question dans ce type d'emplois dynamiques traduirait l'aspect intentionnel d'un acte verbal ou non-verbal. En revanche, si nous employions *genre* à sa place, il introduirait plutôt une précision sur le comportement de quelqu'un que l'énonciateur jugerait d'une manière subjective comme « sévère », « vénère » ou autre.

Comparons :

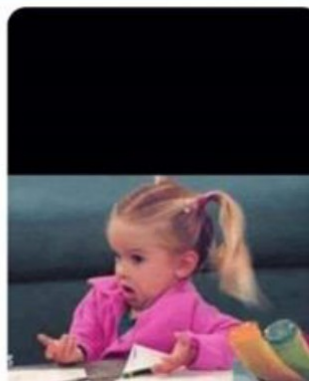
elle nous dit de venir mais genre assez sévère
 * [...] *il me regarde en mode* il en a marre [...]
 (OO10) ≠ [...] *il me regarde genre* il en marre.

Purée elle me soule maman avec ses mails de topchretien de merde la, j'arrive a saturation c'est pire que les spams 😞😞 je suis là en mode



И я просто стою такая WTF? Типа

10:22



(Trad. Et je suis là en mode WTF ? Genre)

3. Introduction d'une reformulation paraphrastique explicative.

* *pfiff c vrmt pas mon genre de personne je te jure elle est lente en mode elle s'endort, elle fait tout au ralenti et tout je deteste mdr* (ERS472).

* [...] *ma coiffure elle m'a juste coupé la longueur en mode* c(e) que j(e) peux faire toute seule en fait [...] (OO32)

* [...] *du coup euh il commençait à glisser dans le bas ge :nre en mode il tombait HYPER lentement tu vois [...]* (OR49)

* *Тебе надо быть неприступной. Такой девочкой-загадкой* (УНКРЯ356).

(Trad. Il faut que tu sois inaccessible, telle une fille mystérieuse).

* *У Миронова есть свой имидж, он такой здоровячок, такой хороший* (ПФ355)

(Trad. Mironov, il a son propre style, tel un bonhomme costaud toujours en mode tout mimi).

Ici, *такой* a une valeur de conceptualisation que le MD *en mode* ne transmet pas.

Tableau 4 : Emplois des MD *en mode* et *такой*

Les MD *genre* et *en mode* qui « envahissent » les discours spontanés entre jeunes français, tout comme *muna* et *такой*, qui sont largement actualisés dans les échanges entre jeunes russes, ne sont pas des équivalents stricts (il y a certaines divergences fonctionnelles liées à leur sémantisme étymologique). Ils apparaissent bien souvent comme co-occurents. Ils ne sont pas complètement pragmatialisés, car demeurent sous l'emprise des sens présidés à leur genèse : de similitude, d'approximation et de non-prise en charge du contenu propositionnel propre à *genre* et à *muna*, et de manière et d'intentionnalité propre à *en mode* et à *такой*. *Genre* et *muna* se manifestent plus « statiques » dans leurs emplois. Selon nous, l'emploi statique est caractéristique des contextes descriptifs dont la modalité est d'ordre subjectif et/ou approximatif : la réalité y est une interprétation qualitativement différenciée de la situation réelle. Le focus est posé sur le locuteur et plus précisément, sur sa perception de cette réalité dont il tient à se détacher. *Genre* et *muna* interviennent donc dans des contextes où le locuteur adopte une position distincte d'observateur extérieur et présente des épisodes en essayant de faire comprendre son ressenti à travers une stratégie de comparaison avec d'autres situations semblables. Les MD en question sont convoqués, entre autres, afin de poser un cadre modal d'incertitude, car le locuteur vise à souligner que sa vision des choses peut être potentiellement différente de la réalité. La valeur de non-prise en charge du contenu propositionnel est omniprésente dans tous les contextes d'apparition.

En ce qui concerne *en mode* et *такой*, ils apparaissent dans des contextes où une dynamique de monstration est privilégiée. L'interlocuteur devient participant des scènes qui se déroulent devant lui, il est fortement impliqué dans le discours et les vit dans leur développement dynamique. Il y agit en se mettant à la place de la personne

délocutive, il est acteur de ces scènes qui sont par la suite recréées de manière théâtralisée afin que l'interlocuteur puisse se sentir partie prenante. L'emploi dynamique se caractérise donc, à notre sens, par une création d'un espace de contextualisation propice à la visualisation, voire à la théâtralisation : la réalité y est conceptualisée de la sorte que les communicants puissent y agir. Le focus est posé sur l'interactivité et l'intentionnalité des actes verbaux et non-verbaux. Bien que les MD en question acquièrent de nouveaux emplois et apparaissent dans de nouveaux contextes, ils gardent une certaine affinité avec leurs unités sources : *genre* et *muna* continuent à véhiculer la facette de comparaison (cf. Vigneron-Bosbach, 2016) et/ou de différenciation, ainsi que la nuance de non-prise en charge du contenu propositionnel (cf. Dufaye, 2014), *en mode* – l'idée d'une attitude consciente et voulue, *такой* -la dynamique de monstration.

Force est de souligner que la valeur de *genre*, *en mode*, *muna* et *такой* doit être analysée en rapport avec d'autres marqueurs qui réalisent dans leur ensemble des actes illocutoires.

V. *tu sais, tu vois, t'imagines* et leurs équivalents russe : *знаешь, понимаешь/видишь, представь/прикинь*

Nous souhaitons évoquer par la suite des unités qui, d'un point de vue morpho-syntaxique, sont figés dans une seule forme, celle de la deuxième personne. C'est le cas des formes *tu sais, tu vois, t'imagines* (Andersen, 2007, p.14) et de leur « parfaits » équivalents *знаешь, понимаешь/видишь, представь/прикинь*⁵, caractéristiques du russe parlé, puisque ces emplois spécifiques des verbes *voir, imaginer* et *savoir* que nous traitons comme MD sont très fréquents en français et en russe. Les linguistes les définissent différemment : Blanche-Benveniste (1991) les classifie comme « phatique », Dehé & Wichman les renvoient à la catégorie des parenthétiques ayant une fonction de marqueurs de discours ou de ponctuels, Traugott (2010) constate l'acquisition d'une fonction phatique interactionnelle, Dostie (2004) les traite comme « marqueurs de balisage », Brinton (2008) estime que ces unités, bien qu'optionnelles au niveau grammatical, sont pragmatiquement nécessaires. Les fonctions que ces marqueurs accomplissent sont explicitées ci-dessous (Tableau 5) à partir des exemples que nous avons pu retenir dans nos corpus.

tu sais , tu vois, t'imagines	знаешь, понимаешь/видишь, представь/прикинь
1. Introduction d'une nouvelle information	
<p><i>t'imagines</i> je viens de faire une lettre a la place d'un pote Reubeu il ma meme disputer (ERS507). [...] ++ mais gros coup d(e) pute mais là je l'ai détestée genre t(u) imagines + le mec euh il va s(e) marier [...] (OO49).</p>	<p>Понимаешь + когда там грузинская свадьба была они приглашали всех людей+ которые идут мимо [...] (УНКРЯ2). (Trad. tu sais, quand j'ai été au mariage géorgien, ils ont invité tout le monde, même les gens qui passaient juste devant). Просто прикинь у нас девочка короче которая типа с Нагиевым-то встречается я тебе рассказывала, ну в общем тупая она короче [...] (Y29). (Trad. Imagine, la fille, celle qui est en couple avec Naguiév genre, je t'en avais parlé, bref, celle qui est trop bête[...]).</p>
2. Fonction de satisfaction d'un besoin d'explication	
<p><i>et ptn je suis choquée</i> tu sais j'ai des amis a Valence pas loin de Lyon que j'ai rencontrer grace a couchsurfing j'ai passer deux semaine chez eux pr mon stage c 3 filles de mon age en coloc (ERS508). <i>Mais du coup les deux filles - là elles m'ont toujours pa:s répondu + t(u) imagines c'est dans dix minutes ba:h à l'aise hein j(e) sais même pas si j'y vais [...] (OO71).</i></p>	<p><i>если будет предлагать заехать за всеми пассажирами, я от мари кюри не откажусь)) ну знаешь бывают такие типа: довезууу (ПСС904)</i> (Trad. Si jamais il propose de venir chercher tous les covoitureurs, je serais d'accord qu'il vienne à Marie-Curie. 'fin tu sais il y a des gens comme ça genre ouaiiii je te déposerai). [...] три ночи подряд + э:возвращались в пять утра и: ну ты прикинь моё состояние естественно я ничё не сделала [...] (Y21). (Trad. On rentrait à cinq heures du matin trois jours d'affilée ; bah imagine mon état, du coup j'ai évidemment rien fait).</p>

⁵ Знаешь - znaech = « tu sais » ; понимаешь - ponimaech = (« tu comprends ») ; видишь - vidich = « tu vois » ; преставь/прикинь - predstav'/prikin' = « t(u) imagines ».

3. Fonction de mobilisation d'un savoir partagé

*là tu vois y'a une pote (française) qui m'envoie par **КАКИЕ**-то дни рождения вечно отмечают! Ну*

<p>sms : ouais attends la je t'écis pendant que je fais caca. On en parle facilement tu vois mdr (ERS312).</p> <p>* [...] genre déjà c'est l(e) matin tout le monde me saoule t(u) vois++ regard dans le vide tu me connais ++ e:t donc j'arri:ve [...] (OO23)</p>	<p>представляешь да, сразу видно наша ассоциация (ПСС710).</p> <p>(Trad. Toujours en train de fêter un anniversaire de quelqu'un, enfin, tu vois, on sait direct que c'est notre association).</p> <p>* [...] я расстроилась потом по- это позвала её в группу типа на день рождения, а она из неё удалилась сразу. ну знаешь как она любит короче [...] (V1).</p> <p>(Trad. J'ai été un peu déçue... parce que je l'ai rajouté dans mon groupe genre pour organiser ma soirée d'anniversaire, bah elle a tout de suite quitté le groupe... enfin tu vois comment elle aime faire).</p>
--	---

4. Fonction de mobilisation d'un espace de contextualisation

<p>* [...] il était persuadé qu'elle avait les pieds noirs elle voulait les cacher mais moi et Laure t' imagines mais on en pouvait plus [...] (ER239).</p> <p>* [...] donc tu vois (il) y a les flics tout autour et tout parce que (il) y avait que le truc attentat fin (il) y a des flics partout genre on doit faire la queue dehors et on doit rentrer un par un dans un sas pour s(e) faire fouiller scanner et tout fin c'est un truc de ouf et donc t(u) vois (il) y a+ c'est vraiment surveillé [...] (OO22).</p>	<p>* Ну короче, там знаешь, ты заходишь, перед тобой сразу холл большой как в театрах, ну представляешь да. Ну и короче, мы заходим!!!! Прикол не могу прям [...] (ПФ 281)</p> <p>(Trad. Bref, du coup, là-bas tu sais tu entres y a une grande halle devant toi comme dans les théâtres tu vois du coup on entre et là jpp c'est trop drôle).</p> <p>* Классно так пообщались, в баре посидели ! На Кирова этот маленький знаешь около книжного ? Короче, мы думали, у него друзья только девочки. Прикинь, одни парни были (ПСС357) .</p> <p>(Trad. C'était cool, on a été dans un bar. Tu sais le petit bar rue Kirova à côté de la librairie. Du coup, nous, on pensait qu'il traînait que avec des filles. Bah imagine, il y avait que des mecs).</p>
---	--

5. Validation de compréhension

<p>* HEY ptn je t'ai pas dit ça fait genre une semaine c trop bizarre j'ai genre des flash de situation vécu tu vois mais vraiment banale (ERS182).</p> <p>* [...] en plus le pote-là qui est v(e)nu en mode eu :h tu vois trop sa tête pervers qui est v(e)nu baiser les fille t(u) vois [...] (OO40).</p> <p>* [...] (il) y avait deu:x deux petits mais vraiment des petits je te jure ils- huit ans tu vois mais les les petits de cité quoi un truc de ouf [...] (OO27)</p>	<p>* L1 Да у неё знаешь был э: такой вид как будто она (сеё)час сознание потеряет ну знаешь такая под кайфом такая ++ да не во(обще) отлично кайфе:ц такая сидит нормально (смех) [...] А: да: она короче мы- включила je suis malade она такая одна знаешь такая (смех) alcoolique такая je suis malade: (смех) мы сидели и ржали + нет она пела хорошо но это было так прикольно такая одна знаешь страдалица такая (V24)</p> <p>(Trad. L1 Mais tu sais elle avait l'air genre elle allait tomber dans les pommes tu sais en mode elle plane genre whaa trop cool au top elle était là en mode posée tranquille [...] Ah mais oui elle- enfin nous on a mis « je suis malade » elle était là en mode toute seule genre alcoolique « je suis malade » on était là ko non mais elle chantait bien mais c'était juste trop drôle genre toute seule tu sais en mode elle souffre genre).</p> <p>* [...] вот в чём видишь дело, палка о двух концах [...] (ПСС14).</p> <p>(Trad. C'est ça la question en fait tu vois c'est une arme à double tranchant).</p>
---	--

Il est intéressant de souligner que si les fonctions que les marqueurs en question accomplissent sont les mêmes en français et en russe, il existe néanmoins quelques divergences quant à la fréquence et à la distribution des marqueurs dans tel ou tel contexte. Nous en parlerons en détails lors de notre présentation orale. Néanmoins, notons que malgré quelques points divergents, ils sont sollicités (dans la majorité des cas) dans des contextes propices à la visualisation que ce soit en russe ou en français.

En guise de conclusion, notons que les MD sont donc un outil linguistique au service des stratégies discursives que l'énonciateur utilise afin de maintenir l'attention du co-énonciateur, de s'assurer de son engagement dans le discours et de sa compréhension, de demander une approbation discursive, de présenter le texte par épisodes, d'introduire des éléments significatifs, de structurer des énoncés. Notre étude met

également en évidence la diversité et la densité de moyens linguistiques utilisés à des fins expressives correspondant aux caractéristiques communicationnelles de jeunes natifs d'un numérique transnational : sujets de préoccupations et univers référentiels des jeunes russes et français apparaissent comme « mondialisés », outils linguistiques et stratégies discursives multimodales adoptés semblent bien en être la manifestation.

Bibliographie

- ANDERSEN H. L. (2007), « Marqueurs discursifs propositionnels », *Langue française*, n° 154, p. 13-28.
- AZZOPARDI, S., (2010). Présentation : « Les » linguistiques de corpus : enjeux spécifiques inhérents à la nature des données traitées. *Cahiers de praxématique* 54-55. Montpellier : Pulm. 209-220.
- BAUDE O. (coord.) (2006), *Corpus oraux - Guide des bonnes pratiques 2006*, CNRS Editions & Presses Universitaire d'Orléans. 203 p.
- BRINTON L. J. (2008), *The Comment Clause in English. Syntactic Origins and Pragmatic Development*, Cambridge: Cambridge University Press/ Studies in English Language.
- CAPPEAU P. & MORENO A. (2017), « Les tendances grammaticales », in F. Gadet (dir.), *Les parlers jeunes dans l'Île-de-France multiculturelle*, Paris : Ophrys, 73-99.
- CAPPEAU, P., GADET, F. (2007). « L'exploitation sociolinguistique des grands corpus. Maître-mot et pierre philosophale », *Revue française de linguistique appliquée* 1/2007, Vol. XII, 99-110, consulté le 25/08/2019 à l'adresse suivante : www.cairn.info/revue-francaise-de-linguistique-appliquee-2007-1-page-99.htm.
- CHANET, C. (2004). « Fréquence des marqueurs discursifs en français parlé : quelques problèmes de méthodologie », *Recherches sur le français parlé*, n°18, pp. 83-107.
- COSNIER J. (2012), « Axiomes de la communication multimodale », in Vion R., Giacomi A., Vargas C., *La corporalité du langage. Multimodalité, discours et écriture*, Aix-en-Provence : Presses Universitaires de Provence, pp.103-106.
- DEKHISSI L. (2013), *Variation syntaxique dans le français multiculturel du cinéma de banlieue*, Thèse de doctorat, Université d'Exeter.
- DEVELOPTE, C., MANGENOT, F. (2010). « L'analyse des corpus multimodaux en ligne : état des lieux et perspectives », in *Actes du congrès de l'Actualité de la recherche en éducation et en formation (AREF)*, Université de Genève, consulté le 25/08/2019 à l'adresse suivante : DOI : 10.1051/shsconf/20120100213.
- DOSTIE G. (2004), *Pragmaticalisation et marqueurs discursifs. Analyse sémantique et traitement lexicographique*, Bruxelles, De Boeck.Duculot.
- DOSTIE G. (2013), « Les associations de marqueurs discursifs. De la cooccurrence libre à la collocation », *Linguistik online*, 62, URL : http://www.linguistik-online.de/62_13/index.html , consulté le 27 février 2017.
- GADET, F., GUERIN, E. (2016). « Construire un corpus pour des façons de parler non standard : "Multicultural Paris French" », *Corpus*, 15, 285-307.
- MARTIN-JUCHAT F. (2008), *Le corps et les médias, La chair éprouvée par les médias et les espaces sociaux*, De Boeck, Culture et Communication, 150 p.
- MELLET S. (2002), « Corpus et recherches linguistiques : introduction », *Corpus*, n°1, pp. 5-12.
- OLIVIERI M., « Syntaxe et Corpus. Présentation », *Corpus* [En ligne], 9|2010, URL: <http://journals.openedition.org/corpus/1950> ; DOI:<https://doi.org/10.4000/corpus.1950>
- PLENAT, M. (2000). « Quelques thèmes de recherche actuels en morphophonologie française », *Cahiers de Lexicologie*, 77, 2, 27-62.
- RASTIER, F. (2005). « Enjeux épistémologiques de la linguistique de corpus », in G. Williams (ed.), *La linguistique de corpus*, Rennes : Presses Universitaires, 31-46.
- SAVREUX, F. (2017). *De la syntaxe à la sémantique du discours : étude en corpus des constructions en parce que*. Thèse de doctorat, Université Toulouse le Mirail - Toulouse II.
- TEUBERT W. (2009), « La linguistique de corpus : une alternative », *Critical Discourse Analysis I*. Les notions de contexte et d'acteurs sociaux, pp. 185-211.
- VIGNERON-BOSBACH J. (2016), *Analyse contrastive des marqueurs genre en français, like en anglais et so en allemand dans des corpus d'oral et d'écrit présentant un faible degré de planification*, Thèse de doctorat, l'Université de Poitiers, 486p.

Temps verbaux et ‘degrés d’oralité’ en français et en occitan

Marianne Vergez-Couret, Université de Poitiers
Janice Carruthers, Queen’s University Belfast

1. Discussion courte des notions d’oral’ et d’écrit’ et des modèles multi-dimensionnels tels que Biber et Conrad (2009) et Koch et Oesterreicher (2001).
2. Discussion courte de la recherche sur les temps verbaux (TV) dans la narration orale et la question de l’alternance entre le passé simple/passé composé et le présent narratif en occitan et en français. Fonctions discursives établies de l’alternance.
3. Notre problématique et nos questions de recherche. Discussion de la notion de ‘degrés d’oralité’ et des paramètres pertinents dans le cas de la narration orale en occitan et en français.
4. La construction de notre corpus : les trois sous-corpus en occitan et le sous-corpus français. Discussion de la question des sources des narrations (tradition orale, textes écrits, mélange des deux), du canal de transmission des contes et récits (oral ou écrit), du contexte de transmission (texte publié, performance orale dans un contexte traditionnel intime, performance publique avec instruments etc.), des facteurs diachroniques et dialectaux, de la relation des locuteurs avec l’occitan et le français, du niveau de spontanéité et de préparation, de la façon de travailler des conteurs etc.
5. Résumé des descripteurs des quatre sous-corpus :
 - OOT – Occitan oral traditionnel
 - OWT – Occitan écrit (written) traditionnel
 - OOO – Occitan oral contemporain
 - FOC – Français oral contemporain

<i>OOT</i> + <i>spontané</i> – <i>planifié</i> + <i>traditionnel</i>	<i>OWT</i> – <i>spontané</i> + <i>planifié</i> + <i>traditionnel</i>	<i>OOO</i> + <i>spontané</i> + <i>planifié</i> +/- <i>traditionnel</i>	<i>FOC</i> + <i>spontané</i> + <i>planifié</i> – <i>traditionnel</i>
<i>sources orales</i>	<i>sources orales</i>	<i>sources écrites et orales</i>	<i>sources écrites</i>
<i>‘oralité mixte’</i>	<i>‘oralité mixte’</i>	<i>‘oralité seconde’</i>	<i>‘oralité seconde’</i>
<i>locuteurs nés entre le milieu du 19ème siècle et le début du 20ème siècle</i>	<i>locuteurs nés au début du 19ème siècle</i>	<i>locuteurs nés à partir du milieu du 20ème siècle</i>	<i>locuteurs nés à partir du milieu du 20ème siècle</i>

6. Brève discussion de notre méthodologie
7. Analyse: (i) des temps verbaux narratifs principaux dans les contes et récits

TV narratif(s) principaux	OOT	OOO	FOC	OWT
PS	5 (narrations)	6	1	13
PC	2	2	8	0
Pnarr	16	4	4	3
PS/Pnarr (50/50)	3	1	0	3

Discussion générale de cette distribution et des différences entre les quatre sous-corpus.

8. Analyse : (ii) de chaque sous-corpus en plus de détail, prenant en compte le concept de degrés d’oralité – basé sur les sources des histoires dans chaque cas, le canal de transmission, le contexte de performance,

le contexte sociétal en ce qui concerne l'alphabétisation et l'oralité ; discussion également des facteurs tels que la variation diachronique et diatopique, la variation individuelle, le contact entre l'occitan et le français, etc. Distributions relativement claires dans OOT et OWT. Discussion de l'hybridité et de la complexité de OOC et FOC.

OOT

<i>Titre</i>	<i>Conteur</i>	<i>Patron des temps verbaux</i>
L'istòria d'un paisan del caulet (M) ⁰	M. Berthoumieu	<i>Uniquement Pnarr</i>
De passatge a Bordeu (M)	Armand Lescouyères	<i>Principalement Pnarr/ Quelques PS</i>
Lo ciror e la filha (M)	Germain Montaubric	<i>Principalement Pnarr/ Quelques PS</i>
La craba qu'èra bièn mentura (M)	Alice Meyrat	<i>Principalement Pnarr/ Quelques PS</i>
Histoires de tours du Drac (M)	Noémie Batifol	<i>Principalement Pnarr/ Quelques PS</i>
Un dròlle pas interessent (M)	Hermine Calastrenc	<i>Principalement Pnarr/ Quelques PS</i>
Fin volur (M)	Hermine Calastrenc	<i>Principalement Pnarr/ Quelques PS</i>
Josèp l'amic de las bestias (M)	Céline Calvet	<i>Principalement Pnarr/ Quelques PS</i>
Lo marat e la marata (C)	Mme Gibli	<i>Principalement Pnarr/ Quelques PS</i>
L'œuf de jument (C)	M. Berthoumieu	<i>Principalement Pnarr/ Quelques PS</i>
Jan de l'ors (M)	M. Berthoumieu	<i>Principalement Pnarr/ Quelques PS</i>
La bèstia de 7 caps (L)	M. Berthoumieu	<i>Principalement Pnarr/ Quelques PS</i>
Le jeune homme et le lion (L)	M. Berthoumieu	<i>Principalement Pnarr/ Quelques PS</i>
Lo conte del nau lops (M)	Marguerite Lagarde	<i>Principalement Pnarr/ Quelques PS</i>
Le mic e mac (L)	André Daste	<i>Principalement Pnarr/ peu de PS</i>
Légende des animaux... (M)	Elizabeth Barbedienne	<i>Principalement Pnarr/ 1 PS</i>
Istòria de Gargantua (S)	Céline Calvet	<i>Principalement PS</i>
Histoire du lac... (S)	Noémie Batifol	<i>Principalement PS</i>
Joan d'Auret (M)	Aurélie Philipot	<i>PS/ Pnarr</i>
L'ueu de cisampa (S)	Storyteller_OOT_1	<i>PS/ Pnarr</i>
Le conte des trois poulettes (L)	Maria Girbal	<i>PS/ Pnarr</i>
L'orra bestia de la coeta verda (M)	Céline Calvet	<i>Principalement PS/ Quelques Pnarr</i>
Istoèra del curè d'Escaudes... (M)	Daniel Lacampagne	<i>Principalement PS/ Quelques Pnarr</i>
Un talhur seguit per un canh (S)	Storyteller_OOT_2	<i>Principalement PS/ Quelques Pnarr</i>
La taupa e lo grapaud (S)	Storyteller_OOT_3	<i>Uniquement PC</i>
Lo bon diu e lo pleu pleu (S)	Storyteller_OOT_3	<i>Principalement PC/ 1 PS</i>

OOC

<i>Titre</i>	<i>Conteur</i>	<i>Patron des temps verbaux</i>
L'ueu de mula (M)	Storyteller_OOC_MC	<i>Principalement PS</i>
La hada e la goja (M)	Storyteller_OOC_MC	<i>Principalement PS</i>
Les fées de la dune de Bombèt (L)	Storyteller_OOC_MC	<i>Principalement PS</i>
Lo peisson-lèbre (M)	Storyteller_OOC_MB	<i>Principalement PS/ Quelques Pnarr</i>
Lo gabarròt (VL)	Storyteller_OOC_PB	<i>Principalement PS/ Quelques Pnarr</i>
L'ase, lo moton, la pola e lo lop	Storyteller_OOC_PB	<i>Principalement PS/ Quelques Pnarr</i>
Lo vaiet e lo diable (L)	Storyteller_OOC_PB	<i>PS/ Pnarr</i>
Lo grelh (M)	Storyteller_OOC_FM	<i>Seulement Pnarr</i>
Lo dròlle logat (L)	Storyteller_OOC_MB	<i>Seulement Pnarr</i>
La bòta (tL)	Storyteller_OOC_MB	<i>Principalement Pnarr/ peu de PC</i>
La caça volenta (M)	Storyteller_OOC_MB	<i>Principalement Pnarr/ 1 PC et 2 PS</i>
La mair deu conte (M)	Storyteller_OOC_PV	<i>Principalement PC/ 1 Pnarr et 2 PS</i>
Las abelhas (tL)	Storyteller_OOC_PV	<i>Principalement PC/ quelques Pnarr et PS</i>

FOC

<i>Titre</i>	<i>Conteur</i>	<i>Patron des temps verbaux</i>
--------------	----------------	---------------------------------

⁰ C=court ; M=moyen ; L=long ; tL=très long – selon la longueur du conte ou récit.

<i>Le fils du roi et le corbeau (L)</i>	Storyteller_FOC_ECv	Principalement PS
<i>Ankou (L)</i>	Storyteller_FOC_PC	Uniquement PC
<i>Histoire de Femmes (L)</i>	Storyteller_FOC_FD	Uniquement PC
<i>Le crapaud et la tortue (M)</i>	Storyteller_FOC_ECa	Uniquement PC
<i>Mange ma graisse (M)</i>	Storyteller_FOC_ECa	Uniquement PC
<i>Poukou I (M)</i>	Storyteller_FOC_GB	Principalement PC/ 1 Pnarr
<i>Le roi des pigeons (L)</i>	Storyteller_FOC_IS	Principalement PC/ peu de PS et Pnarr
<i>Contes de Bourgogne (L)</i>	Storyteller_FOC_CM	Principalement PC/ peu de PS
<i>Le paysan et la paysanne (tL)</i>	Storyteller_FOC_AK	Principalement PC / beaucoup de Pnarr
<i>Le cordonnier de la ville du Caire (L)</i>	Storyteller_FOC_AB	Principalement Pnarr/ beaucoup de PC
<i>La chemise magique (tL)</i>	Storyteller_FOC_RN	Principalement Pnarr/ Quelques PC
<i>La compagnie des loups (M)</i>	Storyteller_FOC_JG	Principalement Pnarr/ peu de PC
<i>Le mari trompé (L)</i>	Storyteller_FOC_GB	Principalement Pnarr/ peu de PC

OWT

<i>Titre</i>	<i>Conteur</i>	<i>Patron des temps verbaux</i>
Lo vielh e los tres volurs (M)	Arnaudin	Principalement Pnarr/ quelques PS
Compair Loison e la mair deu vent (tL)	Arnaudin	Principalement Pnarr/ quelques PS
Lo bon diu e lo diable (L)	Arnaudin	Pnarr/ PS
La pelha arrecastada (M)	Arnaudin	Pnarr/ PS
Lo gojat (M)	Lambert	Pnarr/ PS
Peton-Petet (S)	Lambert	Principalement PS
Las sorcièras (M)	Lambert	Principalement PS
Joan Lo pigre (M)	Bladé	Principalement PS
Los dus bessons (M)	Bladé	Principalement PS
La flauta (M)	Bladé	Principalement PS
Pèl-d'Ase (L)	Bladé	Principalement PS
Lo voiatge de Joanòt (M)	Bladé	Principalement PS
Perqué las fumèlas ... (M)	Lalanne	Principalement PS
La formiga que s'enanava a la fèsta de Sabardu (C)	Séré	Principalement PS
Las tres galinetas (M)	Lambert	Principalement PS/ Quelques Pnarr
Lo lop malaut (M)	Bladé	Principalement PS/ Quelques Pnarr
Lo maset (M)	Montel	Principalement PS/ Quelques Pnarr
Turlendú (M)	Montel	Principalement PS/ Quelques Pnarr
La formiga que s'enanava a ... (M)	Séré	Principalement/ 1 Pnarr

9. Analyse : (iii) de la présence/quantité d'alternance *des temps verbaux* dans les quatre sous-corpus, prenant en compte le concept de 'degrés d'oralité' – basé sur des paramètres tels que les sources des histoires, le canal de transmission, le contexte de performance.

	OOT	OOC	FOC	OWT
<i>Pourcentage des narrations avec alternance des TV (textes avec alternance/nombre de textes)</i>	88.5% (22/26)	61.5% (8/13)	61.5% (8/13)	52.6% (10/19)
<i>Taux moyen d'alternance des TV relatif au nombre de propositions narratives avec verbe</i>	1 alternance pour 4,1 propositions	1 alternance pour 6,7 propositions	1 alternance pour 9,7 propositions	1 alternance pour 3,6 propositions

<i>(seulement les textes avec alternance)</i>				
<i>Fourchette des taux d'alternance pour chaque sous-corpus (i.e. 1 alternance pour X propositions narratives seulement dans les textes avec alternance)</i>	<i>X= entre 2 et 13</i>	<i>X= entre 4 et 50</i>	<i>X= entre 4 et 55</i>	<i>X = entre 2 et 9</i>

10. Quelques réflexions sur l'alternance *des temps verbaux* dans différents types de conte.

<i>Type de conte</i>	<i>Pourcentage des textes avec alternance (uniquement les textes avec alternance/nombre de textes)</i>	<i>Taux moyens d'alternance relatif au nombre de propositions narratives avec verbe (uniquement les textes avec alternance)</i>
Conte d'animaux	80% (8/10)	<i>1 alternance pour 4,8 propositions</i>
Conte de l'ogre dupé	75% (3/4)	<i>1 alternance pour 3,8 propositions</i>
Conte facétieux	70% (7/10)	<i>1 alternance pour 4,3 propositions</i>
Conte formulaire	75% (3/4)	<i>1 alternance pour 3,9 propositions</i>
Conte merveilleux	73% (11/15)	<i>1 alternance pour 5,4 propositions</i>
Total Conte	74% (32/43)	<i>1 alternance pour 4,8 propositions</i>
Récit contemporain	100% (4/4)	<i>1 alternance pour 9,3 propositions</i>
Récit facétieux	100% (5/5)	<i>1 alternance pour 9,6 propositions</i>
Récit légendaire	58% (7/12)	<i>1 alternance pour 4,8 propositions</i>
Total Récit	76% (16/21)	<i>1 alternance pour 7,6 propositions</i>

11. Conclusion : pertinence du concept de 'degrés d'oralité' pour l'analyse de l'emploi des temps verbaux dans la narration orale. Paramètres fondamentaux à prendre en considération dans le contexte de la littérature orale.

Impersonal constructions in Paman: Two puzzles in pronominal morphosyntax

Jean-Christophe Verstraete, University of Leuven & Australian National University

This presentation examines a set of impersonal constructions found in four Paman (< Pama-Nyungan) languages of Cape York Peninsula, Australia (first described in Verstraete 2011a, b). The core function of these structures – defined by the obligatory absence of nominative forms in the pronoun complex – relates to involuntary physical experience, as in (1) below. In one language, moreover, the structures also function as a specialized voice mechanism for human undergoers, highlighting involuntary involvement, as in (2).

(1) Rimanggudinhma (Lamalamic < Paman < Pama-Nyungan)

noka rhi-m=thum

neck spear-PRS-1SG.DAT

‘I have hiccups.’ (Lit. ‘(it) spears me in the neck’) [Godman 1993: 91]

(2) Umpithamu (Middle Paman < Paman < Pama-Nyungan)

kali-n=athuna yinthu

carry-PST=1SG.GEN inside

‘I was carried inside [after fainting].’ (Lit. ‘(They) carried me inside’) [fieldnotes]

The form and function of these structures presents a knotty problem at the interface between description and theory (as well as lexicon and morphosyntax), which I think the dedicatee of this conference would have enjoyed. In this presentation, I try to disentangle two knots, one diachronic and one typological.

The diachronic argument relates to the origins of the impersonal structures in Umpithamu, which are exceptional not just within the Middle Paman subgroup of Paman languages, but also within the broader region. I show that the structures developed as part of a thorough restructuring of pronominal morphosyntax in the language under the influence of its neighbours from the Lamalamic subgroup of Paman (like Rimanggudinhma in (1)). Apart from the historical-comparative and ethnographic evidence used in earlier work (Verstraete 2012), I also provide lexical-semantic and typological evidence to show how a voice mechanism could have developed out of the core impersonal function of the structures.

The typological argument relates to the nature of the resulting construction. Impersonal structures of the type in (1) are found elsewhere in Australia and beyond (see, for instance, Evans 2004, Walsh 1987 and Malchukov 2008), but a category covering both this and a voice mechanism at first sight looks harder to deal with from a typological perspective. However, I use recent work by Inglese (2020, ms) on the typology of the middle voice to show that a structure covering (1) and (2) can be analysed as a type of middle voice, which typically combines a lexical function with a morphosyntactic one. The peculiar diachronic development described here, moreover, demonstrates that lexical uses are not always historically secondary in middle voice systems, as is commonly

assumed (e.g. Kaufmann 2007), but can themselves be the basis for the development of such a system (see Inglese & Verstraete ms).

References

- Evans, N. 2004. Experiencer objects in Iwaidjan languages (Australia). In P. Bhaskararao & K. Subbarao (eds.), *Non-nominative subjects. Volume 1*, 169-192. Amsterdam: Benjamins.
- Godman, I. 1993. *A sketch grammar of Rimanggudinhma. A language of the Princess Charlotte Bay region of Cape York Peninsula*. BA Hons thesis, University of Queensland.
- Inglese, G. 2020. *The Hittite middle voice: Synchrony, diachrony, typology*. Leiden: Brill.
- Inglese, G. Ms. The typology of middle voice systems. Manuscript.
- Inglese, G. & J-C. Verstraete. Ms. Evidence against unidirectionality in the emergence of middle voice systems: Two case studies from Anatolian and Paman. Manuscript.
- Kaufmann, I. 2007. Middle voice. *Lingua* 117: 1677-1714.
- Malchukov, A. 2008. Split intransitives, experiencer objects, and ‘transimpersonal’ constructions: (re-)establishing the connections. In M. Donohue & S. Wichmann (eds.), *The typology of semantic alignment*, 76-100. Oxford: Oxford University Press.
- Verstraete, J-C. 2011a. Impersonal constructions in Umpithamu and the Lamalamic languages. In A. Malchukov & A. Siewierska (eds.), *Impersonal constructions: A cross-linguistic perspective*, 601-618. Amsterdam: Benjamins.
- Verstraete, J-C. 2011b. Experienced action constructions in Umpithamu: Involuntary experience, from bodily processes to externally instigated actions. *Cognitive Linguistics* 22: 275-302.
- Verstraete, J-C. 2012. Contact-induced restructuring of pronominal morphosyntax in Umpithamu (Cape York Peninsula, Australia). *Diachronica* 29: 326-358.
- Walsh, M. 1987. The impersonal verb construction. In R. Steel & T. Threadgold (eds.), *Language topics. Essays in honour of Michael Halliday*, 425-438. Amsterdam: Benjamins.