

## **Le redoublement du sujet en français parlé\* Motivations rythmiques**

Anke Grutschus, Sandra Schwab

Université de Regensburg, Université de Fribourg/Université de Zurich

[kontakt@anke-grutschus.de](mailto:kontakt@anke-grutschus.de), [sandra.schwab@unifr.ch](mailto:sandra.schwab@unifr.ch)

Reçu le 5/3/2020, accepté le 16/5/2020, publié le 5/11/2020 selon les termes de la licence *Creative Commons Attribution 4.0 International* (CC BY 4.0)

**Résumé** : Le phénomène du redoublement du sujet est imputable à un certain nombre de facteurs extralinguistiques (le registre, la provenance géographique du locuteur, etc.) aussi bien qu'intralinguistiques (la personne verbale, le type de phrase, etc.). La présente contribution se propose d'étudier si la présence du redoublement du sujet peut également s'expliquer par des paramètres de nature rythmique tels que le nombre de syllabes dans le sujet ou dans le verbe. L'étude met en regard des contextes doublés et non doublés dans le français valaisan. L'analyse prend uniquement en compte des contextes où le locuteur a effectivement le choix entre les deux variantes, ce qui exclut les dislocations à gauche, les sujets à la 1<sup>ère</sup> et à la 2<sup>ème</sup> personne ou les expressions figées. L'évaluation statistique porte sur l'influence de la longueur du sujet ainsi que sur l'impact de la différence entre le sujet et le verbe en termes de poids syllabique.

**Abstract**: French subject doubling can be attributed to extralinguistic factors (e.g. register, geographic origin of the speaker, etc.) as well as to linguistic factors (e.g. verbal person, sentence type, etc.). The present contribution investigates whether the realization of 'redundant' preverbal subject clitics could be explained by rhythmic factors such as the number of syllables of the noun phrase or the verb phrase. The analysis compares duplicate and simple structures in Valais French, only taking into account contexts where speakers can actually choose between the two variants: left dislocations, first or second person subjects and idioms are thus excluded. The statistical evaluation of the results focuses on the influence of subject length as well as on the impact of the difference in length between subject and object on the doubling rate.

---

\* Nous remercions Olivier Ostrini pour l'extraction manuelle des séquences sujet-verbe du corpus *OFROM* ainsi que Pierre Ménétreay pour l'extraction automatique des énoncés du corpus *OFROM*.

## 1 Introduction

[1] Le redoublement du sujet, qui fait alterner l'énoncé en (1a) avec sa variante non-redoublée en (1b), compte parmi les phénomènes morphosyntaxiques les mieux étudiés du français parlé (Ashby 1980 ; Sankoff 1982 ; Coveney 2005 ; Berrendonner 2007 ; Avanzi 2012).

(1a) Mon frère *il* chante bien.

(1b) Mon frère chante bien.

[2] Les études antérieures se sont focalisées sur des facteurs intra- et extralinguistiques pour expliquer la présence du redoublement, sans pour autant s'appuyer systématiquement sur des corpus. La présente étude se propose de prendre en compte un facteur linguistique négligé jusqu'à maintenant en partant de l'hypothèse selon laquelle des facteurs prosodiques pourraient jouer un rôle dans la variation. S'agissant d'une première étude de cas, l'analyse ne prendra en compte qu'un nombre restreint de facteurs de nature métrique, à savoir le nombre de syllabes dans le sujet et dans le groupe verbal ainsi que la différence entre les deux.

[3] Notre hypothèse sera vérifiée sur la base d'un corpus du français parlé qui représente une variété avec un taux de redoublement relativement élevé, à savoir le français valaisan. Le choix de la variété valaisanne s'explique avant tout par le besoin de disposer d'une masse critique d'occurrences provenant d'un nombre suffisamment élevé de locuteurs. C'est seulement dans un deuxième temps qu'il peut paraître utile d'isoler des facteurs de redoublement spécifiques à une variété en particulier, par exemple en comparant deux variétés. Après un bref état de l'art des facteurs de redoublement identifiés dans les études antérieures (§ 2), on présentera brièvement le corpus utilisé (§ 3.1). Étant donné que l'analyse prendra uniquement en compte des contextes où le locuteur a effectivement le choix entre les deux variantes, il faut exclure un certain nombre de contextes (§ 3.2). L'analyse se propose d'abord de déterminer l'influence de la complexité morphophonologique du sujet sur le taux de redoublement (§ 4.2.1), avant de se focaliser sur l'impact de la différence en termes de poids syllabique entre le sujet et le groupe verbal (§ 4.2.2).

## 2 Facteurs de redoublement

### 2.1 Redoublement vs dislocation à gauche

[4] Avant de dresser un état de l'art des travaux analysant de possibles facteurs de redoublement, il convient de distinguer le redoublement du sujet d'un phénomène étroitement apparenté, à savoir la dislocation à gauche. En dépit de nombreux points communs, les deux phénomènes présentent notamment des différences au niveau de la structure informationnelle : alors que la dislocation a une valeur topicalisante et qu'elle est utilisée pour mettre un élément en relief, le re-

doublément du sujet constitue une construction non marquée dans ce contexte. Au niveau prosodique, la dislocation à gauche est caractérisée par la présence d'un « détachement prosodique » (Berrendonner 2007 : 92). Ainsi, dans l'exemple (2), l'élément disloqué *les jeunes en Allemagne* constitue le topique de l'énoncé. Il est suivi d'une frontière qui le sépare du reste de l'énoncé. L'exemple (3) montre une structure redoublée, dont le syntagme nominal (SN) *ma mère* n'est pas mis en relief. Aucune frontière prosodique n'est détectable entre le constituant doublé et le clitique *elle*.

- (2) ((Les jeunes en Allemagne) (*ils* ont peur)). (Gabriel & Rinke 2010 : 107)  
 (3) ((Ma mère elle m'a téléphoné)). (Gabriel & Rinke 2010 : 107)

[5] Cependant, des études sur corpus ont montré que dans l'usage réel, la distinction entre une construction disloquée d'un côté et un sujet redoublé de l'autre est loin d'être claire (Avanzi 2012 : 178). Il paraît donc plus pertinent de partir de l'idée d'un continuum entre deux pôles, respectivement formés par la dislocation et le redoublement (Buthke, Sichel-Bazin & Meisenburg 2014 ; Stark 1997). Par conséquent, notre analyse de corpus ne tentera pas de distinguer systématiquement les deux phénomènes, mais exclura seulement les cas prototypiques de dislocation (§ 3.2, exemples 11 et 12).

## 2.2 Facteurs linguistiques

[6] De nombreux travaux antérieurs ont permis d'identifier un certain nombre de facteurs linguistiques pouvant entraîner un taux de redoublement du sujet plus élevé. Parmi ces facteurs, on compte d'abord la personne grammaticale ainsi que le nombre grammatical : le redoublement est beaucoup plus fréquent pour les sujets à la 1<sup>ère</sup> et à la 2<sup>ème</sup> personne grammaticale que pour la 3<sup>ème</sup> personne (Koch 1993 : 182). Au pluriel, le taux de redoublement est plus important qu'au singulier (Zahler 2014 : 363). Le type de proposition semble également jouer un rôle, étant donné que le redoublement est plus fréquent en proposition principale qu'en subordonnée (Nagy, Blondeau & Auger 2003 ; Auger & Villeneuve 2010 : 77).

[7] Par ailleurs, le cotexte immédiat peut avoir une incidence sur le redoublement : plus il y a d'éléments entre le sujet et le verbe conjugué, que ce soient des adverbes ou des clitiques préverbaux comme *se* ou *en*, plus il y a de redoublements (Zahler 2014 : 363 ; Auger & Villeneuve 2010 : 78). En outre, on a pu constater une influence des caractéristiques du sujet, avec plus de redoublements pour les sujets définis, spécifiques et/ou animés (Auger & Villeneuve 2010 : 79 ; Gabriel & Rinke 2010 : 112 ; Zahler 2014 : 362s.).

[8] Les sujets syntaxiquement complexes semblent également favoriser le redoublement, avec un taux de redoublement plus important pour les SN compre-

nant des compléments tels que des subordinées relatives (Auger & Villeneuve 2010 : 77). Enfin, le redoublement et le *ne* de négation semblent s'exclure mutuellement (Zahler 2014 : 364). Ce dernier facteur n'est évidemment pas de nature linguistique à proprement parler, mais montre plutôt l'incidence du registre : comme nous le montrerons dans la section suivante, le registre informel est caractérisé aussi bien par un nombre important de redoublements du sujet que par un taux élevé d'absence du *ne* de négation (Massot 2010 ; Zribi-Hertz 2011). Les travaux existants, qui ne s'appuient pas systématiquement sur une analyse statistique de corpus suffisamment larges du français parlé, présentent par ailleurs souvent l'inconvénient de regrouper dislocations et structures redoublées, ce qui rend difficile la comparaison des résultats.

### 2.3 Facteurs extralinguistiques

[9] Parmi les principaux facteurs extralinguistiques favorisant le redoublement du sujet, on compte d'abord la situation de communication, avec un taux de redoublement beaucoup plus important en contexte informel (Sankoff 1982 : 85 ; Massot 2010). Certains facteurs sociolinguistiques ont également une incidence sur le taux de sujets redoublés. Ainsi, l'étude de Coveney (2005 : 106) a montré un taux de redoublement plus important pour les locuteurs jeunes. Quant à l'influence du sexe, les résultats divergent considérablement : alors que dans son analyse du corpus d'Orléans, Ashby (1980) constate une préférence pour la variante redoublée chez les locuteurs masculins, le corpus d'Auger & Villeneuve (2010) ne révèle pas de différence liée au sexe. Le rôle de l'origine sociale semble cependant avéré, avec une préférence claire pour le redoublement par les locuteurs d'origine plus modeste (Sankoff 1982 : 85).

[10] Enfin, l'origine régionale des locuteurs constitue un autre facteur extralinguistique important. Pour de très nombreuses variétés diatopiques, on a constaté un taux variable de redoublement. Ainsi, les variétés parisiennes (Zahler 2014), picardes (Coveney 2003), québécoises (Auger & Villeneuve 2010 ; Sankoff 1982), acadiennes (Beaulieu & Balcom 1998), louisianaises (Girard 2010) et maghrébines (Queffelec 2000 : 790 ; Roberge 1990 : 97s.) semblent privilégier le redoublement. La présence du redoublement en français valaisan n'a pas encore été analysée, mais la seule étude portant sur le français suisse (Fonseca-Greber & Waugh 2003, qui malheureusement ne précisent pas la provenance des locuteurs de leur corpus) atteste également un taux élevé pour ce dernier. Tout comme nous l'avions déjà constaté pour les facteurs linguistiques, les analyses des facteurs extralinguistiques ne sont pas toujours systématiquement corroborées par des évaluations statistiques de corpus.

### 2.4 Facteurs prosodiques ?

[11] Jusqu'à présent, l'incidence de facteurs prosodiques sur la réalisation du redoublement n'a pas encore été prise en compte, même si on peut trouver des re-

marques indiquant une possible influence. Ainsi, Massot (2008 : 334) exprime l'intuition suivante concernant l'alternance entre la structure SVO 'traditionnelle' et une structure à sujet disloqué/redoublé, telle qu'elle se présente fréquemment en registre informel : « Il me semble qu'une hypothèse corrélant l'opposition de type entre SV(O) et (T[opique]) pro[nom]-V[erbe] (X) (A[nti]T[opique]) à des divergences prosodiques est prometteuse ». Avanzi (2012 : 179), dans son étude du marquage prosodique des dislocations à gauche, suggère également la possibilité d'une influence du contexte métrique du syntagme (non-)disloqué.

[12] La présente étude se concentre sur deux facteurs de nature rythmique, ce qui n'exclut évidemment pas la pertinence d'autres facteurs, notamment de nature mélodique, tels que la présence d'un accent initial au début du groupe verbal. D'un côté, elle prendra en compte le poids syllabique du sujet, en émettant l'hypothèse que le redoublement est plus fréquent pour les sujets métriquement 'lourds' tel que dans l'exemple suivant :

- (4) *mon tonton et ma tata* ils les regardaient trop méchamment tu vois genre euh  
(OFROM, unine15b27d\_966)

Cette hypothèse exploite l'observation selon laquelle les sujets syntaxiquement complexes semblent favoriser la reprise du sujet par un clitique (§ 2.2). Elle pré-suppose que la complexité syntaxique équivaut à une complexité morphophonologique (§ 4.2.1), qui se traduit par un nombre élevé de syllabes.

[13] De l'autre côté, nous souhaitons étudier l'influence du principe eurythmique sur la réalisation du redoublement. Partant du principe que les locuteurs tendent à produire des phrases d'une longueur comparable (Dell 1984 ; Martin 1987 ; Delais-Roussarie 1995), nous explorons l'hypothèse selon laquelle la présence ou l'absence du clitique contribue à l'eurythmie des groupes accentuels comprenant le sujet et le verbe. Le redoublement interviendrait donc pour équilibrer le poids syllabique des deux groupes. Ainsi, le clitique *elle* dans l'exemple (5) contribue à équilibrer le nombre de syllabes du sujet (S) et du groupe verbal (GV). En revanche l'exemple (6) présente 'd'emblée' une eurythmie entre sujet et groupe verbal, ce qui fait que le recours au clitique n'est pas nécessaire.

- (5) [la mezzanine]<sub>S</sub> [elle est très haute]<sub>GV</sub> (OFROM, unine11c18d\_187)  
(6) [le pilote]<sub>S</sub> [nous disait]<sub>GV</sub> on va atterrir à Genève (OFROM, unine11c05m\_169)

Nous cherchons à vérifier ces hypothèses au moyen d'une analyse de corpus, dont nous présenterons les principes ci-dessous.

N° loc.	Code locuteur	Sexe	Âge	Profession	Nb. occ.	% re-doubl.
L1	unifr11-cra	F	19	Étudiante	13	100%
L2	unifr14-jba	F	22	Infirmière	7	29%
L3	unine11-csa	F	22	Étudiante	10	30%
L4	unine11-smb	H	22	Étudiant	12	83%
L5	unine15-083	F	22	Étudiante	13	31%
L6	unine15-936	F	22	Étudiante	9	78%
L7	unine11-apa	F	23	Étudiante	10	90%
L8	unine15-935	F	23	Étudiante	6	50%
L9	unine11-fdb	H	24	Laborant chimie	5	40%
L10	unine15-070	F	24	Junior manager	7	71%
L11	unine15-937	F	25	Étudiante	5	100%
L12	unine11-jra	H	27	Facteur	12	50%
L13	unine15-092	H	27	Marketing	5	25%
L14	unine15-104	F	27	Pédagogue	16	13%
L15	unine11-sda	F	32	Femme au foyer	12	42%
L16	unifr11-aaa	F	34	Traductrice	7	43%
L17	unine11-cma	H	38	Employé de commerce	11	82%
L18	unine11-gma	H	38	Employé de commerce	16	94%
L19	unine11-fda	H	42	Infirmier	4	25%
L20	unine15-091	F	43	Infirmière	3	100%
L21	unine15-090	H	44	Ambulancier	12	83%
L22	unine11-yfa	F	47	Secrétaire patronale	10	10%
L23	unine11-eza	F	51	Assistante de direction	21	10%
L24	unine14-pra	F	51	Maîtresse enfantine	12	17%
L25	unine11-gpa	H	52	Instituteur	11	0%
L26	unine11-joa	H	56	Avocat	12	9%
L27	unine11-jma	F	57	Employée de commerce	9	33%
L28	unine11-pgb	F	58	Retraitée enseignante	15	0%
L29	unine11-pga	H	60	Juge	9	13%
L30	unine15-105	F	60	Responsable de ménage	12	50%
L31	unine11-gaa	F	77	Retraitée empl. de bureau	13	23%
L32	unine11-lva	F	78	Retraitée pharmacienne	11	27%
L33	unine11-vra	H	79	Retraité pharmacien	12	42%
L34	unifr11-dla	F	80	Retraitée	15	93%
L35	unine11-cpa	F	80	Retraitée sommelière	10	30%
L36	unifr14-nva	H	88	Retraité ouvrier chantier	7	71%
L37	unine15-096	F	94	Retraitée	12	83%
L38	unifr11-dba	H	NR	NR	20	55%
L39	unifr12-tda	F	NR	NR	3	33%

Tableau 1: locuteurs du corpus *OFROM/Valais* (novembre 2018)

### 3 Étude de corpus

#### 3.1 Présentation du corpus

[14] Étant donné que l'étude de phénomènes syntaxiques tels que le redoublement du sujet ne peut s'effectuer que sur la base d'un corpus suffisamment large, nous avons décidé de nous appuyer sur un corpus préexistant en étudiant les enregistrements de locuteurs originaires du Valais contenus dans le *Corpus oral de français de Suisse romande (OFRON)*. Le corpus est constitué d'entretiens guidés d'une durée entre 30 et 40 minutes qui ont été effectués, pour la partie valaisanne du corpus, entre 2011 et 2015. Seule une dizaine de minutes de chaque entretien a été transcrite, de sorte que le nombre de token de la partie valaisanne s'élève à environ 111000. Les thèmes abordés concernent pour la plupart la vie quotidienne des interviewés, le registre employé peut être caractérisé d'informel (Avanzi, Béguelin & Diémoz 2016 : 4). Il faut cependant noter que la nature de la relation entre les intervieweurs et les interviewés, qui pourrait contribuer à déterminer le degré de familiarité des échanges avec plus de précision, n'est pas renseignée. Le tableau 1 (voir supra), qui comprend uniquement les locuteurs dont plus d'un exemple a été retenu pour l'analyse, montre que la répartition des 39 locuteurs est relativement équilibrée en ce qui concerne les tranches d'âge, le sexe et le niveau socio-éducatif.

#### 3.2 Considérations méthodologiques

[15] Afin de pouvoir isoler le rôle des facteurs prosodiques, il convient de prendre uniquement en considération des énoncés où les locuteurs ont effectivement le choix entre la variante 'simple' et la variante redoublée. Pour exclure l'incidence des facteurs linguistiques connus (§ 2.2), favorisant soit le redoublement, soit la réalisation 'simple' du sujet, l'analyse prendra uniquement en compte des énoncés avec sujet lexical (SN) et verbe conjugué, aussi bien simples que redoublés. Nous excluons les constructions du type SN + *c'est* (voir l'exemple 7), étant donné qu'elles apparaissent souvent dans des contextes prédicatifs ou correspondent à des constructions *hanging topic* (Dufter & Gabriel 2016) et ne sont donc pas susceptibles d'être remplacées par la variante simple.

(7) l'argot c'est quelque chose qui a disparu (OFRON, unifr11a02m\_29\_unifr11-dba)

[16] Pour les mêmes raisons, nous excluons les constructions figées comme *les bras m'en tombaient*. Par ailleurs, les phrases avec des pronoms indéfinis comme *tout le monde* ou *quelqu'un* à la place du sujet ont été écartées, car même si l'on trouve des constructions comme *tout le monde il est beau* dans des contextes d'oralité simulée (voir le titre de la comédie éponyme), leur emploi en français parlé paraît marginal et n'est pas documenté dans notre corpus. Seuls les

pronoms indéfinis de la série *un(s)/une(s)*, très courants en français valaisan, ainsi que les pronoms indéfinis *certain(e)s* et *la plupart* ont été pris en considération, étant donné qu'une variation entre la structure 'simple' (voir l'exemple 8) et la structure redoublée (voir l'exemple 9) est tout à fait possible.

- (8) un est au collège (*OFROM*, unine11c10m\_223\_unine11-cpa)  
 (9) une elle sait pas faire (*OFROM*, unine15b57m\_184\_unine15-105)

En outre, les phrases avec hésitation ou pause dans le sujet et entre le sujet et le verbe (voir l'exemple 10) n'ont pas été prises en compte.

- (10) les gens ils tre/ ils te f/ ils se foutent de ta gueule (*OFROM*, unine15z62d\_107\_unine15-936)

Par ailleurs, les exemples avec une hésitation au sein du groupe verbal ou avec un groupe verbal incomplet ont été écartés de la deuxième partie de l'analyse, qui compare le poids syllabique du sujet avec celui du groupe verbal. Ainsi, l'exemple en (11) n'a pas pu être pris en considération car l'indication de lieu a été effacée lors de l'anonymisation des enregistrements.

- (11) un de ses fils il est à # (*OFROM*, unine11c10m\_065\_unine11-cpa)

[17] Enfin, nous avons exclu de l'analyse des cas 'clairs' de dislocation à gauche. Pour les identifier, nous nous sommes appuyées sur deux types de critères (§ 2.1) : d'un côté, un critère syntaxique, stipulant qu'il ne devait y avoir aucun constituant entre sujet et verbe conjugué, à l'exception d'autres clitiques comme *se* ou *en*. Ainsi, l'exemple en (12) a été écarté en raison du constituant *en fait*.

- (12) euh ma cousine en fait elle était déjà sur place (*OFROM*, unine12a04m\_136, unine12-avb)

De l'autre côté, nous nous sommes basées sur un critère de nature prosodique, selon lequel les dislocations prototypiques étaient marquées soit par une pause (voir l'exemple 13), soit par un coup de glotte, dont la présence a été évaluée de manière perceptive (voir l'exemple 14).

- (13) mais mon papa [0,21 sec.] il était très intelligent à l'école (*OFROM*, unifr11b02m\_378\_unifr11-dla)  
 (14) mais la vie elle [ʔɛl] a tellement changé (*OFROM*, unifr11b02m\_054\_unifr11-dla)

### 3.3 Analyse des données

[18] Dans le cadre de l'analyse, nous avons identifié manuellement, sur la base des transcriptions orthographiques, les sujets (lexicaux) redoublés et non redoublés. Ensuite, le nombre de syllabes phonétiques dans le sujet a été déterminé grâce à un examen acoustique des énoncés, qui a notamment permis de prendre en compte les élisions effectivement réalisées. Ainsi, le sujet de l'exemple (15) comprend deux syllabes, tout comme celui de l'exemple (16), dans lequel le *e* muet a été élidé.

- (15) *le gars* il nous r'posait une question on répondait tu vois (OFROM, uni-fr11a02m\_180\_unifr11-dba)  
 (16) *l'machin* il était tout branlant (OFROM, unine11b09m\_211\_unine11-smb)

[19] Afin de dégager des eurythmies entre sujet et groupe verbal, nous cherchons à comparer le nombre de syllabes du groupe accentuel comportant le sujet avec celui du groupe accentuel comportant le verbe. Par *groupe accentuel*, nous entendons une « suite de syllabes dont la dernière est marquée par un accent primaire », tout en tenant compte de la « réalité syntactico-sémantique du groupe » (Lacheret-Dujour & Beaugendre 1999 : 45). L'annotation des groupes accentuels a été effectuée avec l'aide du logiciel *Analor* (Avanzi et al. 2011), qui a été conçu pour détecter des prééminences syllabiques de manière semi-automatique et dont nous nous sommes servies pour déterminer les accents primaires de chaque groupe accentuel.

[20] Pour ce faire, les transcriptions en orthographe standard des occurrences sélectionnées (voir *supra*) ont d'abord été phonétisées puis alignées semi-automatiquement sur les sons à l'aide du logiciel *EasyAlign* (Goldman 2011). Les résultats de la segmentation automatique ont été vérifiés manuellement et modifiés le cas échéant. Les transcriptions phonétiques ainsi obtenues ainsi que les fichiers son correspondants ont ensuite été traités à l'aide d'*Analor*, qui a détecté les prééminences dans les segments analysés. La figure 1 illustre la détection des prééminences de l'énoncé (17) : le logiciel a détecté quatre prééminences au total, dont trois correspondent à des accents primaires à la fin des différents groupes accentuels. La quatrième prééminence (*BArrage*) correspond à l'allongement de la pénultième, très caractéristique du français de Suisse romande (Avanzi et al. 2012).

- (17) et pis des gens ils restaient quand même sur le barrage (OFROM, uni-fr14a04m\_368\_unifr14-nva)

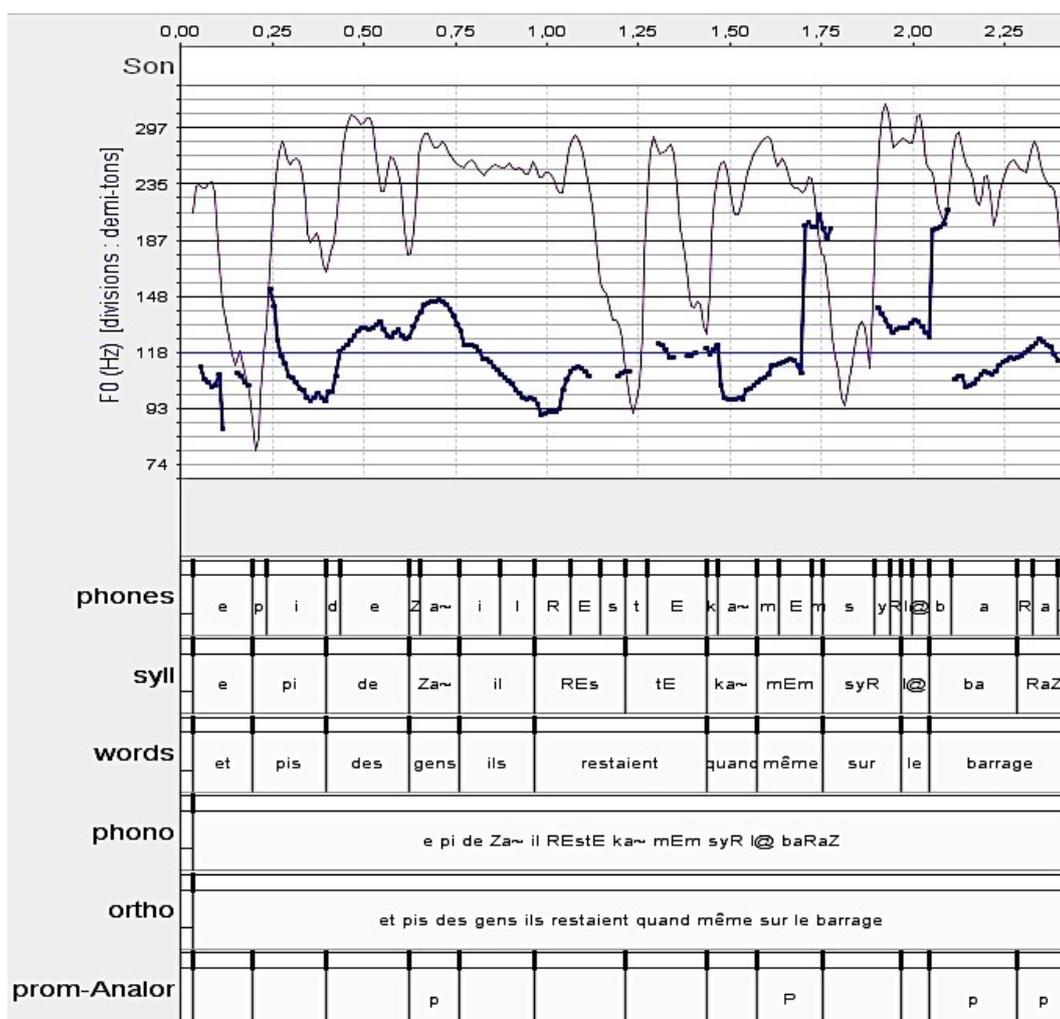


Figure 1: copie d'écran *Anamor*. Détection de proéminences de l'énoncé *et pis des gens ils restaient quand même sur le barrage* (OFROM, unifr14a04m\_368\_unifr14-nva)

Sur la base de ce résultat, nous avons annoté les groupes accentuels comme indiqué en (17'). L'analyse porte donc uniquement sur deux groupes accentuels, à savoir celui qui contient le sujet et le groupe qui le suit immédiatement. Ce deuxième groupe accentuel comprend le verbe ainsi que, les cas échéant, le clitique.

(17') [et pis des gens]<sub>GA\_sujet</sub> [ils restaient quand même]<sub>GA\_groupe\_verbal</sub>

[21] La hiérarchie des proéminences indiquée par *Anamor*, qui différencie entre différents degrés de saillance syllabique (*p* vs *P*, voir la figure 1) n'a pas été

prise en compte. La détection des proéminences s'est parfois avérée erronée, dans le sens où le logiciel a parfois détecté des proéminences là où il n'y en avait pas ou inversement, que le programme n'a pas réussi à détecter des proéminences qui étaient pourtant perceptibles à l'écoute. Les résultats ont donc été contrôlés et, lorsqu'ils paraissaient erronés, soumis à une analyse perceptive par les auteures. Suite à l'annotation en groupes accentuels, le nombre de syllabes effectivement réalisées a été déterminé pour chaque groupe.

[22] Les données ont été examinées en deux temps au moyen de modèles de régression logistique avec effets mixtes. La variable dépendante était la même dans tous les modèles, à savoir la présence/l'absence d'un sujet redoublé. Par ailleurs, l'effet aléatoire, caractéristique d'un modèle à effets mixtes, correspondait dans tous les modèles à la variable *Locuteur*. L'utilisation de modèles mixtes nous permettait ainsi de prendre en compte la variabilité qui pouvait être observée entre les différents locuteurs de notre corpus. Nous avons ainsi tout d'abord examiné l'influence du sujet syntaxique sur la présence de redoublement à l'aide de deux modèles dans lesquels le prédicteur était la complexité syntaxique et la longueur (en syllabes) du sujet, respectivement. Nous avons ensuite étudié l'impact de la différence de poids syllabique entre le sujet et le groupe verbal sur la présence du redoublement à l'aide d'un modèle dans lequel le prédicteur était la différence de syllabes entre le groupe accentuel (GA) comprenant le sujet et le GA comprenant le groupe verbal. La significativité des effets a été évaluée en comparant le modèle contenant le prédicteur en question avec le modèle sans ce dernier. Comme il est de coutume pour ce genre de modèles, les valeurs statistiques que nous rapportons dans les sections suivantes sont des valeurs de chi-carré. Notons encore que, bien que les résultats présentés dans les sections suivantes soient exprimés en nombre d'occurrences ou en pourcentages, toutes les analyses ont été effectuées sur les valeurs brutes de présence/absence de redoublement.

## 4 Résultats

### 4.1 Observations préliminaires

[23] Avant de présenter les résultats de l'analyse, il convient de souligner certaines caractéristiques des données. La partie analysée du corpus *OFROM* comporte 407 exemples, le taux global de redoublement s'élève à 47 %. La variabilité entre les locuteurs est très importante : Comme l'indique la figure 2, on trouve aussi bien des locuteurs qui ne redoublent que très rarement, voire jamais, alors que d'autres privilégient la variante redoublée dans la majorité des cas, voire de manière systématique.

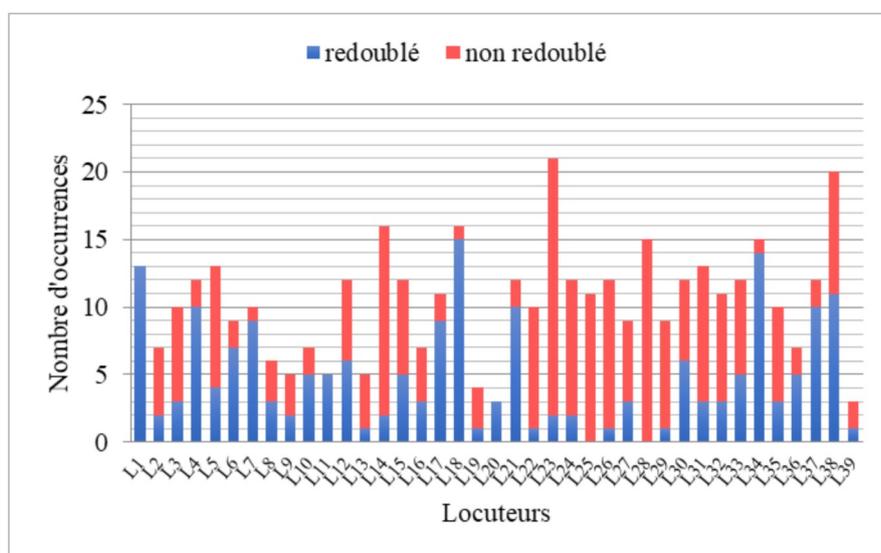


Figure 2: redoublement chez les locuteurs

## 4.2 Résultats

[24] Dans le cadre de l'analyse, nous avons d'abord cherché à vérifier notre première hypothèse concernant l'impact de la longueur du sujet sur la présence du redoublement (§ 4.2.1). Par la suite, nous avons étudié l'incidence du facteur eurythmique (i.e. différence de syllabes entre le GA sujet et le GA groupe verbal) sur la présence ou l'absence du redoublement (§ 4.2.2).

### 4.2.1 Longueur du sujet

[25] Comme nous l'avons indiqué dans la section 2.2, la complexité syntaxique du sujet compte parmi les facteurs linguistiques favorisant le redoublement du sujet. Afin de déterminer si ce facteur de nature syntaxique ne cache pas plutôt un facteur de nature rythmique dans le sens où le redoublement interviendrait de préférence après des sujets lexicaux comprenant un nombre élevé de syllabes, nous avons étudié l'impact de la complexité morphophonologique du sujet sur le taux de redoublement. Selon les résultats des études antérieures, nous nous attendons d'un côté à un taux de redoublement plus élevé pour des sujets syntaxiquement complexes et de l'autre côté à une augmentation du taux de redoublement en fonction du nombre croissant de syllabes du sujet.

[26] Afin d'établir le degré de complexité syntaxique qui caractérise les sujets de notre corpus, nous nous sommes appuyées sur la catégorisation établie par Auger & Villeneuve (2010 : 77). Il s'agit d'une catégorisation très fine qui convient d'autant mieux aux spécificités des sujets de notre corpus que la gamme de complexité syntaxique n'y est pas très large : Étant donné que nous avons

seulement pris en compte des sujets lexicaux, les sujets peu complexes comme les pronoms personnels ont été exclus d'avance. Le tableau 2 rassemble les différents types de sujet présents dans le corpus. Ils ont été classés en fonction de leur complexité syntaxique croissante, en commençant par la classe fermée des pronoms indéfinis, en passant par des sujets à complexité moyenne comme des syntagmes comprenant un adjectif ou un participe pour aboutir à des syntagmes très complexes. Le tableau 2 montre que dans trois quarts des cas, les sujets de notre corpus correspondent à un syntagme nominal 'simple'. Notre corpus ne contient aucun exemple d'un sujet comprenant une relative, ce qui s'explique probablement par le fait que nous avons exclu d'avance les structures disloquées.

Type de syntagme nominal (SN)	Exemple(s)	Nb. occ.	% Redoub.
Pronom indéfini	<i>une, la plupart</i>	20	45%
Nom sans déterminant	<i>papa, Constantin, Neuchâtel</i>	10	60%
Déterminant + nom	<i>ma sœur, la différence</i>	303	48%
SN + modificateur pré- ou postnominal	<i>les trois garçons, cette femme-là</i>	17	18%
SN + adjectif/participe pré- ou postnominal	<i>des vaches laitières, les vieilles maisons</i>	13	38%
SN + syntagme prépositionnel	<i>les noms de lieux, le prix du terrain</i>	34	53%
SN + SN	<i>mon tonton et ma tata</i>	5	60%
Combinaisons de plusieurs types	<i>les plus forts à l'école, mes grands-parents du côté de mon père</i>	5	20%
Total		407	47%

Tableau 2: taux de redoublement en fonction de la complexité syntaxique du sujet lexical

Contrairement à nos attentes, le taux de redoublement n'augmente pas avec la complexité du sujet. Pour les sujets très complexes comme celui illustré en (18), on constate même une baisse du taux de redoublement, qui ne s'élève qu'à 20%. Les analyses statistiques confirment par ailleurs qu'il n'y a pas de relation entre la complexité syntaxique et le taux de redoublement ( $\chi^2(1) = 6.45$ ,  $p = .26$ ).

(18) et pis *mes grands-parents du côté de mon père* étaient très très très catholiques aussi (OFROM, unine15b27d\_937\_unine15-083)

[27] Par la suite, nous avons déterminé le nombre de syllabes du sujet syntaxique, afin d'établir une possible relation entre le poids prosodique du sujet et la présence ou l'absence d'un redoublement du sujet. Comme le montre la figure 3, le taux de redoublement se situe autour de 50% pour les sujets comprenant jusqu'à trois syllabes (ce qui correspond *grosso modo* au taux moyen de redoublement du corpus en entier). Pour les sujets très longs, qui comprennent entre 6 et 9 syllabes,

le taux de redoublement chute jusqu'à 29% des occurrences, alors qu'on se serait attendu à une augmentation plus ou moins nette pour ce type de sujet. Les résultats suggèrent que, contrairement à notre hypothèse de départ, le taux de redoublement n'augmente pas avec le poids syllabique du sujet. De manière plus générale, on constate que le poids syllabique n'a pas une incidence significative sur le taux de redoublement, ce qui est confirmé par l'analyse statistique ( $\chi^2(1) = 0.26$ ,  $p = .61$ ).

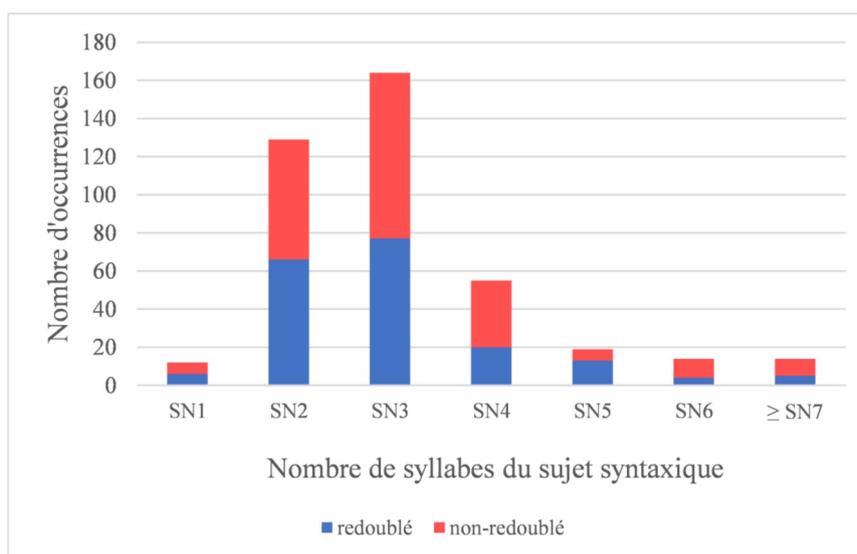


Figure 3: redoublement et poids syllabique du sujet

#### 4.2.2 Relation entre longueur du sujet et longueur du groupe verbal

[28] Afin d'étudier plus en détail la relation entre la longueur du sujet et la longueur du groupe verbal ainsi que l'incidence de ce facteur eurhythmique sur le taux de redoublement, nous avons d'abord déterminé, comme nous l'avons déjà mentionné à la section 3.3, le nombre de syllabes des groupes accentuels comprenant le sujet d'un côté et le verbe de l'autre. Seules les syllabes effectivement réalisées ont été considérées. Ensuite, nous avons comparé la longueur des deux groupes accentuels en déduisant le nombre de syllabes du groupe accentuel comprenant le verbe de celui du groupe accentuel comprenant le sujet. Pour ne pas gommer d'avance de possibles différences entre les occurrences avec et sans redoublement et afin d'éviter une circularité, nous n'avons pas inclus les clitiques sujet dans la détermination du nombre de syllabes dans le groupe verbal. Nous avons ainsi obtenu, pour un exemple sans redoublement comme celui en (19), où un sujet bisyllabique est suivi d'un groupe verbal trisyllabique, une différence de -1. Comme le clitique sujet n'est pas compté parmi les syllabes du groupe verbal,

nous obtenons également une différence de -1 pour l'exemple avec redoublement illustré en (20).

- (19) [les gens]<sub>GA\_sujet</sub> [se connaissent]<sub>GA\_groupe\_verbal</sub> et puis il y a une partie conviviale aussi après les concerts (*OFROM*, unine11b02m\_201\_unine11-gpa)  
 (20) [les gens]<sub>GA\_sujet</sub> [ils écoutent pas]<sub>GA\_groupe\_verbal</sub> (*OFROM*, unine15b64d\_1133\_unine15-105)

[29] Comme nous avons émis l'hypothèse selon laquelle le clitique sujet sert à équilibrer le nombre de syllabes du sujet et du groupe verbal, nous nous attendons donc à un taux de redoublement plus élevé pour les exemples où le nombre de syllabes dans le groupe accentuel comprenant le sujet excède le nombre de syllabes du groupe accentuel comprenant le verbe (i.e. différence positive). La syllabe supplémentaire que fournit le clitique sujet contribuerait donc à augmenter le poids rythmique du groupe verbal et établirait ainsi une eurythmie. Inversement, nous nous attendons à un taux de redoublement plus faible lorsque le nombre de syllabes dans le groupe accentuel comprenant le sujet est égal ou même inférieur au nombre de syllabes du groupe accentuel comprenant le verbe (i.e. différence négative).

[30] La figure 4 montre la distribution des différences de syllabes entre le groupe accentuel comprenant le sujet et celui comprenant le groupe verbal dans le corpus étudié.

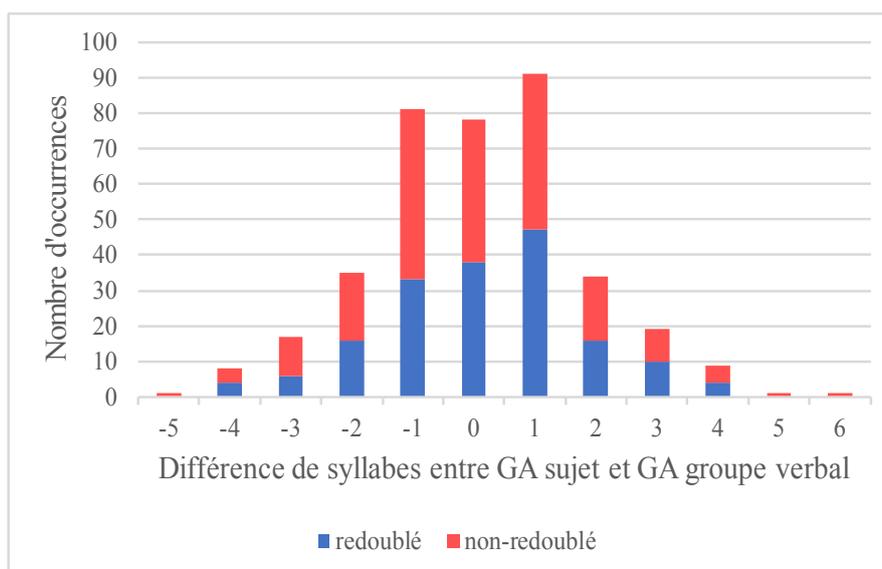


Figure 4: différence de poids syllabique entre GA sujet et GA groupe verbal

Une différence positive indique que le GA comprenant le sujet est plus long que le GA comprenant le groupe verbal, tandis qu'une différence négative indique que le sujet est plus court. Une différence de 0 indique que les deux groupes accentuels ont la même longueur. Étant donné qu'il n'y a qu'un seul exemple respectif pour les différences de -5, 5 et 6 syllabes (voir l'exemple 21), nous avons décidé de les exclure pour la suite des analyses.

(21) peut-être même soixante-trois [quand mon dernier des p'tits frères]<sub>GA\_sujet</sub> [est né]<sub>GA\_groupe\_verbal</sub> (*OFROM*, unine15b57m\_574\_unine15-105)

[31] Nous avons ensuite étudié si le taux de redoublement variait en fonction de la différence de poids syllabique entre le groupe accentuel comprenant le sujet et celui comprenant le groupe verbal. Les résultats illustrés dans la figure 5 montrent que, si l'on ne considère que les occurrences allant de -3 (i.e. les GA sujet contiennent 3 syllabes de moins que les GA groupe verbal) à 3 (i.e. les GA sujet contiennent 3 syllabes de plus que les GA groupe verbal), on remarque que le taux de redoublement augmente au fur et à mesure que la différence augmente positivement ( $\chi^2(1) = 4.28$ ,  $p = .04$ ).

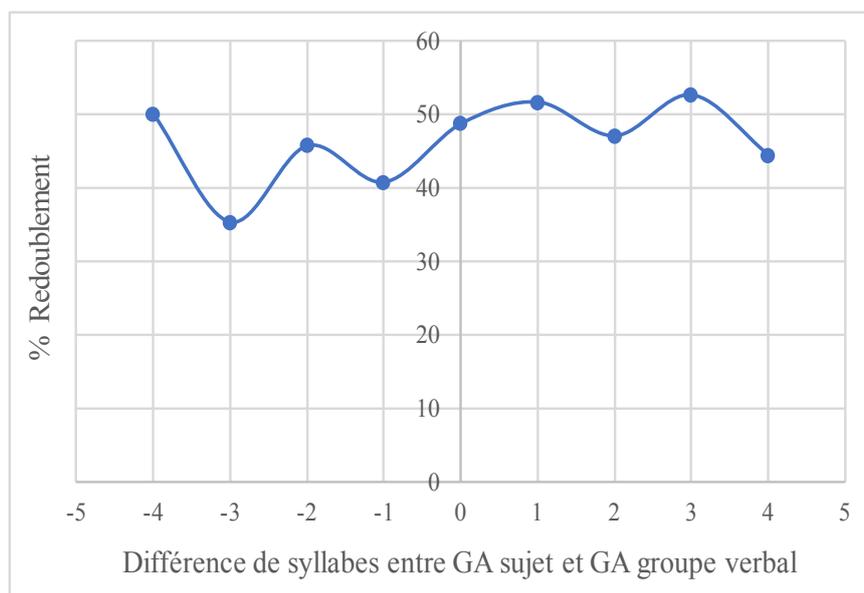


Figure 5: redoublement et différence de poids syllabique entre GA sujet et GA groupe verbal

Le taux de redoublement atteint même jusqu'à 53% pour les exemples comme celui illustré en (22), où le groupe accentuel comprenant le sujet compte trois syllabes de plus que le groupe accentuel comprenant le verbe.

- (22) pis euh enfin justement [une de mes cousines]<sub>GA\_sujet</sub> [elle disait]<sub>GA\_groupe\_verbal</sub>  
(*OFROM*, unine15b27d\_116\_unine15-083)

[32] Il faut cependant souligner que la tendance au redoublement observée vaut uniquement pour les exemples avec une différence maximale de trois syllabes entre le sujet et le verbe. Lorsque la différence de longueur entre sujet et verbe équivaut à quatre syllabes, on constate même une inversion de cette tendance : pour des exemples comme en (23), où la longueur du groupe verbal excède la longueur du sujet de quatre syllabes, le taux de redoublement s'élève à 50%. Ce taux relativement élevé est d'autant plus étonnant que le clitique contribue à faire augmenter la différence en termes de poids syllabique entre sujet et verbe au lieu de l'équilibrer.

- (23) donc elle doit refaire en plus enfin un truc vraiment bizarre [sa mère]<sub>GA\_sujet</sub>  
[elle m'a un peu expliqué]<sub>GA\_groupe\_verbal</sub> (*OFROM*, unine15z62d\_727\_unine15-936)

[33] Même si ces cas sont certes peu nombreux (une quinzaine d'exemples), ils ne constituent pourtant pas une quantité négligeable. L'exemple en (24) montre que l'inversion de la tendance vaut également pour les cas dans lesquels le groupe accentuel sujet excède de quatre syllabes le groupe accentuel comprenant le verbe. Ici, le clitique aurait pu contribuer à diminuer la différence de longueur, mais le locuteur a choisi de ne pas redoubler le sujet.

- (24) [le niveau professionnel]<sub>GA\_sujet</sub> [m'attire pas]<sub>GA\_groupe\_verbal</sub> (*OFROM*, unine15b40m\_510\_unine15-092)

Comment peut-on expliquer cette divergence dans la réalisation du redoublement, qui fait que ce sont justement les groupes accentuels les plus hétérogènes qui n'ont pas recours au clitique en tant que mécanisme équilibreur ? On pourrait éventuellement suggérer que, lorsque la différence de longueur entre sujet et verbe est trop importante, le clitique ne 'fait plus le poids' dans la balance et son emploi ne suit donc plus la tendance attendue. Ainsi, dans l'exemple (24), le clitique n'aurait pas réussi à estomper le poids syllabique très important du sujet, qui atteint avec sept syllabes la limite observée pour la taille du groupe accentuel en français (Martin 2004 : 55). Le potentiel eurythmique du clitique s'avère ainsi neutralisé dans des contextes qui présentent une irrégularité rythmique inhérente.

## 5 Conclusion et perspectives

[34] Notre étude de corpus a montré que le redoublement du sujet est très fréquent en français valaisan : dans le corpus *OFROM*, le taux de redoublement

est de 47%, c'est-à-dire que les locuteurs du corpus emploient la variante redoublée pratiquement dans la moitié des cas dans lesquels ils ont le choix entre les deux variantes.

[35] Malgré l'importante variabilité entre les locuteurs (voir supra), il conviendrait donc d'explorer les raisons pouvant expliquer la fréquence élevée du redoublement dans la variété valaisanne : est-elle liée à d'autres particularités rythmiques ou prosodiques de la variété, telles que l'accent initial ou le débit, par exemple ? Ou doit-on plutôt l'attribuer à la nature du corpus et aux caractéristiques des interactions qui y sont représentées ? On pourrait notamment partir de l'hypothèse que les interactions ont lieu dans un registre plus informel que dans d'autres corpus.

[36] En ce qui concerne les caractéristiques rythmiques étudiées, les hypothèses de départ n'ont pu être confirmées qu'en partie. D'une part, nos résultats ont montré que la longueur du sujet n'a pas d'effet sur le taux de redoublement. En revanche, nous avons pu établir, d'autre part, une tendance dans l'incidence de la différence de longueur entre sujet et groupe verbal sur le taux de redoublement : les données du corpus révèlent une augmentation du taux de redoublement au fur et à mesure que la différence de longueur entre sujet et verbe augmente positivement.

[37] On peut donc souligner que des facteurs eurythmiques ont une incidence sur la présence ou l'absence du clitique sujet, dont l'emploi contribue à l'eurythmie des groupes accentuels. Cependant, nous avons constaté que cette tendance était valable uniquement pour une différence allant jusqu'à trois syllabes, ce qui nous a amené à suggérer que le clitique ne pouvait pas déployer son potentiel eurythmique au-delà d'un certain seuil. Il serait envisageable de vérifier cette hypothèse au moyen de jugements d'acceptabilité. Pour ce faire, on construirait des phrases dont la différence de poids syllabique entre sujet et groupe verbal varierait entre -4 et 4 syllabes. Les phrases seraient ensuite présentées à des auditeurs dans deux versions – avec et sans redoublement du sujet – et les auditeurs donneraient un jugement d'acceptabilité pour chacune des versions.

[38] Étant donnée la grande variabilité entre les locuteurs, il serait intéressant de déterminer si le facteur eurythmique a la même importance pour tous les locuteurs. Une première exploration des données de notre corpus a révélé que chez les 'faibles redoubleurs' affichant moins de 30% de redoublement (Auger & Villeneuve 2010 : 75), le taux de redoublement tend à augmenter dans des contextes où l'ajout du clitique peut pleinement exercer son potentiel eurythmique. En revanche, le rôle du facteur eurythmique est moins marqué pour les locuteurs avec un taux de redoublement plus élevé (entre 31 % et 99 %). Il se peut par conséquent que l'étude d'un corpus plus large révèle l'existence de différentes 'grammaires' chez les locuteurs.

[39] S'agissant d'une première exploration des possibles motivations prosodiques du redoublement du sujet, la présente étude s'est focalisée sur des facteurs rythmiques. Des analyses futures pourront davantage s'appuyer sur des paires

(quasi-)minimales (voir les exemples 25 et 26 provenant de la même locutrice) et prendre en compte des facteurs mélodiques.

(25) *ma belle-sœur s'était mariée* quinze jours avant (OFROM, unine11c05m\_019\_unine11-sda)

(26) *ma belle-sœur elle s'est mariée* à quatre heures de l'après-midi (OFROM, unine11c05m\_025\_unine11-sda)

Au moyen d'analyses acoustiques, on pourra ainsi étudier l'incidence de facteurs mélodiques comme l'accent secondaire ou le rôle du degré de la frontière prosodique précédant le clitique. Il conviendrait également de tester si l'effet du différentiel syllabique dans la limite des trois syllabes ainsi que son absence au-delà de ce seuil se confirment lorsqu'on augmente la quantité de données prises en compte.

## Abréviations

GA = groupe accentuel

GV = groupe verbal

SN = syntagme nominal

## Bibliographie

- Ashby, William J. 1980. Prefixed conjugation in Parisian French. Herbert J. Izzo (éd.). *Italic and Romance. Linguistic studies in honor of Ernst Pulgram*. Amsterdam : Benjamins, 195-207.
- Auger, Julie, Anne-José Villeneuve 2010. La double expression des sujets en français saguenéen : étude variationniste. Wim Remysen, Diane Vincent (éds.). *Hétérogénéité et homogénéité dans les pratiques langagières. Mélanges offerts à Denise Deshaies*. Québec : Presses de l'Université Laval, 67-85.
- Avanzi, Mathieu 2012. *L'interface prosodie/syntaxe en français. Dislocations, incises et asyndètes*. Bruxelles : Lang.
- Avanzi, Mathieu et al. 2011. Vers une modélisation continue de la structure prosodique : le cas des proéminences syllabiques. *Journal of French Language Studies* 21, 53-71.
- Avanzi, Mathieu et al. 2012. La prosodie de quelques variétés de français parlées en Suisse romande. Anne-Catherine Simon (éd.). *La variation prosodique régionale en français*. Bruxelles : De Boeck/Duculot, 89-118.
- Avanzi, Mathieu, Marie-José Béguelin, Federica Diémoz 2016. *OFROM. Corpus oral de français de Suisse Romande*. [http://www11.unine.ch/uploads/Documents/AM-MJB-FD\\_OFROM.pdf](http://www11.unine.ch/uploads/Documents/AM-MJB-FD_OFROM.pdf). Consulté le 20 mai 2019.
- Beaulieu, Louise, Patricia Balcom 1998. Le statut des pronoms personnels sujets en français acadien du nord-est du Nouveau-Brunswick. *Linguistica atlantica* 20, 1-27.
- Berrendonner, Alain 2007. Dislocation et conjugaison en français. *Cahiers de praxématique* 48, 85-110. <https://journals.openedition.org/praxematique/779>.
- Buthke, Carolin, Raféu Sichel-Bazin, Trudel Meisenburg 2014. Dislokation im gesprochenen Französisch: zwischen Emphase und Grammatik. In: Ingrid Isenhardt, Ralf Steinert, Gert G. Simon (éds.). *Emotionen, Expressivität, Emphatische Strukturen*. Berlin : Springer, 11-21.
- Coveney, Aidan 2003. Le redoublement du sujet en français parlé. Anita B. Hansen, Maj-Britt M. Hansen (éds.). *Actes du colloque international, Université de Copenhague du 22 au 23 juin 2001*. Copenhague : Museum Tusulanum Press, 111-143.
- Coveney, Aidan 2005. Subject doubling in spoken French: a sociolinguistic approach. *The French Review* 79, 96-111.
- Delais-Roussarie, Elisabeth 1995. *Pour une approche probabiliste de la structure prosodique. Étude de l'organisation prosodique et rythmique de la phrase française*. Thèse de doctorat, Université Toulouse-le-Mirail.
- Dell, François 1984. L'accentuation dans les phrases en français. François Dell, Daniel Hirst, Jean-Roger Vergnaud (éds.). *Forme sonore du langage : structure des représentations en phonologie*. Paris : Hermann, 65-122.
- Dufter, Andreas, Christoph Gabriel 2016. Information structure, prosody, and word order. Susann Fischer, Christoph Gabriel (éds.). *Manual of grammatical interfaces in Romance*. Berlin : De Gruyter, 419-455.
- Fonseca-Greber, Bonnie, Linda R. Waugh 2003. The subject clitics of conversational European French. Morphologization, grammatical change, semantic change, and change in progress. Rafael Núñez, Luis López, Richard Cameron (éds.). *A Romance perspective on language*

- knowledge and use*. Amsterdam : Benjamins, 99-117.
- Gabriel, Christoph, Esther Rinke 2010. Structure informationnelle et statut morphosyntaxique des clitiques : évidence diachronique du doublement pronominal espagnol et français. Andreas Dufter, Daniel Jacob (éds.). *Syntaxe, structure informationnelle et organisation du discours dans les langues romanes*. Frankfurt am Main : Lang, 95-116.
- Girard, Francine 2010. Le statut des clitiques sujets cadiens. Franck Neveu et al. (éds.). *CMLF 2010 - 2ème Congrès mondial de Linguistique française*, 2103-2112. [https://www.linguistiquefrancaise.org/articles/cmlf/abs/2010/01/cmlf2010\\_000209/cmlf2010\\_000209.html](https://www.linguistiquefrancaise.org/articles/cmlf/abs/2010/01/cmlf2010_000209/cmlf2010_000209.html).
- Goldman, Jean-Philippe 2011. *EasyAlign*: an automatic phonetic alignment tool under Praat. *Interspeech 2011. 12th Annual Conference of the International Speech Communication Association. Florence, Italy, 27-31 August 2011*. [https://www.isca-speech.org/archive/interspeech\\_2011/i11\\_3233.html](https://www.isca-speech.org/archive/interspeech_2011/i11_3233.html).
- Koch, Peter 1993. Le « chinook » roman face à l'empirie. Y a-t-il une conjugaison objective en français, en italien et en espagnol et une conjugaison subjective prédéterminante en français ? Gerold Hilty (éd.). *Actes du XX<sup>e</sup> Congrès International de Linguistique et Philologie Romanes. Université de Zurich (6-11 avril 1992). Tome III*. Tübingen : Niemeyer, 171-190.
- Lacheret-Dujour, Anne, Frédéric Beaugendre 1999. *La prosodie du français*. Paris : CNRS éditions.
- Martin, Philippe 1987. Prosodic and rhythmic structures in French. *Linguistics* 25, 925-949.
- Martin, Philippe 2004. Intonation de la phrase dans les langues romanes : l'exception du français. *Langue française* 141, 36-55.
- Massot, Benjamin 2008. *Français et diglossie. Décrire la situation linguistique française contemporaine comme une diglossie : arguments morphosyntaxiques*. Thèse de doctorat, Université Paris VIII Vincennes-Saint Denis.
- Massot, Benjamin 2010. Le patron diglossique de variation grammaticale en français. *Langue française* 168, 87-106.
- Nagy, Naomi, Hélène Blondeau, Julie Auger 2003. Second language acquisition and 'real' French: An investigation of subject doubling in the French of Montreal Anglophones. *Language Variation and Change* 15, 73-103.
- OFROM = Mathieu Avanzi, Marie-José Béguelin, Federica Diémoz (éds.) 2012-2019. *Corpus oral de français de Suisse romande*. <http://www.unine.ch/ofrom>.
- Queffélec, Ambroise 2000. Le français au Maghreb. Gérald Antoine, Bernard Cerquiglini (éds.). *Histoire de la langue française 1945-2000*. Paris : CNRS éditions, 765-796.
- Roberge, Yves 1990. *The syntactic recoverability of null arguments*. Montréal : McGill-Queen's University Press.
- Sankoff, Gillian 1982. Usage linguistique et grammaticalisation : les clitiques sujets en français. Norbert Dittmar, Brigitte Schlieben-Lange (éds.). *Die Soziolinguistik in romanischsprachigen Ländern*. Tübingen : Narr, 81-85.
- Stark, Elisabeth 1997. *Vorstellungsstrukturen und topic-Markierung im Französischen. Mit einem Ausblick auf das Italienische*. Tübingen : Narr.
- Zahler, Sara 2014. Variable subject doubling in spoken Parisian French. *University of Pennsylvania Working Papers in Linguistics* 20, 361-371. <https://repository.upenn.edu/pwpl/vol20/iss1/38/>.
- Zribi-Hertz, Anne 2011. Pour un modèle diglossique de description du français : quelques implications théoriques, didactiques et méthodologiques. *Journal of French Language Studies* 21, 231-256.