

Toute i'est à vitesse : effet de la fréquence verbale sur la grammaticalisation des pronoms sujets en français parlé

Alphonse Kabano

Université Nationale du Rwanda et Université d'Ottawa (chercheur invité)

alphonsekabano@gmx.com

Résumé : L'émergence d'une grammaire du français où le pronom sujet se grammaticaliserait en une marque verbale d'accord est testée en contexte de redoublement du SN sujet dans un corpus de données vernaculaires d'une variété canadienne de français parlé. Une analyse de l'impact de la fréquence verbale suggère qu'une petite portion de formes verbales fléchies à morphologie irrégulière sont impliquées dans le patron pronom sujet + verbe. Ce n'est ni des formes verbales lemmatiques ni la catégorie verbale comme un tout qui sont concernées, d'où la généralisation que les formes verbales sont sélectionnées déjà dotées du pronom sujet est mise en cause.

Mots clés : pronom sujet, verbe, usage, fréquence, grammaticalisation

Abstract : The emergence of a grammar of French where the subject pronoun is getting grammaticalized as an agreement verbal prefix is tested in the context of subject doubling within a corpus of vernacular data of a variety of Canadian spoken French. An analysis of the impact of the inflected verb frequency suggests that a small portion of verb forms with irregular morphology are involved in the instantiation of the pattern. Neither all the verb lemmas nor the overall verb category are concerned, and this rules out the idea that verbal forms are selected already inflected with the subject pronoun.

Keywords : subject pronoun, verb, usage, frequency, grammaticalization

1 Introduction

Le redoublement du sujet en français parlé (*mes enfants ils vont à l'école*) engendre une structure linguistique mystérieuse. Presque aussi vieille que la langue française selon des études historiques (Blasco-Dulbecco 1999 pour ne citer que celle-là), longtemps stigmatisée par la grammaire comme étant le signe d'un « langage populaire », « vulgaire » (Sandfeld 1928), d'un « style pauvre » (Judge et Healey 1983), etc., cette structure reste cependant aujourd'hui « une forme tout à fait normale dans la langue parlée » (Manzini et Savoia 2002) et semble relever d'un processus « naturel » (Givón 1975). Il est d'ailleurs prédit qu'elle se généralisera pour faire passer le français (parlé) du statut de langue à sujet non nul au statut de langue à sujet nul (Friedemann 1995). La grammaticalisation (Heine, Claudi et Hunnemeyer 1991; Hopper et Traugott 1993; Bybee, Perkins et Pagliuca, 1994, etc.), mécanisme de portée universelle, interviendrait pour transformer le pronom

personnel sujet sous forme de marque d'accord intégrée dans la morphologie verbale (Givón 1976, 1984, 1995; Bubenik 1993; Lyons 1990; Lehmann 1987, 1995, etc.), autorisant ainsi à redoubler le sujet. Cela dit, la controverse subsiste quant au statut réel du pronom personnel redoubleur. Une analyse alternative voudrait ainsi qu'il ne soit en rien une marque d'accord mais un constituant à part entière s'analysant dans la composante syntaxique (de Cat 2005, entre autres). Dans cette dernière perspective, la réanalyse (de la dislocation) est justement postulée par certains auteurs (par exemple Auger 1994), comme responsable de la génération du redoublement du sujet

Dans cet article, nous approfondissons une recherche déjà menée sur le français parlé dans les années 1980 à Ottawa-Hull (région de la capitale fédérale du Canada) (Kabano 2003). Il ressort de cette étude que la variation caractéristique du redoublement du sujet est stable dans cette région, et que, en temps apparent, on ne trouve pas de signes d'un changement imminent. Voici des exemples :

- (1) a. *en tous cas, les enfants _ parlent toutes français* (010/0258) ¹
b. *les enfants _ ont toujours été à école française* (023/0016)
c. *mes enfants là ils sont dans, tsé, bien des choses* (042/0653)
d. *c'est que les enfants ben ils sont tout seuls le soir* (042/0631)
e. *ma mère _ nous a élevés pas mal toute seule* (052/0922)
f. *non, mes parents ils m'ont élevé du bon côté* (100/1047)
g. *ses enfants ils ont tous été élevés en français* (010/0378)
h. *les enfants _ étaient élevés en anglais* (035/1335)

Il s'agit de voir si la grammaticalisation supposée du pronom personnel sujet en français fait partie d'une grammaire réellement émergente, en construction, grâce à l'opérativité de la fréquence et de la routinisation (Hopper 1998). L'approche est centrée sur l'usage et vise à déterminer le type de schéma possédé par les usagers de la langue, qui permettrait de générer le pronom personnel en avant du

¹ Les codes réfèrent au numéro du locuteur et à la ligne dans la transcription du *Corpus du français parlé à Ottawa-Hull* (Poplack 1989). Le corpus est un projet de la professeure S. Poplack bénéficiant d'une subvention du CRSHC; il est gardé au Laboratoire de sociolinguistique, Département de linguistique, à l'Université d'Ottawa. Les exemples sont des reproductions verbatim du discours des locuteurs. Le tiret (_) indique l'absence du pronom sujet.

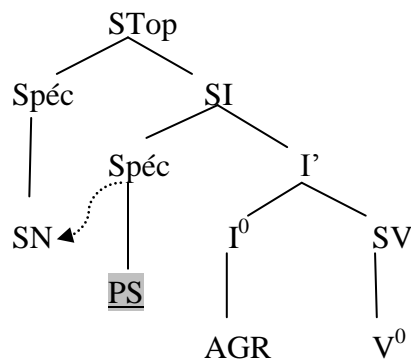
verbe et de redoubler le sujet. L'objectif est d'évaluer la productivité du patron pronom+verbe lorsque la position sujet est occupée par le syntagme nominal.

Dans la section 2, nous formulons de manière détaillée la problématique en rapport avec le fonctionnement et le devenir du pronom sujet en redoublement. Dans la section 3, nous parlons du lien étroit entre changement linguistique, variation et mécanismes constructeurs des patrons linguistiques, dont la fréquence. Dans la section 4, nous exposons notre méthode de collecte des données, et les aspects relatifs à l'analyse. Dans la section 5, nous présentons et discutons nos résultats par rapport aux effets de la fréquence des formes verbales.

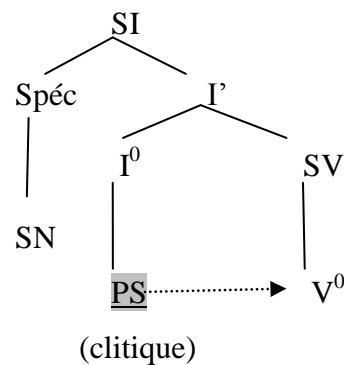
2 La problématique

Le statut et la position du pronom sujet redoubleur suscite de vifs désaccords : occupe-t-il la même position syntaxique que le SN qu'il est censé remplacer? Notons que la co-occurrence des deux éléments est interdite dans le même espace prédicatif (sauf dans l'inversion interrogative (*mes enfants vont-ils à l'école?*)). Ou le pronom fait-il partie intégrante de la morphologie verbale, ce qui autorise alors le redoublement du SN? En (2a), la position du spécificateur du syntagme inflexionnel (SI), qui revient au SN sujet, est occupé par le pronom sujet (PS) anaphorique qui le remplace suite à son placement en position topique (2a). Quand le SN et le PS cohabitent en dehors de la mise en relief (2b), c'est que le PS est en position I réservée, entre autres, aux marques d'accord (AGR), et que le SN a 'retrouvé' sa position de base. On pose ainsi que (2a) est réanalysée en (2b).

(2) a.



b.



Dans une approche minimaliste, Auger (1998 : 55), applique au français parlé à Montréal « un traitement vraiment morphologique de la morphologie flexionnelle », et situe à l'interface lexique/syntaxe la génération du pronom sujet au sein du verbe, ou plus largement du SV. Le rôle de cette interface serait de « transformer les entrées lexicales en formes fléchies munies des traits morphosyntaxiques nécessaires à leur insertion dans une structure syntaxique ». La proposition de Auger rappelle étonnamment des intuitions anciennes, par exemple, l'idée d'une table de conjugaison comprenant des formes du genre *je-parle, tu-parles, il-parle* (Brunot et Bruneau 1949 : 276). Sous la version forte de cette proposition, le pronom sujet n'est pas inséré dans la structure de la phrase tardivement, juste pour remplacer le SN, mais est déjà intégré à la forme verbale dès sa sélection du lexique.

La possibilité que la structure subjectale redoublée se généralise, comme en picard (Auger 1994, 2003), en français d'Afrique du Nord (Lanly 1970), ou dans les dialectes du Nord de l'Italie cités comme le prototype est problématique. Il est pourtant plus judicieux d'envisager les choses de manière variable. Friedemann (1997 : 148–149) prédit ainsi un changement d'envergure : « Une fois la position du sujet lexical dans les phrases à redoublement clitique sujet réanalysée comme position A, le pas à franchir pour engendrer le clitique sujet directement en I⁰ (et donc de l'analyser comme une marque d'accord) n'est pas considérable. Le redoublement jusqu'alors massif devient obligatoire et *pro* est légitimé (de façon générale) ». Blasco-Dulbecco (1999 : 52–53) recommande cependant la prudence : « ...vu la présence des dislocations dans l'ancien français, quel intérêt trouverions-nous à affirmer son imposition dans la langue et sa grammaticalisation au détriment d'autres structures? Pourquoi ne nous contenterions-nous pas d'y déceler une simple variation? »

Notre étude sur le français parlé dans la région Ottawa-Hull (Kabano 2003), est arrivée, comme nous l'avons annoncé, à la conclusion que le redoublement du sujet fait l'objet d'une variation stable, qui ne montre pas de signes de changement, du moins à court et à moyen terme. La structure est en effet réalisée dans 31% des énoncés; ce pourcentage n'est pas très parlant, car nous ne disposons pas

de données pouvant aider à déterminer en temps réel depuis quand le phénomène existe et quelle est sa trajectoire exacte au cours des siècles. Pour ce qui est de l'impact des paramètres sociométriques (âge, sexe, niveau d'étude, occupation, prestige du quartier de résidence, indice de maîtrise de la langue anglaise), ils ne sont réellement pas significatifs, ce qui laisse penser que la variation est inhérente à la structure linguistique, et qu'aucun groupe ou sous-groupe social ne se démarque pour initier et orchestrer un changement.

Sur le plan linguistique, il apparaît, premièrement, que la proximité du SN sujet par rapport au verbe est le facteur prépondérant qui favorise le redoublement du sujet. Or cette contrainte est rapportée et tolérée depuis au moins 4 siècles dans nombre de grammaires du français (Nyrop 1903; Blinkenberg 1928; etc.). Nous avons donc un état de langue qui perdure depuis longtemps. Deuxièmement, le rôle de marquage d'accord que le pronom sujet viendrait assumer en compensation de l'affaiblissement des désinences verbales (Moignet 1965), ne s'avère pas pertinent. Heap (2000) a observé lui-même que le pronom sujet est plutôt déclenché en proportion inverse du syncrétisme verbal en roman central. Troisièmement, en testant l'impact de la fréquence des formes verbales hôtes sur le pronom sujet, nos résultats ont indiqué que la distribution du pronom est restreinte à un petit nombre de verbes, ce qui ne semble pas être en faveur de la productivité du patron pronom+verbe à titre d'un schéma général enraciné ou en voie de l'être dans le système. Or, l'idée de regarder l'impact des verbes individuels rejoint l'intuition de Brunot (1922), qui écrit que dans l'ancienne langue, le pronom, sans que rien de spécial l'appelle, était réalisé « quand le verbe était *avoir, faire, être* ». Des verbes précis se seraient donc montrés favorables au pronom sujet. Encore une fois, nous avons un état de langue qui aurait prévalu depuis belle lurette.

3 Considérations théoriques

3.1 *Le changement linguistique*

Le changement est inhérent à la structure des langues, et toute théorie de la grammaire doit prédire les possibilités et les limites du changement. Ce qui permet le changement c'est l'hétérogénéité et la variabilité dans l'application des règles

linguistiques. Selon Weinreich, Labov et Herzog (1976 : 100), « Une maîtrise quasi innée des structures hétérogènes ... fait partie de la compétence linguistique de l'individu unilingue [...] c'est l'absence d'une hétérogénéité structurée qui se révélerait dysfonctionnelle ». Depuis Labov, l'on sait que la variation n'est ni aléatoire ni anarchique, mais est gouvernée par des règles (Winford 1997).

Lorsque deux formes (le redoublement et le non-redoublement du sujet, par exemple) sont en compétition dans l'usage, une peut en venir à dominer l'autre, en général dans un processus transitionnel et graduel selon le modèle inspiré par Haugen ($A > Ab > aB > B$), comme elles peuvent rester toutes les deux en usage de manière stable sur une longue période, car la variation ne conduit pas nécessairement au changement. La manière dont une innovation linguistique se propage peut s'observer de deux façons : en temps réel, historiquement et/ou en temps apparent, synchroniquement (Labov 1994). Dans tous les cas, le cheminement est dicté par le fonctionnement cognitif de l'homme : tout commence par des préférences, qui deviennent des tendances, puis des habitudes, puis des contraintes plus fermes, et finalement des règles rigoureuses, voire catégoriques (Aitchison 1996 : 38). Le parcours est accompli inconsciemment, mais d'une manière non moins gouvernée par des règles, et le cycle de changement est susceptible de recommencer dès que de nouvelles préférences se déclarent.

Plusieurs aspects intéressent l'étude des changements linguistiques : le niveau concerné dans l'architecture de la langue, les causes internes et externes, les motivations, les mécanismes cognitifs impliqués, le mode et le rythme de propagation dans le système et au sein de la communauté, le potentiel de dissémination inter-dialectale, les jugements de valeur associés aux changements, etc. Un aspect crucial est de savoir si des changements multiples sont reliés à une même contrainte, ou si un changement qui a lieu dans une région de la grammaire (par exemple, la morphologie) enchevêtre un autre dans une autre région (par exemple, la syntaxe).

3.2 *La construction et l'enracinement*

L'abandon d'une forme A et l'adoption d'une forme B est un processus mesurable en termes de récurrences. L'usage d'une forme qui tombe en désuétude diminue au moment où la forme qui s'installe voit sa réalisation prendre plus d'ampleur.

Pour cerner et prédire la productivité d'un élément, d'un patron ou d'une règle linguistique, l'approche centrée sur l'usage prônée entre autres par Bybee, Langacker, Burlow et Kemmer, etc. offre un cadre opératoire intéressant. Deux dimensions sont importantes. D'abord, la connaissance des règles linguistiques se construit (d'où la 'Grammaire de construction') et s'organise en représentations abstraites et en réseaux schématiques qui, grâce à un travail continu de généralisation, émergent (d'où la 'Grammaire Émergente') des expériences et des usages particuliers que les locuteurs cumulent (Hopper 1998:164). Chaque fois qu'un locuteur ouvre la bouche pour parler, il contribue, sinon à la construction, du moins au renforcement des structures de la langue. L'opération s'effectue de bas en haut ('Bottom-up', Langacker 2000), et non de haut en bas ('Top-down') comme dans la grammaire générative qui insiste sur les règles générales, les principes universels et innés, au détriment des sous patrons de niveau inférieur (Langacker 2000).

Ensuite, l'enracinement ('Entrenchment', Langacker 1987) d'un élément ou d'un patron ou d'une règle dans la langue donne la mesure de sa disponibilité, de la facilité et de la rapidité de son accessibilité procédurale. Il s'agit de la valeur d'activation d'un élément au sein d'un modèle de type connexionniste. Cette valeur est fonction de la répétition et de l'automatisation. La fréquence avec laquelle un élément apparaît en usage influe sur la nature de sa représentation mentale, et fait partie des propriétés intrinsèques des unités linguistiques, à côté de leurs traits de nature qualitative, tels que la prononciation, le sens, le type d'environnement syntaxique d'un mot, etc. Ainsi, plus un mot sera fréquent, plus il aura tendance à être phonologiquement raccourci (Zipf), et plus il tendra sémantiquement à la polysémie ('Reducing Effect'). La fréquence a ainsi un effet immédiat sur la forme et la substance des éléments de la langue. Ainsi, la perte des effets pragmatiques caractéristique de la grammaticalisation, est, selon Haspelmath (1999), une conséquence directe de l'habituation à travers la fréquence d'usage. D'autre part, des recherches psycholinguistiques ont démontré que plus deux entités syntaxiques s'associeront, plus elles tendront à être stockées et retrouvées dans la mémoire sous une forme non décomposée. S'il s'avère alors que le pronom sujet per-

siste en avant du verbe dans des contextes nouveaux par rapport au français standard, l'hypothèse d'une tendance à la sélection des formes verbales déjà porteuses du pronom est en soi logique dans le cadre du processus d'enracinement.

3.3 *La configuration fréquentielle*

On distingue deux sortes de fréquence : la fréquence d'occurrence qui réfère à la fréquence d'une expression dans l'usage concret, et la fréquence type qui est en rapport avec le nombre d'items ou expressions linguistiques qui instancient un patron ou un schéma constructionnel donné (MacWinney 1978, Bybee 1985, entre autres). Les effets des deux types de fréquence dans la structuration, le stockage et l'activation des expressions et patrons linguistiques sont différents. Plus on enregistre des occurrences d'une expression lexicale, plus sa récurrence est facilitée. Dans la reconnaissance et l'accès lexical, les mots les plus fréquents sont les plus rapidement et les plus facilement retrouvés. Des collocations fixées par la routine sont activées sans analyse décompositionnelle. En revanche, un schéma constructionnel, qui est abstrait, sera d'autant enraciné dans la grammaire qu'elle s'appliquera à un plus large éventail de types d'expression. Une fréquence élevée d'une expression peut cependant rendre moins dense son rattachement à un patron, surtout quand ce dernier est une innovation, et donc résister au processus d'analogie et de changement diachronique. Ainsi, des verbes irréguliers de haute fréquence peuvent échapper à la régularisation parce qu'ils sont automatisés. Kroch (1989) montre, par exemple, que lors de l'introduction de l'élément support *do* en anglais dans les transformations interrogatives et négatives au 14^e siècle, le système verbal était affecté, excepté les verbes *have, be, can, may, need, ought*, et quelques autres qui continuent de se construire sans l'élément support. Or, au moment du développement du nouveau patron avec *do*, ces verbes non affectés étaient de grande fréquence. C'est qu'une expression enracinée dans la langue pourra fonctionner de manière autonome et soustraite à la contrainte d'une innovation (Bybee et Thompson (1997). Des membres d'une classe, en dépit de leur fréquence, mais en raison de leur autonomie, peuvent donc ne pas contribuer à la propagation d'un changement et tendre à la conservation ('Conserving Effect').

Deux cas de figure se présentent : nous pouvons avoir un schéma faiblement enraciné dans le système de la langue avec un petit nombre d'instances, qui elles, sont fortement enracinées, ou, par contre, un schéma fortement enraciné dans le système avec de nombreuses instances qui, elles, sont faiblement enracinées. Pour qu'un changement d'un schéma A vers un schéma B soit effectif, il faut que la fréquence type de B soit accrue, autrement dit que B soit productif et facile à activer comme choix par défaut en cas de schémas en compétition, et s'applique donc à de nouveaux items ou à de nouveaux contextes (Langacker 2000). L'hypercorrection est le revers de la médaille d'une productivité qui déborde le champ d'application 'normal' d'un schéma. On fixera à l'esprit qu'une fréquence élevée d'un ou de quelques éléments isolés enracinés n'implique pas nécessairement que le schéma constructionnel général soit productif puisqu'il peut s'agir d'une sous-composante autonome par rapport aux autres éléments de la langue.

4 Méthodologie

4.1 Le corpus

Nos données proviennent du corpus du *Français parlé à Ottawa-Hull* (Poplack 1989) gardé au Laboratoire de Sociolinguistique à l'Université d'Ottawa. D'une taille de 3.5 millions de mots, le corpus est constitué d'interviews sociolinguistiques menées dans les années 1980 dans 5 quartiers à Ottawa majoritairement anglophone et à Hull (aujourd'hui Gatineau) majoritairement francophone, auprès de 120 locuteurs. Les bandes sonores, d'une bonne qualité, sont accompagnées de transcriptions. Le corpus est aussi organisé sous forme de concordances faciles à consulter. En plus des textes, véritable mine des discours vernaculaires de la région, des indications sociométriques (âge, sexe, niveau d'études, occupation, performance en anglais) sont fournies. Sur les 120 locuteurs, nous avons sélectionné 38 répartis d'une manière qui reflète la diversité humaine inhérente à la communauté. Nous n'en donnons pas de détails ici puisque nous ne nous intéressons pas à l'effet des facteurs socio-économique sur le phénomène que nous étudions.

Le corpus a donné lieu à d'excellents travaux de recherche variationniste : Poplack (1992) sur la sélection du subjonctif en proposition subordonnée suite à

certaines verbes matriciels, LeBlanc (1999) sur le conditionnel dans les propositions hypothétiques, Willis (2000) sur l'alternance des auxiliaires *être* et *avoir*, Klapka (2002) sur le processus d'accord, Kabano (2003) sur le redoublement du sujet. Un fait épatant qui ressort de ces travaux, même s'ils sont basés sur une approche en temps apparent, est que le français parlé à Ottawa-Hull, exhibe des zones de stabilité depuis plus de 3 siècles. Ceci n'exclut pas qu'il y ait eu des changements dans le fonctionnement particulier du pronom personnel sujet.

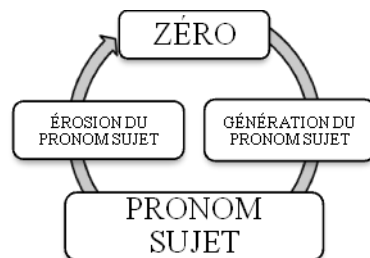
4.2 Les énoncés

Dans Kabano (2003 : 81-94), nous exposons les critères de sélection des énoncés dans lesquels le redoublement est variable, c'est-à-dire se présente comme une option entre la présence et l'absence du pronom sujet. Globalement, nous nous limitons à la structure SN-(PP)-SV minimale que nous avons exemplifiée en (1). Nous excluons tout contexte qui ne permet pas la variation, par exemple, des locutions figées, enregistrées telles quelles dans la mémoire sans aucune marge de variation. Sont écartés aussi les énoncés qui ne respectent pas l'ordre SN (PP) SV, comme les phrases à dislocation à droite (*ils étaient bien stricts les frères* (052/0212)) dans lesquelles la contrainte sur la présence vs. l'absence du pronom est à l'évidence différente du contexte qui nous intéresse, ou les énoncés avec extraposition du SN sujet (*c'est que mes enfants quand ils ont venu au monde...* (004/1191)) ou encore les énoncés, quasi inexistant, avec inversion interrogative (*les frères étaient-ils stricts?*). Nous écartons aussi le deuxième, troisième, nième membre de la coordination où le SN sujet de la première proposition n'est pas repris sous sa forme lexicale (*son père _ s'était marié une première fois, puis il avait eu trois filles, puis il s'était remarié encore* (050/0611)). Il en va de même des énoncés sans le SN dans lesquels le pronom sujet est anaphorique, ou joue tout autre rôle, comme l'explétif (*il est bien sur les nerfs le gars, il travaille plus* (007/0921); *il n'y en a pas d'ouvrage* (087/1102). Les énoncés dans lesquels le SN n'est pas au cas nominatif ou est introduit par un terme 'isolant' ou une marque virtuelle de topicalisation sont aussi mis de côté (*j'avais un oncle il venait de—de France, je crés bien* (069/1764); *comme pour mes enfants, pour mes enfants moi-même là ils vont aller à école* (013/0584), *même l'église catholique, ils forment*

des lois d'église (004/1073)). Nous ne tenons pas compte non plus des énoncés aux personnes autres que la troisième puisque le pronom sujet est obligatoire (*moi j'ai des petits bébés, six mois, un mois, deux mois, ...* (007/0136)). Enfin, les énoncés où le verbe n'est pas tensé n'entrent pas en ligne de compte puisque le pronom sujet ne peut pas être licencié (*ça, c'est effrayant, les universités_ faire des choses comme ça* (020/0462); *les enfants_ étant plus jeunes, ben, à ce moment-là, tsé, on est* – (115/0114)).

Le travail de recherche des énoncés à sujet redoublé fait face à quelques difficultés techniques que la consultation conjointe des bandes sonores et de la transcription du corpus permet toutefois de surmonter. En effet, la reconnaissance auditive du pronom sujet est laborieuse dans certains contextes où il subit une érosion phonétique, processus typique de la grammaticalisation et résultat logique de sa haute fréquence. De fait, *il* et *ils* sont réalisés [i] (King et Nadasdi 1995; Léard 1997), avec semi-vocalisation à la liaison avec une voyelle : [j]. Selon Roberge et Vinet (1989 : 53), l'effacement de *l* est un indice du « haut degré de dépendance qui existe entre le clitique et le verbe ». Le *l* est aussi optionnellement effacé avec le pronom sujet féminin à la troisième personne du singulier, qui se réalise [a(l)], la forme du pluriel étant homophone de la forme au masculin [i]. Selon Laforest et coll. (1997), la 'déformation' ne revient pas aux Québécois, on la retrouve depuis longtemps dans de nombreuses variétés orales du français. Ce qui est singulièrement embarrassant, c'est qu'à un niveau plus avancé de la grammaticalisation, et avec la hausse de la liaison structurelle avec le verbe, le pronom sujet tend à s'effacer (Givón 1976; Lehmann 1995), pour atteindre « le maximum d'abstraction structurelle », « le degré zéro de la grammaticalisation, qui annonce la reprise du processus » (De Castillo 1998 : 145). Le scénario est visualisé en (3)

(3) Cycle de grammaticalisation du pronom sujet



Dans la meilleure des situations, il faudrait donc savoir distinguer les contextes d'un zéro représentant une position vide fondamentale non encore occupée par le pronom sujet (redoubleur) et ceux d'un zéro, d'une espèce de trace, où ce dernier a atteint le terminus de la grammaticalisation. Mais cela s'avère très difficile étant donné que même au terme de la grammaticalisation maximale, le processus de régénération du pronom reste possible. Le mouvement cyclique qu'implique la grammaticalisation contredit ici le caractère unidirectionnel (Haspelmath 1998). Dans notre étude, nous optons donc pour un traitement similaire de tous les cas zéro, qu'il s'agisse de zéros fondamentaux ou de zéros résultant d'une grammaticalisation antérieure.

Une autre difficulté a rapport avec la distinction entre la dislocation où le pronom sujet obligatoire s'analyse comme élément de reprise (2a), et le 'vrai' redoublement où il s'interprète comme en train de se grammaticaliser au sein du verbe. Comme nous l'avons dit, la grammaticalisation entraîne une perte de la saillance pragmatique suite à l'effet de la routinisation (Haspelmath), si bien qu'il devient difficile, à l'analyse, de savoir si un locuteur procède à une mise en relief du SN sujet ou pas. Des tests ont été pratiqués pour démêler la situation mais sans être concluants : recherche d'une évidence prosodique, acoustique (Deshaies, Guilbaut et Paradis 1993, Ashby 1994), vérification de l'impact des traits sémantiques tels que la définitude, la quantification, le caractère générique, le statut informationnel, etc. du SN sujet. Par exemple, les sujets indéfinis peuvent être dédoublés par le pronom sujet dans notre corpus : *toute il est à vitesse* (035/0635), *quelqu'un il a mis ça* (013/1722). Quand il s'agit de faits relevant de l'intention inobservable du locuteur, il devient hasardeux d'affirmer certaines choses. Dans les *enfants ils-vont à l'école*, Auger (1998) écrit que la forme 'mixte' *ils-vont* a le trait [+accord fort], mais de quoi s'agit-il au fait? C'est toute la question délicate concernant le rapport forme/fonction qui est soulevée (Lavandera 1978, Labov 1978, Romaine 1984, Milroy 1987, entre autres). Ce qui est en cause, ce sont des traits reliés au SN sujet ainsi que sa position syntaxique, qui exercent, on ne saurait le nier, une influence certaine sur la génération du pronom sujet. Mais il est fondé de focaliser l'attention sur le rapport que le pronom sujet entretient avec le

verbe. La dépendance des pronoms atones a été démontrée par beaucoup d'auteurs (voir Heap et Roberge 2001). Pour renchérir, nous citons Nyrop (1903 : 212) : « Le pronom sujet a tendu de plus en plus à se souder à son verbe; on peut dire qu'aujourd'hui il fait partie intégrante de la forme verbale et peut être considéré comme une sorte de flexion verbale. Ce dernier fait contribue à expliquer son emploi pléonastique dans **mon père il vient**. On dit *j'viens* pour **venio** et de même *i'vient* pour **venit**; la combinaison *mon père i'vient* était donc inévitable ». La polarisation du pronom sujet dans le sillage du verbe est clair chez Wartburg et Zumthor (1947 : 332) : « La langue populaire emploie en général le pronom alors même que le sujet est un nom : *mon frère il chante*; on voit par là à quel degré l'ancien pronom conjoint est devenu un morphème préfixal du verbe. Aucune pause ne sépare, dans la prononciation de cette phrase, le nom du pronom. La phrase se compose de deux mots phonétiques : *mon frère* et *il chante* ». En tout état de cause, ceci nous amène à retenir, non sans risque, nous l'avouons, tous les énoncés, le SN étant ou non disloqué. Une telle démarche a d'ailleurs été suggérée par Coveney (2002) qui juge non pertinente la distinction entre la dislocation et ce qui est souvent considéré comme 'vrai' redoublement du sujet.

4.3 *L'analyse*

5004 énoncés ont été extraits des interviews avec 38 locuteurs. On trouvera dans Kabano (2003) la distribution de ces données selon les locuteurs, et l'analyse quantitative que nous avons effectuée à l'aide de GoldVarb 2001, un programme informatique d'analyse des règles variables. Les étapes de cette analyse, dès la phase de codage des énoncés selon les facteurs linguistiques et extralinguistiques dûment justifiés jusqu'au calcul des fréquences et des probabilités relatives de ces derniers sont exposées en détails. Nous nous intéressons ici à l'effet de la fréquence (d'occurrence et type) des formes verbales sur la génération du pronom personnel sujet. En (4), nous avons trois niveaux de traitement des formes verbales dont la fréquence va être corrélée à la distribution du pronom.

(4) a. FORMES VERBALES FLÉCHIES

(*est, sont, étai(en)t, a, ont, avai(en)t, fait, font faisai(en)t, etc.*)

b. FORMES LEMMATIQUES

(être, avoir, faire, etc.)

c. CATÉGORIE VERBALE GÉNÉRALE

Le premier niveau est celui des formes verbales fléchies individuelles : *va, vont, est, sont, étai(en), fait, font, étai(en)t, veut, veulent, travaille(n)t*, etc. Les formes syncrétiques sont regroupées (par exemple, *était* et *étaient*); nous n'entrons pas dans la discussion à propos du mode d'enregistrement dans la mémoire, problème qui revient à la psycholinguistique. Les formes homophones en français standard mais hétérophones dans le parler que nous étudions sont dégroupées : *joue(nt)* en français standard, mais *joue* et *jousent* en français d'Ottawa-Hull, invention par analogie courante en français canadien (Chambers 1979, Morin 1998). Les formes que l'on pourrait juger 'incorrectes' sont enregistrées telles quelles : par exemple, le même locuteur dit : *parce que si la police _ le sait, c'est d'eux autres d'agir* (010/ 1227), et, un peu plus loin : *c'est ça que la police _ font* (010/ 1243), l'accord se faisant tantôt au singulier, tantôt au pluriel. La démarche consiste donc à coller de près à chaque occurrence effective d'un élément pour faire valoir ce qu'il apporte dans la construction du système. On s'attendra naturellement à un nombre élevé de formes avec 5004 énoncés. Signalons que dans les constructions composées, c'est l'auxiliaire qui est retenu et non la base lexicale qui n'entre pas en accord avec le SN sujet : ainsi dans *si quelqu'un il aurait revenu* (031/0146), c'est *aurait* et non *revenu* qui est retenu. Il va sans dire que le répertoire des formes fléchies ne reflète pas la richesse lexicale de l'univers de discours. Le deuxième niveau est le niveau lemmatique où les formes verbales fléchies sont rattachées à leur lemme, pour abstrait que celui-ci puisse être : *est, sont, étai(en)t, sera, seront*, etc. sont rattachés au verbe *être*; *a, ont, avai(en)t, aura, auront*, etc. sont reliés à *avoir*, etc. Le nombre de lemmes est généralement inférieur à celui des formes fléchies. Il est à prévoir que les groupes lemmatiques n'auront pas le même nombre de membres pour différentes raisons, dont la régularité vs. l'irrégularité des paradigmes de la conjugaison. Enfin, le troisième niveau équivaut à la catégorie générale des verbes, qui est hautement abstraite et qui regroupe en fait tous les ensembles lemmatiques utilisés dans le corpus. Nous nous intéresserons à

ce niveau pour faire ressortir le caractère abstrait et, tout compte fait, très peu opérationnel de l'hypothèse que le verbe en français, en général, est en train de compenser la perte des marques d'accord par le recours au pronom sujet.

5 Les résultats

5.1 *La fréquence d'occurrence*

487 formes verbales différentes sont utilisées dans le corpus. Elles ont été retenues pour des raisons structurelles que nous avons données. Il faut rappeler que ces formes ne représentent pas la richesse lexicale de l'univers de discours qui comprend aussi des formes non tensées (infinitives et participiales). Dans le tableau 1, nous présentons seulement les 22 formes verbales fléchies ayant $N > 30$ dans le corpus; elles ouvrent un total de 3559 énoncés, soit 71% du corpus. Un ensemble de 465 'autres' formes éparses avec chacune $N < 30$ couvrent 1445 énoncés soit 29%; nous ne les énumérons pas ici pour éviter d'être trop fastidieux.

Tableau 1

Distribution des formes verbales et du pronom sujet

| | Formes verbales | | Pronom sujet | |
|------------------------|-----------------|-----|--------------|-----|
| | N | % | N | % |
| <i>est</i> | 795 | 22% | 202 | 25% |
| <i>a</i> | 561 | 16% | 249 | 44% |
| <i>étai(en)t</i> | 498 | 14% | 101 | 20% |
| <i>sont</i> | 331 | 9% | 20 | 6% |
| <i>avai(en)t</i> | 242 | 7% | 87 | 35% |
| <i>ont</i> | 218 | 6% | 82 | 37% |
| <i>va</i> | 141 | 4% | 44 | 31% |
| <i>vont</i> | 83 | 2% | 29 | 34% |
| <i>dit</i> | 73 | 2% | 60 | 62% |
| <i>vient</i> | 64 | 2% | 28 | 43% |
| <i>travaille(nt)</i> | 56 | 2% | 20 | 36% |
| <i>allai(en)t</i> | 54 | 2% | 19 | 35% |
| <i>travaillai(en)t</i> | 52 | 1% | 17 | 32% |
| <i>disai(en)t</i> | 52 | 1% | 23 | 44% |
| <i>parle(nt)</i> | 50 | 1% | 18 | 36% |
| <i>fait</i> | 48 | 1% | 19 | 39% |
| <i>faisai(en)t</i> | 47 | 1% | 17 | 36% |
| <i>venai(en)t</i> | 44 | 1% | 13 | 29% |
| <i>peut</i> | 43 | 1% | 11 | 25% |

| | | | | |
|--------------------|-----------|-----|------|-----|
| <i>reste(nt)</i> | 37 | 1% | 20 | 54% |
| <i>aurai(en)t</i> | 36 | 1% | 13 | 36% |
| <i>voulai(en)t</i> | 34 | 1% | 7 | 20% |
| TOTAL | 3445/5004 | 71% | 1099 | 31% |

Comme on le voit, la distribution des formes verbales est dispersée, mais obéit au principe qui dicte qu'un petit nombre de formes (les 6 premières formes les plus fréquentes dérivent des verbes *être* et *avoir*, suivies de 2 formes du verbe *aller*) soient d'une grande fréquence, et qu'un grand nombre de formes soient d'une basse fréquence. On voit aussi que les formes les plus supplétives du français se maintiennent à une haute fréquence, et que la majeure partie des formes verbales qui entrent en accord avec le SN sujet sont en réalité morpho-lexicalement encodées au point de vue du nombre. Ceci met sérieusement en péril l'hypothèse que le pronom sujet se grammaticaliserait sur le verbe pour remédier à l'incapacité du verbe de marquer l'accord (en termes de coindexation des traits). À partir d'une telle distribution, on se demandera si, de façon générale, les 8 formes verbales ne tendent pas à constituer une composante autonome soumise à un certain nombre de règles grammaticales qui les rapprochent. L'autonomie serait d'autant mieux circonscrite pour deux raisons. D'une part, de nombreuses formes fléchies dérivées des 3 verbes *être*, *avoir* et *aller* (exemples : *soi(ent)*, *sera*, *seront*, *aie(nt)*, *ira*, *iront*, *aille(nt)*, etc.) ne font pas partie du peloton de tête. D'autre part, pour les 8 formes les plus fréquentes, nous avons une disparité de fréquence entre les formes dérivées du même verbe. Par exemple *est* apparaît 795 fois, mais *sont* n'apparaît que 331 fois, moins de la moitié. On aurait des sous-sous-composantes de plus en plus autonomes qui s'emboîtent dans un réseau de connexion complexe.

Observons la fréquence d'occurrence de l'association pronom+verbe. Le fait est qu'elle n'est pas directement corrélée à la fréquence des formes verbales dans le corpus. Ce n'est pas vrai que plus une forme verbale est fréquente plus le pronom sujet a tendance à s'y agglutiner. Une forme comme *dit*, qui réalise 62% d'association avec le pronom sujet, ce qui est un bon score pour un enregistrement de la fusion dans le lexique mental, est malheureusement d'une faible fréquence dans le corpus (2%), et ceci ne peut avoir un effet robuste au sein du système. À

l'inverse, *sont* a le score de 6% seulement, *étai(en)t* a 20%, *est* a 25%, alors que ces 3 formes du verbe *être* couvent 45% des formes fléchies dans le corpus. Ceci n'est pas non plus de nature à contribuer de manière significative à la propagation de la règle d'appariement du pronom et du verbe à l'échelle du système. Le blocage du pronom sujet avec les formes *être* est certainement à approfondir en rapport avec l'hypothèse du degré zéro de la grammaticalisation.

5.2 *La fréquence type*

Si l'on distribue les 1564 occurrences du patron pronom sujet + verbe fléchi selon 4 catégories de verbes déterminés sur la base de leur ordre décroissant de fréquence dans le corpus, le tableau 2 montre la configuration de la fréquence type. Notons que les 465 formes dont $N < 30$ sont cette fois-ci pris en compte, ce qui signifie que notre analyse va inclure tous les énoncés du corpus. Il est important de souligner que les 4 catégories ont été déterminées sur la base, non de propriétés linguistiques, mais du poids numérique relatif des verbes dans le corpus.

Tableau 2

**Distribution du pronom sujet par rapport aux 4 catégories
fréquentielles de verbe**

| | Énoncés | Verbes | Lemmes | Pro sujet |
|--|---------------|--------|--------|--------------|
| <u>Catégorie 1</u> (<i>a, est, étai(en)t</i>) | 1854 (37%) | 3 | 2 | 552 (35%) |
| <u>Catégorie 2</u> (<i>ont, sont, ...</i>) | 1296 (25%) | 10 | 7 | 413 (26%) |
| <u>Catégorie 3</u> (<i>peut, fait, font, ...</i>) | 824 (16%) | 32 | 19 | 257 (16%) |
| <u>Catégorie 4</u> (<i>fera, finit, ...</i>) | 1030 (21%) | 442 | 289 | 342 (21%) |
| TOTAL | 5004 | 487 | 289 | 1564 |

Nous avons fait état d'une distribution inconsistante du patron pronom sujet+verbe fléchi par rapport aux formes verbales individuelles. Insistons sur deux tendances opposées: des formes de grande fréquence dans le corpus ou bien ne s'impliquent pas dans l'accroissement du volume des formes touchées par la fusion (par exemple *est, étai(en)t, sont*) ou bien s'y impliquent, fût-ce à un niveau

légèrement supérieur à la moyenne (par exemple *a, ont*). Cette asymétrie est en soi porteuse d'une tendance au maintien d'un certain équilibre dont on ne pourrait cependant établir la stabilité qu'en diachronie. Nous supposons que ces verbes, vu leur fréquence élevée, ont acquis une autonomie suffisante dans la grammaire pour leur permettre de résister à des changements pour un bon bout de temps. Auger (1998) a peut-être raison quand elle dit qu'il faudra attendre «quelques siècles» encore pour voir la généralisation du redoublement du sujet complétée.

En termes de fréquence type, le tableau 2 montre que 35% des cas d'application de la règle d'agglutination du pronom sujet au verbe concernent 3 formes fléchies (*a, est, étai(en)t*), soit 0,6% du corpus, et, à l'autre extrême, que 21% de l'application de la règle sont éparpillés sur 442 formes verbales différentes, soit 90,7%. Les catégories 1 et 2 mises ensemble, ce sont 13 formes qui s'accaparent 61% de la règle. Une grande partie de l'instanciation de la structure dont on pense qu'elle est en passe de se généraliser est à vrai dire solidement reliée à une petite portion de formes verbales, alors qu'une petite partie des occurrences du pronom sujet sont reliées à une vaste étendue de formes dans une relation qui ne peut être que lâche avec chacune d'elle. D'une manière générale, plus le nombre d'items lexicaux concernés par un patron sera élevé, moins celui-ci sera associé à chacun (Bybee et Thompson 1997; Poplack 2001). L'illusion d'une règle enracinée découle du fait qu'une grande fréquence type, qui se manifeste par un grand nombre de cas d'application de la règle, n'est pas accompagnée de fréquences d'occurrence assez fortes pour les différents items.

Quand on envisage les unités lemmatiques, le constat est le même : 2 verbes (*être, avoir*) sur 289 lemmes, soit 0,69% du corpus sont concernés par 35% des pronoms sujets agglutinés, 7 verbes soit 2,43%, le sont par 61% du patron. Remarquons que les 289 verbes de la catégorie 4 concernées par 21% de pronoms sujets comprennent le reste des verbes, y compris cependant les mêmes lemmes des trois catégories précédentes pour ce qui est de leurs formes conjuguées les plus rares dans le corpus. Des formes verbales reliées à un même lemme sont en effet dégroupées en raison de leur inégalité de fréquence. D'ailleurs, l'effet de la fréquence se répercute, comme on l'a dit, sur la nature linguistique des éléments

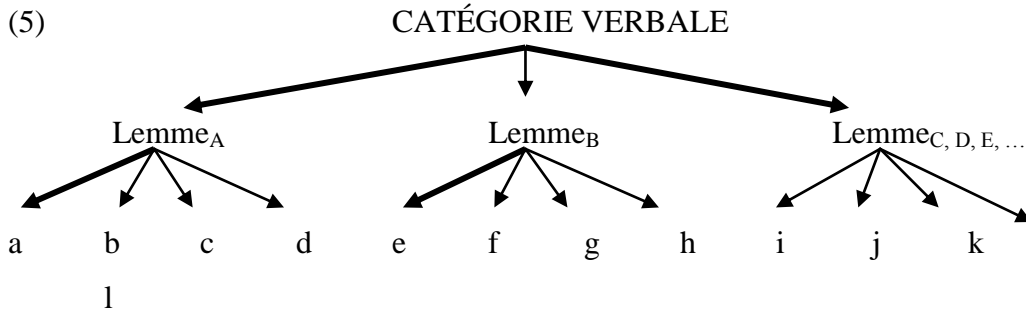
de la langue. Miller (1994) donne un exemple d'idiosyncrasies très intéressant : en français, *je suis* est réalisé *chuis* en France et *chu* au Québec pour le verbe *être*, très fréquent, mais pas pour le verbe *suivre*, peu fréquent, ce qui donne deux unités phonétiques distinctes à la place de deux homophones. Dans tous les cas, l'analyse lemmatique ne rend pas bien compte de ce qui se passe dans la performance pour deux raisons : d'une part, la forme lemmatique reste une donnée abstraite, elle ne s'actualise pas telle quelle, et une contrainte majeure est qu'elle ne licencie pas le pronom sujet, et, d'autre part, le lemme conduit à l'illusion que toutes les formes verbales sous le même lemme ont la même fréquence et se comportent de la même manière, ce qui est réfuté.

Au niveau le plus abstrait, celui de la catégorie verbale, il est en quelque sorte erronée d'affirmer que nous assistons aujourd'hui à un changement (en cours) destiné à rendre obligatoire le pronom sujet en avant du verbe en français. La question qui nous préoccupe relève de la performance et consiste à savoir de quels verbes il s'agit, et une fois que l'on en a une certaine idée, de quelles formes fléchies en usage il s'agit exactement. Nous ne remettons pas en cause le bien-fondé de tout le travail effectué dans le domaine de la théorie qui essaie de capter les lois linguistiques générales avant leur contextualisation. Nous avons simplement montré que des formes spécifiques étaient ciblées dans l'application du patron qui génère le pronom sujet dans la composante morphologique du verbe.

S'il est vérifié que des formes verbales spécifiques, en tout petit nombre, manifestent une certaine tendance à incorporer le pronom sujet dans leur composante lexicale, peut-on dire qu'il s'agit d'une innovation en français aujourd'hui? Nous rappelons que Brunot et Bruneau (1949) mentionnait le fait que les verbes *être*, *avoir* et *faire*, semblaient s'impliquer plus dans le redoublement. Nos résultats indiquent partiellement que cette tendance reste en vigueur, avec certaines formes fléchies des verbes *avoir* et *être*, mais pas avec celles du verbe *faire*. Nous concluons que sur au moins 80 ans, des formes verbales identifiables, et non toute la catégorie verbale, tendent à former corps avec le pronom sujet. La liste des formes verbales se réorganise avec le temps, les taux peuvent aller en augmentant ou

en diminuant, mais il y a quelque chose de stable dans la mécanique sous-tendant l'organisation interne du système et la distribution de ses éléments.

La configuration fréquentielle de la catégorie verbale dans l'usage ainsi que les mécanismes d'activation du patron pronom+verbe sont assez complexes étant donné différents niveaux de leur enracinement (flèche en gras). En (5) nous schématisons le réseau de connexion à l'œuvre.



La littérature psycholinguistique est la mieux placée pour donner des indications expérimentales précises dans les domaines de la production et de la compréhension verbales. Pour ce qui est de la génération du pronom sujet dans le verbe, nos résultats montrent que son exécution reste variable, et s'enracine à différents niveaux. Nous retiendrons trois cas : i) avec la forme lemmatique A, la tendance à incorporer le pronom sujet est enracinée dans la langue, permettant à la forme (a) d'être également enracinée (par exemple *a* du verbe *avoir*), mais pas aux formes (b), (c) ou (d) (par exemple, *aurai(en)t*, *aie*, etc.); ii) la forme lemmatique B n'est pas enracinée dans la langue à titre de site du pronom sujet, mais une forme dérivée comme (e) exhibe tant soit peu une propension pour la règle (par exemple, la forme *dit* qui a un taux élevé de fusion avec le pronom sujet, mais qui a une basse fréquence en discours, au moment où d'autres formes (f), (g) et (h) dérivées de *dire* ne se manifestent pas); iii) dans le cas de C, soit une forme lemmatique enracinée est actualisée par des formes fléchies multiples qui ne sont pas, elles, enracinées en raison de leur faible fréquence; soit différentes formes éparses (la catégorie 4 (tableau 2)) contribuent en vrac à l'enracinement du patron en raison de leur fréquence type élevée, sans être reliées à une seule forme lemmatique.

Conclusion

Nous avons tenté de circonscrire le domaine verbal d'application de la règle selon laquelle les formes verbales tendraient à être sélectionnées déjà dotées du pronom sujet en français, selon la version forte de la thèse minimaliste. Notre investigation a porté sur des données de performance actuelles et non sur des exemples fabriqués. Que pourrions-nous conclure à propos de la productivité à venir du système? Une chose est claire : on ne peut demander à la règle, comme des linguistes en ont l'habitude, ce qu'on pourrait appeler la prédictibilité absolue (Langaker 1987). La variation dans la fusion pronom sujet+verbe (en contexte de redoublement) est loin de répondre à une prescription universelle puisque, réalisée dans 31% des énoncés, cette fusion, qui n'est pas complétée, ne concerne qu'une petite portion parmi les formes verbales fléchies de haute fréquence et une multitude d'autres formes avec un enracinement plutôt très faible. Même si nous supposons que la grammaticalisation du pronom sujet s'amplifiait, ce ne serait qu'avec des formes toujours précises, étant donné qu'il est clair que la majeure partie des verbes sont voués à s'utiliser en dehors de la position gouvernée par AGR situé sous I et n'entrent pratiquement jamais en accord avec le SN sujet. Le fait de tableur sur des données d'usage nous empêche donc de généraliser sur la classe des verbes, et même sur des unités lemmatiques, une tendance, si tant elle existe, qui reste concentrée sur un petit sous-ensemble. Il est des faits de langue comme de la distribution des richesses : quelques riches **ils** deviennent de plus en plus riches, et un grand nombre de pauvres _ deviennent de plus en plus démunis.

Références

- Aitchison, Jean (1996). *The Seeds of speech : language origin and evolution*. Cambridge : Cambridge University Press.
- Ashby, William (1982). The syntax, pragmatics and sociolinguistics of left- and right-dislocation in French. *Lingua* 71 (2-3) : 203-229.
- Auger, Julie (2003). Le redoublement des sujets en picard. *Journal of French Language Studies* 13 (3). 381-404.

- Auger, Julie (1998). Le redoublement des sujets en français parlé informel. *Revue québécoise de linguistique*. 24 (1) : 21–26.
- Auger, Julie (1994). *Pronominal clitics in Québec colloquial French: a morphological analysis*. University of Pennsylvania (Thèse de doctorat).
- Blasco-Dulbecco, Mylène (1999). *Les dislocations en français contemporain : études syntaxique*. Paris : Honoré Champion.
- Blinkenberg, Andreas (1928). *L'ordre des mots en français moderne I*. Copenhagen : Blanco Lunos Bogtrykkeri.
- Brunot, Ferdinand (1922). *La pensée et la langue. Méthodes, principes et plan d'une théorie nouvelle du langage appliquée au français*. Paris : Masson et Cie.
- Brunot, Ferdinand et Charles Bruneau (1949). *Précis de grammaire historique de la langue*. Paris : Masson.
- Bubenik, Vit (1993). Inflectional morphology and clitics in functional grammar. Typology and diachrony. *Diachronica* 10 (2). 165–190.
- Bybee, Joan (1995). Regular morphology and the lexicon. *Language and cognitive processes*. 10 (5). 425–455.
- Bybee Joan, Revere Perkins et William Pagliuca (1994). *The evolution of grammar : Tense, aspect, and modality in the languages of the world*. Chicago et London : The University of Chicago Press.
- Bybee, Joan (2003). Cognitive processes in grammaticalization. Dans Michael Tomasello (dir.) *The New Psychology of Language 2*. New Jersey : Lawrence Erlbaum Associates Inc. 145–167.
- Bybee, Joan (2001). Frequency effect on French liaison. Dans Joan Bybee et Paul Hopper (dirs.) *Frequency and the emergence of linguistic structure*. Amsterdam : John Benjamins. 337–359.
- Bybee, Joan et Sandra Tompson (2000). Three frequency effects in syntax. *Berkely Linguistic Society* 23. 378–388 .
- DeCastillo, Ataliba Teixeira (1998). Langue parlée et processus grammaticaux. Dans Mireille Bilger, Karel van den Eynde et Françoise Gadet (dirs.).

Analyse linguistique et approche de l'oral. Leuven, Paris : Peeters. 141–148.

De Cat, Cécile (2005). French subject clitics are not agreement markers. *Lingua* 115 (9). 1195-1219.

Deshaies, Denise, Christian Guilbault et Claude Paradis (1993). Prosodie et dislocation à gauche par anaphore en français québécois spontané. Dans André Crochetière, Jean-Claude Boulanger et Conrad Ouellon (dir.), *Actes du XVe Congrès international des linguistes 2*. Québec : Presses de l'Université Laval. 31–34.

Friedemann, Marc-Ariel (1997). *Sujets syntaxiques. Positions, inversion et pro*. Berne : Peter Lag.

Givón, Talmy (1995). *Functionalism and grammar*. Amsterdam : John Benjamins.

Givón, Talmy (1984). *Syntax : a functional-typological introduction 1*. Amsterdam : John Benjamins.

Givón, Talmy (1976). Topic, pronoun, and grammatical agreement. Dans Charles N. Li (dir.). *Subject and topic*. New York : Academic Press. 149–188.

Grévisse, Maurice (1975). *Le bon usage. Grammaire française avec des remarques sur la grammaire française d'aujourd'hui*. Gembloux : Duculot.

Haspelmath, Martin (1998). Does grammaticalization need reanalysis? *Studies in languages* 22 (2). 315–351.

Heap, David et Yves Roberge (2001). Cliticisation et théorie syntaxique, 1971–2001. *Revue québécoise de linguistique*. 30 (1) : 63–90.

Heap, David (2000). *La variation grammaticale en géolinguistique: les pronoms sujets en roman central*. München : Lincom Europa.

Heine, Bernd, Ulrike Claudi, et Friederike Hünemeyer (1991). *Grammaticalization: A conceptual framework*. Chicago : University of Chicago Press.

Hopper, Paul et Elizabeth Traugott (1993). *Grammaticalization*. Cambridge, UK : Cambridge University Press.

- Hopper, Paul (1998). The paradigm at the end of the universe. Dans Anna Giacalone Ramat et Paul Hopper (dirs.). *The limits of grammaticalization (Typological studies in Language 37)*. 147–158.
- Jolivet, Remi (1982). *Descriptions quantifiées en syntaxe du français : approche fonctionnelle*. Genève, Paris : Slaktine.
- Judge, Anne et Frank George Healey (1983). *A reference grammar of modern French*. London : E. Arnold.
- Kabano, Alphonse (2004). *Le pronom personnel sujet de la troisième personne dans le parler d'Ottawa-Hull : grammaticalisation, réanalyse, variation*. Ottawa : Université d'Ottawa (Thèse de doctorat).
- King, Ruth et Terry Nadasdi (1995). La puissance des pronoms faibles en français terre-neuvien. Dans Donna L. Lillian (dir.), *Proceedings of the Annual Meeting of the Atlantic Provinces Linguistic Association* 19. 129–137.
- Klapka, (2002). *Étude comparative: l'accord du genre en français parlé québécois au 19^{ème} et au 20^{ème} siècle*. Ottawa : Université d'Ottawa (Ms).
- Kroch, Anthony (1989). Reflexes of grammar in patterns of language change. *Language variation and change* 1. 199–244.
- Labov, William (1994). *Principles of linguistic change I*. Oxford, Cambridge : Blackwell.
- Laforest, Marty et coll. (1997). *États d'âme, états de langue. Essai sur le français parlé au Québec*. Québec : Nuit Blanche Éditeur.
- Langacker, Ronald (2000). A dynamic usage-based model. Dans Michael Burlow et Suzanne Kemmer (dir.). *Usage based models of language*. Stanford : CSLI. 24–63.
- Langacker, Ronald (1983). *Foundations of cognitive grammar*. Stanford : Stanford University Press.
- Lanly, André (1970). *Le français d'Afrique du Nord. Étude linguistique*. Paris, Montréal : Bordas.
- Lavandera, Beatriz (1978). Where does the sociolinguistic variable stop? *Language in society* 7. 171–182.

- Leard, Jean-Marcel (1997). *Grammaire québécoise d'aujourd'hui. Comprendre les québécismes*. Montréal : Guérin Universitaire.
- LeBlanc, Carmen (1999). *Du conditionnel dans les propositions hypothétiques en si : cet intrus*. Ottawa : Université d'Ottawa (Ms).
- Lehmann, Christian (1995). *Thoughts on grammaticalization*. München : Lincom Europa.
- Lehmann, Christian (1987). On the function of agreement. Dans Michael Burlow et Charles A. Ferguson (dirs.). *Agreement in natural language: approaches, theories, descriptions*. Stanford : CSLI. 55–65.
- Lyons, John (1990). An agreement approach to clitic doubling. *Transactions of the Philological Society* 88 (1). 1–57.
- McWinney, Brian (1978). *Mechanisms of language acquisition*. Hillsdale N.J : Lawrence Erlbaum Associates.
- Manzini, M. Rita et Leonardo M. Savoia (2002). Parameters of subject inflection in Italian dialects. Dans Peter Svenonius (dir.). *Subjects, Expletives and the EPP*. Oxford : Oxford University Press. 157–200.
- Miller, Philip (1992). *Clitics and constituency in Phrase Structure Grammar*. New York : Garland.
- Milroy, (1987). *Observing and analyzing natural language: a critical account of sociolinguistic method*. Oxford : Blackwell.
- Moignet, Gérard (1965). *Le pronom personnel français : essai de psychosystématique historique*. Paris : Klincksieck.
- Nadasdi, Terry (1995). NP doubling, matching, and minority French. *Language variation and change* 7. 1–14.
- Nyrop, Kristoffer (1903). *Grammaire historique de la langue française*. Copenhagen : Gyldendalske Boghandel.
- Poplack, Shana (2001). Variability, frequency, and productivity in the irrealis domain of French. Dans Joan Bybee et Paul Hopper (dirs.). *Frequency and the emergence of linguistic structure*. Amsterdam : John Benjamins. 405–428.

- Poplack, Shana (1992). The inherent variability of French subjunctive. Dans Chris-tiane Laeufer et Terrell A. Morgan (dirs.). *Theoretical analyses in Romance linguistics (Current issues in linguistic theory)*. Amsterdam : John Benja-mins. 235–262.
- Poplack, Shana (1989). The care and handling of a mega-corpus : the Ottawa-Hull French projet. Dans Ralph Fasold et Deborah Schiffrin (dirs.). *Language change and variation*. Amsterdam : John Benjamins. 411–451.
- Roberge, Yves et Marie Thérèse Vinet (1989). *La variation dialectale en gram-maire universelle*. Montréal, Sherbrooke : Les Presses de l'Université de Montréal, Les Presses de l'Université de Sherbrooke.
- Sandfeld, Kraus (1928). *Syntaxe du français contemporain*. Paris : Honoré Cham-pion.
- von Wartburg, Walter et Paul Zumthor (1958). *Précis de syntaxe du français con-temporain*. Bern : A. Francke AG. Verlag.
- Weinreich, Uriel, William Labov et Marvin I. Herzog (1968). Empirical founda-tions for a theory of language change. Dans Winfred Philipp Lehmann et Yakov Malkiel (dirs.). *Directions for historical linguistics: a symposium*. Austin : University of Texas Press. 95–189.
- Winford, Donald (1996). The problem of syntax variation. Dans Jennifer Arnold *et coll.* (dirs.). *Sociolinguistic variation : data, theory, and analysis*. Selected papers from NWAV 23 at Stanford. Stanford: CSLI. 177–192.
- Willis, L.A. (2000). *Être ou ne pas être : auxiliary alternation in Ottawa-Hull French*. Ottawa : University of Ottawa, Ms.