

Qu'en est-il des clitiques sujet en français oral contemporain?¹

Jenny Culbertson & Géraldine Legendre

Johns Hopkins University

1 Introduction

L'histoire de la Grammaire Générative se confond avec celle d'un type de données empiriques privilégiées, à savoir les jugements de grammaticalité provenant le plus souvent de linguistes natifs de la langue étudiée. Même si elle a ses détracteurs, la méthode s'est révélée très féconde au cours des dernières décennies. Elle a permis en autres de développer des argumentations complexes et de jeter la lumière sur des interactions jusque-là insoupçonnées.

Mais la méthode a aussi ses limites et il arrive qu'elle obscurcisse des faits linguistiques qui portent sur des questions théoriques majeures. C'est le cas, par exemple, des pronoms clitiques sujet en français dont le statut d'argument en position sujet reste controversé (e.g. Auger 1995, Roberge 1990, Rizzi 1986, Kayne 1975, Miller & Sag 1997, Miller & Monachesi 2003). Mis à part quelques études comme Fonseca-Greber & Waugh (2003) et Zribi-Hertz (1994), les données sur lesquels s'appuient des conclusions contradictoires proviennent de jugements de grammaticalité portant sur le français dit standard. S'il s'avérait que les clitiques sujet ont le statut d'affixe flexionnel (plutôt que celui d'argument en position sujet) alors le français serait à classer comme langue à sujet nul et bien des analyses syntaxiques basées sur les corrélations entre sujet nul, inversion libre du sujet, etc. seraient à revoir.

Nous proposons de réexaminer les clitiques sujet à la lumière de nouvelles données provenant de sources autres que les intuitions de linguistes, y compris des analyses de parole adressée à l'enfant, des analyses prosodiques de stimuli créés expérimentalement, et des considérations typologiques. Nous montrerons que dans le registre *oral et spontané* du français contemporain, les clitiques sujet ont définitivement changé de statut. Dans un premier temps nous nous concentrerons sur le phénomène du redoublement du sujet et montrerons que le sujet est bien redoublé, et non disloqué, en français oral contemporain. Si le sujet est redoublé par un clitique c'est que le clitique n'a pas le statut de sujet mais celui d'affixe flexionnel. Nous proposerons une analyse inspirée des travaux de Suñer (1992) et Poletto (2000) selon laquelle le redoublement est un phénomène d'accord grammatical. Dans un deuxième temps nous montrerons que les faits empiriques qui sont le plus souvent cités comme évidence du statut syntaxique et non morphologique des clitiques sujet, à savoir la rétention de la particule négative *ne*, l'inversion du clitique sujet dans les questions, et la répétition non-obligatoire du clitique sujet dans la coordination, ne survivent pas à un examen détaillé de corpus de parole adressée à l'enfant et qu'au contraire, les clitiques sujet ont clairement le statut d'affixe flexionnel dans le langage adressé aux jeunes enfants. Le phénomène qui nous intéresse est donc un phénomène diachronique au terme de son évolution alimenté par l'acquisition du français génération après génération.

2 Le redoublement clitique du sujet

L'analyse de la construction au cœur du débat concerne celle illustrée en (1). S'il s'agit de redoublement du SN en position sujet par un clitique alors le clitique lui-même ne peut avoir de statut autre que celui d'affixe flexionnel. Par contre, si (1) est un exemple de dislocation (à gauche) du SN sujet, alors le clitique peut en principe occuper la position sujet (SpecIP).

- (1) a. Jean **il** parle.
b. Personne (***il**) ne parle.

Le redoublement du sujet existe bien dans des langues apparentées au français tels que le parler pied noir (Roberge 1990) et le picard (Auger 2003) mais son existence en français reste à prouver. En fait, le contraste en (1) est fréquemment cité comme preuve que la construction en (1a) est un cas de dislocation et non un cas de redoublement. Puisque les SN quantifiés ne sont jamais des topiques, ils ne peuvent pas être disloqués; ils doivent occuper la position sujet (SpecIP). Les SN non-quantifiés tels que *Jean* en (1a) ne sont pas soumis à la même restriction. Lorsque le sujet est un topique, specIP est occupé par un pronom résomptif. L'apparition du clitique correspond donc à la possibilité de disloquer le sujet. Conclusion: la construction en (1a) est une dislocation (Rizzi 1986, De Cat 2007). En l'absence de dislocation, le pronom clitique et le SN sont en compétition pour la même position, specIP ; Il en résulte deux possibilités :

- (2) a. Jean parle.
b. Il parle.

Toute hypothèse auxiliaire mérite examen et il est donc en principe possible que l'agrammaticalité de (1b) ne soit pas due au statut discursif des syntagmes quantifiés. Par exemple, on peut noter en (1b) que le clitique *il* porte le trait [défini] alors que *personne* porte le trait [indéfini]. Il suffirait qu'on impose une contrainte d'appariement de traits pour rendre compte du contraste entre (1a) et (1b). Il importe donc d'établir que la dislocation a d'autres propriétés caractéristiques, par exemple une prosodie particulière, et de confirmer qu'elle est bien restreinte aux topiques. Et même si l'analyse se révélait correcte pour le registre de langue dit standard, il est loin d'être sûr qu'elle s'applique au registre oral contemporain.

2.1 Analyse prosodique du redoublement

Il est généralement admis que la dislocation à gauche se distingue par une prosodie particulière et certains indices acoustiques. En particulier, Doetjes et al (2002) ont identifié une signature acoustique des éléments disloqués à gauche, à savoir (i) une montée de F0 sur l'ensemble de l'élément disloqué (avec la cible F0 atteinte au niveau du noyau ou à la fin de la syllabe finale accentuée) et (ii) un allongement de la durée de la syllabe finale:

- (3) *A mon prof de linguistique H% [cont]] je lui ai donné un livre L%]*
(Doetjes et al. 2002: 3)

Mais Doetjes et al. ne différencient pas les éléments disloqués en fonction sujet de ceux en fonction objet, ce qui nécessite une étude expérimentale complémentaire présentée ci-dessous.

Notons d'abord les prédictions générales concernant la prosodie comparée de (1a) et (2a) selon qu'il s'agit en (1a) d'une structure disloquée ou d'un redoublement du sujet. Si le clitique occupe la position de sujet (specIP), la construction est une dislocation comparable à celle où un syntagme objet est disloqué, et on s'attend à une prosodie particulière de la construction en (1a), comparée à (2a) et (3). Par contre, si le clitique est un affixe flexionnel, la prosodie de (1a) devrait être similaire ou même identique à celle de (2a) et différente de celle identifiée par Doetjes et al. en (3). Contrairement à Doetjes et al. qui ne cherchent pas à établir la robustesse de leurs résultats, nous nous baserons sur des tests statistiques pour étayer nos conclusions.

2.1.1 Matériel et Procédure

Les stimuli comprennent des paires quasi-minimales de trois types de phrases cibles: SN sujet non-redoublé (4a), SN sujet redoublé (4b), et SN objet disloqué (4c). Le souci principal a été de minimiser au maximum les différences parmi ces phrases et de les intégrer individuellement à des petits dialogues quasi-identiques.

Chaque dialogue concerne un échange d'information entre deux amis. Cinq dialogue-types ont été créés ainsi que trois versions de chaque dialogue-type pour favoriser l'interprétation voulue pour chaque phrase cible puisque la phrase elle-même peut être ambiguë (4b vs. 4c).

- (4) a. David l'a déjà invité. (David = sujet)
b. David il l'a déjà invité. (David = sujet)
c. David il l'a déjà invité. (David = objet)

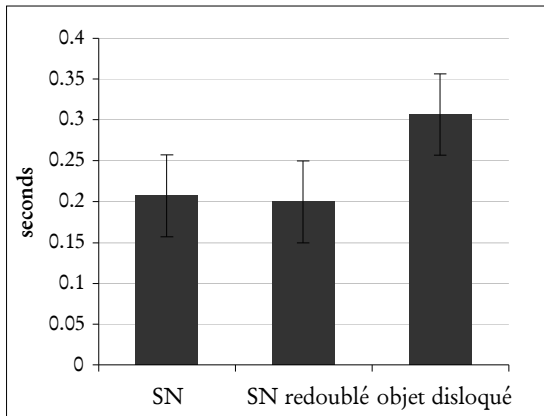
Huit sujets de moins de 30 ans, de langue maternelle française, récemment arrivés aux Etats-Unis, ont été recrutés. L'expérience s'est déroulée à Johns Hopkins University, Baltimore (USA). Les sujets ont été testés par paires, chacun assumant un rôle dans les dialogues présentés par écrit. Ils ont reçu l'instruction de prétendre qu'ils échangeaient une conversation « spontanée » entre amis. Pour faciliter la tâche les dialogues contenaient les marques typiques d'oral spontané (*David i l'a déjà invité*, élision de *ne*, certains *e*, vocabulaire familier, etc.) Pour chaque paire, un seul rôle contenait les phrases cibles; un des sujets ne servait qu'à donner la réplique. Les données correspondent donc à quatre locuteurs.

Le SN initial (sujet ou objet) de chaque phrase cible a été analysé au moyen du logiciel Praat. Chacun des quatre locuteurs a produit cinq exemplaires de chacun des trois types de phrase cible donc 20 exemplaires de chaque type pour un total de 60 syntagmes initiaux analysés. Trois mesures ont été faites : durée de la syllabe finale, montée de F0 sur toute la syllabe, et augmentation d'intensité sur toute la syllabe. La troisième mesure, l'augmentation d'intensité, est un ajout par rapport à Doetjes et al. (2002) car ce facteur est parfois mentionné comme pertinent (De Cat 2007, Deshaies et al. 1993). Deux traits acoustiques supplémentaires ont été étudiés, à savoir la présence d'une pause entre le SN initial et le reste de la phrase, et l'enchaînement ou déplacement de la frontière syllabique à gauche (*sept amis* => [se-ta-mi]).

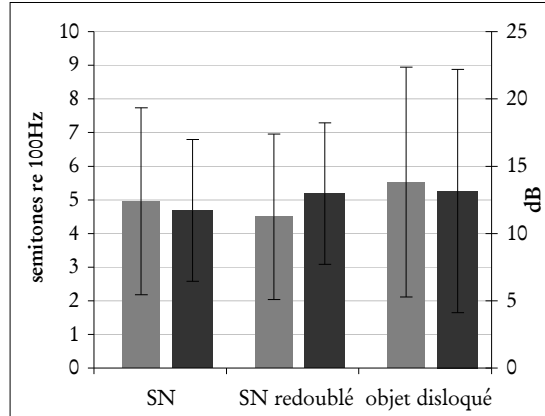
2.1.2 Résultats

Des ANOVAs ont été effectuées sur chacune des trois mesures. Comme le montre le Graphe 1, un effet significatif du facteur type de phrase a été trouvé pour la durée de la syllabe finale ($F(2,57)=17.4$, $p<0.001$). Des comparaisons planifiées ont confirmé que la durée moyenne pour la condition objets disloqués était significativement plus longue que celle pour les conditions sujets redoublés et non-redoublés ($F(1,57)=35.5$, $p<0.01$). Il n'y avait pas de différence significative entre les conditions sujets redoublés et non-redoublés ($F(1,57)=0.14$, $p=0.71$). Ces résultats suggèrent qu'en terme de durée de la syllabe finale, seuls les objets disloqués montrent l'effet d'allongement attendu pour les dislocations, alors que les sujets redoublés *ne montrent pas* cet effet. En fait, les sujets redoublés ne se distinguent pas des sujets non-redoublés sur cette dimension acoustique.

Par contre, les ANOVAs n'ont pas fait apparaître de différences significatives entre conditions pour les mesures de F0 ($F(2,57)=0.63$, $p=0.54$) et de montée d'intensité ($F(2,57)=0.27$, $p=0.67$), comme le montre le Graphe 2 présentant les moyennes pour la F0 et la montée d'intensité à travers la syllabe finale de chaque DP pour les 3 conditions. Ces résultats suggèrent que la F0 et l'intensité pourraient ne pas être utilisées pour marquer de façon robuste les différences entre ces trois types de structures.



Graph 1. Allongement de durée syllabique en position finale de SN



Graph 2. F0 (gris) et montée d'intensité (noir) en position finale de SN

Pris ensemble, ces résultats suggèrent que l'indice le plus fiable pour distinguer les SNs disloqués à gauche de SNs non-disloqués serait la durée de la syllabe finale (accentuée). Cette conclusion est en fait suggérée par De Cat (2007) mais notre étude est la première à apporter des éléments la confirmant quantitativement. De plus, nos données soutiennent la prédiction que les SN objets dissociés à gauche, mais pas les SN sujets redoublés, sont marqués par des effets d'allongement, en accord avec une analyse des clitiques sujet en termes d'affixes flexionnels. De façon évidente, ces résultats sont incompatibles avec les prédictions d'une analyse argumentale des clitiques sujet selon laquelle les sujets redoublés devraient se comporter comme d'autres syntagmes dissociés à gauche.

2.2 Les contextes à redoublement

En l'absence de marques prosodiques/acoustiques, il est possible de maintenir que d'un point de vue syntaxique, la construction en (1a) est une dislocation, peut-être d'un type différent de celui de la dislocation à gauche du SN objet. En fait, De Cat (2007) observe que la construction qui nous concerne est agrammaticale lorsque le sujet est interprété comme focus ou information nouvelle.

(5) Question: Qui a fini son travail ? Réponse: Jean (*il) a fini son travail.

Il est en général admis que les énoncés à focus sont de simples IPs en français (e.g. Samek-Lodovici 2005) mais identifiés comme focus large ou focus étroit par "des tons de frontière illocutoires reflétant le partage entre information nouvelle et information ancienne" (Marandin 2004).

(6) Qu'est-ce qui se passe?
 a. Bernadette est venue L%
 b. #Bernadette L% est venue L%

(7) Qui est venu?
 a. #Bernadette est venue L%
 b. Bernadette L% est venue L%

D'après les résultats d'une tâche de jugement de grammaticalité conduite par De Cat (2007) les locuteurs français n'acceptent pas le redoublement clitique dans les énoncés à focus; ce qui la mène à conclure que la présence d'un clitique signale une dislocation. Cependant, le redoublement clitique est obligatoire en français oral contemporain en présence d'un pronom fort (*moi*), comme le montre (8a).

- (8) a. Moi **je** parle beaucoup.
b. Question: Qui a appelé Jean ? Réponse: Moi **je** l'ai appelé.

Fonseca-Greber & Waugh (2003) n'identifient aucun exemple de pronom fort sans redoublement clitique dans leur corpus de français oral contemporain, et nous n'avons rencontré qu'un seul exemple de pronom fort (*lui*) sans redoublement clitique parmi 589 exemples de pronoms forts dans le corpus de Lyon (Demuth & Tremblay 2008; CHILDES, MacWhinney 2000). A noter: le redoublement clitique est possible même lorsque le pronom fort est interprété comme focus (étroit) comme en (8b). Cependant, les questions et réponses à focus en français oral contemporain sont le plus souvent des structures clivées (voir 9); (4) et (8) n'y sont donc pas les structures les plus communes. Dans le corpus de Lyon, par exemple, toutes les questions à focus en position sujet et objet sont des structures clivées.

- (9) Question : Qui c'est qui tire la flèche ? R : C'est Kiwiwi. (Ana40, line 812)

Il n'en reste pas moins que le redoublement clitique obligatoire en présence d'un pronom fort et la possibilité de redoublement clitique dans les énoncés à focus montre que les clitiques sujet peuvent en fait coexister avec un SN dans de simples IPs, et donc qu'ils ne sont pas en compétition avec le sujet nominal en ce qui concerne la position (specIP) et le rôle d'argument. Toutefois, le fait que le redoublement clitique n'est pas permis dans les énoncés à focus comme (4) suggère que des contraintes existent sur le phénomène. Des considérations typologiques révèlent qu'il s'agit de contraintes générales sur l'accord grammatical.

2.3 Considérations typologiques

D'après Brandi & Cordin (1989) et Rizzi (1986) le redoublement est permis avec les sujets quantifiés et indéfinis dans les groupes dialectaux du nord de l'Italie, ce qui les différencie du français (1b). Dans une étude plus récente et plus large, Poletto (2000) note un comportement variable parmi ces groupes dialectaux. Les dialectes en question sont organisés sur la base d'une hiérarchie implicationnelle des traits du SN sujet:

- (10) Variables Q → SN quantifiés → SN non-quantifiés → pronoms forts

Il résulte de (10) que dans les dialectes du nord de l'Italie le redoublement peut être limité aux seuls pronoms forts; si le redoublement est possible avec des SN non-quantifiés il est aussi possible avec les pronoms forts. De même, si le redoublement est possible avec un sujet quantifié ou indéfini il est aussi possible avec les SN non-quantifiés et les pronoms forts. Enfin, si le redoublement est possible avec une variable Q telle qu'un pronom relatif, il s'en suit que le redoublement est possible avec tous les types de sujet.

La possibilité de redoublement ou non ne peut être réduite à une simple question de dislocation du sujet. Pour accommoder la distribution du redoublement dans les dialectes du nord de l'Italie et le français (dont ils sont proches), il faudrait corrélérer la dislocation avec une position donnée dans la hiérarchie en (10). Cela reviendrait à dire que dans certaines langues seuls les pronoms forts *doivent* être disloqués alors que dans d'autres langues seuls les syntagmes nominaux non-quantifiés et les pronoms forts *doivent* être disloqués, etc. D'autre part, le fait que (10) sépare les variables Q des syntagmes quantifiés montre bien que le statut des clitiques sujet (argument ou affixe flexionnel?) dans une langue donnée ne peut être établi sur la base du simple redoublement ou non en présence d'un SN quantifié, contrairement à la tradition (De Cat 2007, Auger 1995, Brandi & Cordin 1989, Rizzi 1986). Il nous paraît plus judicieux de rendre compte des propriétés des clitiques sujet et du redoublement en d'autres termes.

De plus, les groupes dialectaux du nord de l'Italie et l'espagnol suggèrent que d'autres traits des syntagmes nominaux entrent en jeu dans le redoublement. En espagnol (y compris divers dialectes) la possibilité de redoubler un SN en fonction d'objet est liée aux traits [+/- animé] et [+/- défini] du syntagme. Suñer (1992) propose d'y ajouter le trait [+/- accessible] ce qui voudrait dire que la hiérarchie en (9) devrait en fin de compte inclure les SN [+/- accessible].

D'un point de vue typologique le français oral contemporain occupe une place intermédiaire parmi les dialectes du nord de l'Italie, entre par exemple le vénitien qui restreint le redoublement aux seuls pronoms forts, le dialecte trentin qui nécessite le redoublement avec les SN quantifiés mais pas les SN non-quantifiés, et le florentin qui nécessite le redoublement dans tous les contextes.

2.4 L'hypothèse d'appariement des traits

A la suite de Suñer (1988, 1992) nous proposons que le redoublement clitique est sujet à une contrainte de d'appariement (Matching Hypothesis) : Chaque clitique a des traits spécifiés qui doivent s'apparier à ceux du SN qu'il redouble. En français oral contemporain, le clitique sujet est une marque d'accord avec le sujet présente lorsque les traits du sujet correspondent aux traits du clitique et obligatoirement absente dans le cas contraire.

Nous posons que les traits des clitiques sujet comprennent la personne, le genre, le nombre, mais aussi la spécificité (Nadasdi 1995) et l'accessibilité au niveau du discours (Suñer 1992). Selon Lambrecht (1981) un référent de discours est accessible s'il est évoqué textuellement ou situationnellement, ou s'il peut être inféré du discours précédent. A la suite de Lambrecht (1981), le trait [+/- défini] est défini comme un corrélat du trait [+/- accessible].

En vertu de cette analyse, l'absence de redoublement avec les SN sujets quantifiés et indéfinis (sauf de rares exemples comme *tout le monde il est beau, tout le monde il est gentil* (Zribi-Hertz 1994) n'est pas due à l'impossibilité de disloquer ces syntagmes mais a un clash fatal entre leurs traits [-défini, -spécifique] et ceux des clitiques sujet comme *il* qui porte les traits [+accessible, +spécifique]. De même, les pronoms relatifs et interrogatifs ne permettent pas le redoublement parce qu'ils sont spécifiés pour des traits similaires à ceux des syntagmes quantifiés.²

Notons que l'analyse prédit les restrictions observées dans les énoncés à focus. Porteur d'information nouvelle, le syntagme en focus est [-accessible] et son redoublement avec un clitique [+accessible] donc impossible. Par contre, si un énoncé à focus concerne un référent qui est [+accessible] tel que les pronoms forts *moi, toi*, il est prédit que le redoublement devient possible, ce qui est confirmé en (8).

L'accord est donc exprimé par le moyen d'un clitique avec tout sujet portant les traits [+défini, +spécifique, +accessible], y compris le topique discursif, les pronoms forts (qui sont nécessairement [+accessible]), ainsi que les SN [-défini] qui sont non-quantifiés et donc [+accessible].³ Il est à noter que le français a un clitique sujet désigné comme [+accessible, -spécifique]: *ça*, et qu'en vertu de notre hypothèse il permet le redoublement avec les SN génériques du type (11).⁴

(11) Un enfant *ça* pleure beaucoup.

Dans cette perspective, l'optionalité du redoublement en français oral contemporain n'est qu'apparente. Le redoublement, en fait, est nécessaire lorsque les traits s'accordent. Il résulte de cette analyse que le français oral contemporain est à classer comme une langue à sujet nul du même type que les dialectes du nord de l'Italie (Poletto 2000, Brandi & Cordin 1989, Rizzi 1986). De plus, le français oral contemporain est une langue qui permet la redondance dans la marque de l'accord sujet-verbe; malgré l'appauvrissement phonologique considérable de la marque d'accord en position suffixale pour plus de 90% des verbes (Dietiker 1978) il reste des verbes irréguliers où la marque d'accord en position suffixale a survécu et coexiste avec la marque d'accord préfixale. Cette redondance n'est pas unique au français oral contemporain puisqu'elle caractérise aussi la plupart des dialectes du nord de l'Italie ainsi que le picard où le redoublement est obligatoire pour tous types de sujets et coexiste avec une morphologie suffixale relativement riche selon Auger (2003).

3 La parole adressée à l'enfant et le rôle de l' acquisition dans l'évolution diachronique des clitiques sujet

Il est indéniable que l'acquisition d'une langue joue un rôle dans son évolution diachronique (Lightfoot 1999; Croft 2000). Par conséquent, il est nécessaire d'examiner la nature de l'input langagier que reçoivent les apprenants pour comprendre un processus évolutif en cours ou à son terme. Une spécification de la grammaire adulte est nécessaire mais elle demeure insuffisante pour deux raisons principales: (i) la fréquence d'utilisation d'un tour est rarement en rapport étroit avec la grammaire synchronique mais elle joue un rôle majeur dans l'évolution de la langue (Kroch 2000), (ii) la parole adressée à l'enfant peut avoir des propriétés uniques distinctes de la parole entre adultes, ce qui peut influencer le cours ou le résultat du processus acquisitionnel.

L'évolution diachronique des pronoms indépendants – de leur statut d'argument à celui d'affixe flexionnel – suit un cycle de réduction sur plusieurs dimensions. D'une part, leur évolution est guidée par une réduction phonologique continue, et de l'autre par une perte de traits référentiels (Fuß 2005, Hopper & Traugott 1993). C'est le cas en français oral contemporain où les clitiques sujet sont phonologiquement extrêmement réduits (*je* = [ch] dans *j'suis*; *tu es* = *t'es*; *il va* = *i va*, etc.) et où selon notre hypothèse ils ont perdu leur statut d'argument. Nous suggérons qu'une évolution croissante pourrait mener à une perte du trait [+accessible] avec peut-être comme seul résidu le trait [+spécifique].

Le débat concernant les clitiques sujet révèle deux statuts différents associés à des situations langagières différentes. Les études favorables au statut d'argument reposent en général sur des jugements de grammaticalité portant sur le français dit standard (e.g. Kayne 1975, Rizzi 1986, Brandi & Cordin 1989, De Cat 2007) alors que les études favorables au statut d'affixe flexionnel portent sur divers 'registres' de la langue orale (e.g. Auger 1995, Roberge 1990, Ashby 1981, Fonseca-Greber & Waugh 2003). Nous avons montré que dans la langue orale spontanée le redoublement clitique est systématique et donc que le clitique lui-même ne peut avoir d'autre statut que celui de marque d'accord. Ce qui importe c'est de reconnaître cette double réalité.

L'écart entre le français oral contemporain et le français dit standard va grandissant comme le montre (ci-dessous) une étude détaillée de la parole adressée à l'enfant comparée à la parole entre adultes. D'une part, la très haute fréquence du redoublement clitique est établie. D'autre part, un examen des trois diagnostics qui ont servi par le passé à étayer l'hypothèse argumentale (à savoir l'usage de *ne*, l'inversion du clitique sujet, et la répétition du clitique sujet dans la coordination) se révèle étayer l'hypothèse opposée. Le corpus de parole adressée à l'enfant est celui de Lyon (Demuth & Tremblay 2008) mis à la disposition du public par CHILDES (MacWhinney 2000). Il comprend environ 106,000 phrases. Sauf mention contraire, les fréquences portent sur le corpus entier.

3.1 Le redoublement dans la parole adressée à l'enfant

Dans la parole adressée à l'enfant, les SN en fonction sujet sont presque toujours redoublés. Sur la base d'un large échantillon du corpus de Lyon nous estimons que la fréquence totale du redoublement du sujet dans les énoncés déclaratifs est de 85% (154/182). Le redoublement n'étant pas permis avec les sujets indéfinis ou quantifiés – ce que notre analyse attribue à l'absence d'appariement de traits – la fréquence du redoublement du sujet [+défini] est même plus élevée: 89% (154/174). Cette proportion n'inclut pas les pronoms forts qui, comme il est noté en section 1.2 ont un taux de redoublement de presque 100% (588/589). En comparaison, le redoublement documenté dans divers corpus de parole entre adultes est beaucoup moins élevé, de 20% à 35%, selon Ashby (1980), Campion (1984) sur la base de divers contextes sociolinguistiques dans des variétés de français qui peuvent inclure un stade moins avancé de français oral contemporain. La fréquence du redoublement chez 23 locuteurs de Paris et Lyon du corpus PFC (Phonologie du français contemporain,

Durand et al. 2002, 2005) est de 60% avec les SN en fonction sujet. Il est à noter que l'âge est un facteur: les locuteurs de moins de 35 ans ont un taux de redoublement de presque 70% alors que le taux est de 50% chez les locuteurs de plus de 35 ans. Cet effet est attendu si l'évolution du statut du clitique sujet en français est en cours comme nous l'avons suggéré plus haut; les locuteurs les plus âgés et les plus jeunes sont présumés avoir une grammaire différente et le traitement des clitiques sujet dépend du registre utilisé. Le taux très élevé de redoublement dans la parole adressée à l'enfant peut justement jouer un rôle critique dans la grammaire acquise par de nouvelles générations de locuteurs. Bien qu'il soit probable que les taux élevés soient dus à un usage excessif de topiques ou SN [+accessible] dans le contexte d'un échange avec de jeunes enfants (alors que ces SN seraient peut-être absents entre adultes), il n'en reste pas moins que les apprenants acquièrent leur grammaire sur la base de l'input qu'ils reçoivent. Si l'input est stable et inambigu parce qu'il traite le redoublement clitique comme obligatoire avec tous les sujets définis, il est inévitable que les générations successives d'apprenants puissent alimenter l'évolution diachronique des clitiques vers de simples affixes flexionnels.

Il reste un point important qui concerne une différence plus fine à faire encore quant au statut des clitiques sujet sur la base du trait de personne. Klausenburger (1999) et Fuß (2005) discutent le rôle pionnier que les pronoms personnels des première et deuxième personnes jouent dans l'évolution vers une simple marque d'accord. C'est de toute évidence le cas en français oral contemporain. Le redoublement des clitiques sujet des première et deuxième personnes est obligatoire dans tous les contextes où il est possible, contrairement aux clitiques sujet de la troisième personne qui sont soumis à la restriction supplémentaire d'appariement des traits vis-à-vis des SN qu'ils redoublent. Pour les pronoms des première et deuxième personnes il suffit pour un apprenant de poser des traits de personne et de nombre. Donc, à la réduction phonologique et une distribution particulière de ces clitiques s'ajoute la perte de traits référentiels, ce qui peut entraîner une évolution générale du système. Ces observations concernant l'asymétrie de personne sont confirmées par le fait qu'elles sont aussi caractéristiques des groupes dialectaux du nord de l'Italie (Poletto 2000, Suñer 1992).

3.2 La rétention de *ne* en français oral contemporain

Un argument en faveur de l'analyse argumentale des clitiques sujet en français en SpecIP repose traditionnellement sur la position du marqueur de portée négative *ne* (Rizzi 1986; Brandi & Cordin 1989).

- (12)a. Jean/il n'aime pas le café.
b. Jean/il ne l'aime pas.

Le fait que *ne* intervienne entre le clitique sujet et le verbe, mais pas entre le clitique objet et le verbe – (12a) versus (12b) – est souvent interprété comme indice de la position structurale plus élevée des clitiques sujet, comparée à celle plus basse des clitiques objet (Rizzi 1986, Brandi & Cordin 1989, Zanuttini 1997). Un phénomène similaire dans certains dialectes du nord de l'Italie (où les SN sujets sont en SpecIP) a mené Poletto (2000), Zanuttini (1997), et Rizzi (1986) à créer une position supplémentaire en I (Infl) pour les clitiques sujet plus haute que celle réservée aux clitiques objet (dans d'autres dialectes du nord de l'Italie, le clitique sujet suit la particule négative, ce qui place le clitique immédiatement à la gauche du verbe. Pour ces dialectes il est proposé que les positions du clitique et de la particule négative sont inversées.) Cette analyse pourrait être appliquée au français oral contemporain mais elle impliquerait que la particule négative ait des propriétés catégorielles identiques aux clitiques sujet et donc que *ne* soit un affixe et non un mot indépendant. Les faits semblent tendre vers une conclusion différente: *ne*, contrairement au clitiques sujet, peut dépendre d'un adverbe comme le montre des exemples tirés de yahoo.fr (*elle dit ne souvent pas être satisfaite*); *ne* est un marqueur de portée; sa position est donc déterminée par la syntaxe, ce qui est difficile à réconcilier avec le statut d'affixe (Zwicky & Pullum 1983). Il faut noter cependant que le statut non-affixal de *ne* n'est un problème pour l'analyse affixale des clitiques sujet que si *ne* apparaît bien entre le clitique sujet et le verbe en français oral contemporain.

Il est clair que *ne* est en perte d'usage (voir Ashby 1976, Coveney 1996, et de nombreux auteurs sur ce point) mais il manque une description détaillée du phénomène. Pour De Cat (2007) l'utilisation de *ne* reste obligatoire dans certaines conditions, en particulier avec les sujets négatifs (*personne, rien, aucun SN, etc.*) et pour marquer la portée négative. Cependant, ces restrictions indicatives d'un *ne* encore actif se révèlent largement absentes en français oral contemporain. Dans la parole adressée à l'enfant (corpus de Lyon), *ne* est omis en présence de sujets négatifs (13a) et *pas* marque la portée négative (13b).

- (13)a. Personne m'en a parlé. (Nat25, line 719)
 b. Tu essaies de pas colorier sur le jaune. (Tim20, line 1439)

Dans la parole entre adultes le taux de rétention de *ne* est très bas, 16% selon Ashby (1981) sur la base d'un corpus de français parlé en région parisienne, 18.8% selon Coveney (1996) pour le français parlé en Picardie, et 1.1% selon Armstrong (2002) pour le français parlé en Lorraine. L'écart entre ces trois résultats est probablement lié à des registres de langue différents. Par exemple, le corpus analysé par Coveney (1996) comprend des interviews avec une personne non-familière (l'auteur) alors que le corpus analysé par Armstrong (2002) est fait de conversations entre des personnes qui se connaissent bien. Enfin, le corpus de Ashby (1981) est relativement plus ancien.

Le taux de rétention de *ne* dans l'ensemble du corpus de Lyon est de 7%, en accord avec les observations précédentes. Mais le taux en lui-même est loin d'être aussi révélateur que la distribution de *ne* elle-même. Comme le montre le tableau 1, le taux de rétention de *ne* dépend en fait du type de sujet présent. La rétention est la plus élevée lorsqu'il s'agit d'un SN sujet (y compris les sujets négatifs), suivi par les sujets nuls (impératif, *y'a existentiel, faut*). La rétention est la moins élevée en présence d'un clitique sujet. Ce qui revient à dire que *ne* est de préférence absent précisément dans les contextes où il interviendrait entre le clitique sujet et le verbe. Ceci est vérifié par une distribution asymétrique pour les sujets redoublés versus non-redoublés. Le taux de rétention est d'environ 83% avec les sujets non-redoublés, comparé à 7% pour les sujets redoublés. Tenant compte du taux de rétention de *ne* avec les sujets nuls (deux fois plus élevé qu'avec un clitique), ces fréquences suggèrent qu'un facteur important dans l'omission de *ne* n'est pas tant la présence d'un SN sujet que le statut des clitics sujet évoluant vers celui d'affixe verbal. Le tableau 2 montre que, par comparaison, les taux de rétention sont différents dans le même corpus lorsque l'adulte parle à partir d'un livre d'enfant au lieu de s'adresser à lui/elle spontanément (même s'il est impossible de savoir si l'adulte lit strictement le texte du livre ou l'adapte dans sa parole adressée à l'enfant). Cela indique une différence de registre mais les tendances générales de rétention de *ne* restent les mêmes.

	ne...		TOTAL
Clitique sujet	323	6%	5092
Sujet nul	104	14%	737
SN	20	83%	24
SN redoublé	9	7%	137

Tableau 1. Rétention de *ne* en parole spontanée

	ne...		TOTAL
Clitique sujet	27	73%	37
Sujet nul	6	100%	6
SN	14	100%	14
SN redoublé	0	--	0

Tableau 2. Rétention de *ne* en parole lue

Comme avec toute évolution en cours, il y a un résidu, en l'occurrence 6% de cas où *ne* intervient entre le clitique et le verbe. Ce taux de résidu nous paraît indiquer que *ne* a, d'une part, presque complètement disparu et que, d'autre part, les locuteurs français disposent encore de plusieurs grammaires correspondant à des registres de langue différents. L'enfant peut donc construire à la fois une grammaire du français oral contemporain (sur la base de l'input auquel il est exposé) et une grammaire du français dit standard correspondant au registre écrit (dont une maîtrise parfaite peut nécessiter une éducation formelle)

3.3 L'inversion en français oral contemporain

Un autre argument fréquent contre une analyse affixale des clitiques sujet provient du fait qu'ils semblent se déplacer en syntaxe comme tout autre mot indépendant (De Cat 2007, Zwicky & Pullum (1983). Parmi les diverses stratégies qu'utilise le français pour exprimer une question il est possible d'inverser le verbe et son clitique sujet. Quoique des analyses existent réconciliant le statut d'affixe des clitiques sujet et l'inversion (voir Poletto 2000 pour les dialectes du nord de l'Italie et Auger 1995 pour le français du Québec) nous montrerons que la stratégie d'inversion est en voie de disparition en français oral contemporain et donc qu'il n'a aucune incompatibilité entre une analyse syntaxique de l'inversion et une analyse affixale des clitiques sujet.

La fréquence de l'inversion dans les corpus de parole entre adultes est extrêmement peu élevée. Dans son corpus de français de Lorraine, le taux d'inversion dans les questions oui-non est de 0% versus 6.6% dans les questions wh selon Coveney (2002). Pour De Cat (2007) le taux est de 0% dans les questions oui-non versus 2% dans les questions wh. Les taux sont encore plus bas dans le corpus de Lyon: 0% dans les questions oui-non (0/2811) versus .9% dans les questions wh (13/1388). De plus, parmi les 13 cas d'inversion dans les questions wh, 8 sont des cas d'inversion complexe comprenant un SN sujet inversé. Au total donc 5/1388 cas d'inversion d'un simple clitique. L'évolution de l'inversion est donc identique à celle de *ne*; les deux montrent que la grammaire du français oral contemporain est sensible au statut affixal, plus précisément préfixal, des clitiques sujet.

3.4 La répétition des clitiques sujet dans la coordination

Pour beaucoup de linguistes (Fonseca-Greber and Waugh 2003, Poletto 2000, Brandi & Cordin 1989, Sportiche 1992, Auger 1995, Rizzi 1986, Kayne 1975) le facteur déterminant de l'analyse des clitiques sujet provient de leur comportement dans la coordination. En contraste avec les clitiques objet en français et les clitiques sujet dans de nombreux dialectes du nord de l'Italie qui doivent être répétés lorsque deux verbes sont coordonnés, la répétition des clitiques sujet est caractérisée comme non-obligatoire en français. Cela semblerait porter un coup fatal à l'analyse affixale car il est difficile d'imaginer qu'un affixe de mot ne soit pas obligatoire (lorsque l'appariement des traits est respecté). Le fait est que dans le corpus de Lyon (parole adressée à l'enfant), la répétition du clitique sujet est quasi-obligatoire, contrairement à l'opinion reçue. Nous n'avons identifié que 3 cas (sur 183) de non-répétition sous coordination de SV, dont l'un concerne un sujet indéfini qui, en vertu de notre analyse, ne permet pas l'accord préfixal. Le taux de répétition sous coordination est donc de 98.9%. L'input auquel sont exposés les apprenants soutient donc l'hypothèse que les clitiques doivent être répétés dans la coordination, conformément à leur statut d'affixe de mot. Là encore, on note une différence entre la parole spontanée (98.9%) et la parole lue (38%, 14/37) adressées à l'enfant. Cette dernière reflète bien sûr le registre écrit de la langue.

4 Conclusions

Le français poursuit son évolution vers une langue à sujet nul. Toutes les indications sont là pour conclure qu'en français oral contemporain les clitiques sujet ont le statut de préfixe marquant l'accord avec le SN en position sujet:

- (i) Le redoublement clitique est un véritable redoublement sans aucune marque prosodique/acoustique qui soutiendrait l'analyse traditionnelle en tant que dislocation du SN sujet à gauche.
- (ii) Le redoublement clitique en français oral contemporain a sa position propre dans la typologie du redoublement dans les langues romanes lorsqu'il est analysé à la suite de Suñer (1988, 1992) en termes d'appariement de traits communs à la marque d'accord (le clitique) et le SN argumental référentiel.

(iii) L'input auquel sont exposés les jeunes français est inambigu et il encourage la convergence vers une grammaire du français où les clitiques sujet ont toutes les propriétés caractéristiques d'affixes flexionnels.

Nous avons insisté sur le fait qu'il est impossible de documenter cet état de fait si l'on se contente d'examiner les intuitions de linguistes qui portent sur le registre écrit ou français dit standard. Il est nécessaire de s'appuyer sur des études quantitatives de corpus appropriés et des analyses expérimentales et par là même, d'enrichir la base de données traditionnelle en Grammaire Générative synchronique.

Références

- Armstrong, N. (2002). Variable deletion of French *ne* : a cross-stylistic perspective. *Language Sciences* 24, 153-173.
- Ashby, W. (1981). The Loss of the Negative Particle *ne* in French : a Syntactic Change in Progress. *Language* 57, 674-687.
- Auger, J. (1995). *Pronominal Clitics in Québec Colloquial French : A Morphological Analysis*. Ph.D. Dissertation, University of Pennsylvania.
- Auger, J. (2003). Le redoublement des sujets en picard. *French Language Studies* (13), 381-406.
- Brandi, L., & Cordin, P. (1989). Two Italian dialects and the null subject parameter. In Jaeggli, O. & Safir, K. (éds.), *The null subject parameter*. Dordrecht: Kluwer, 111-142.
- Coveney, A. (1996). *Variability in Spoken French : a sociolinguistic study of interrogation and negation*. Exeter : Elm Bank Publications.
- Croft, W. (2000). *Explaining Language Change: An Evolutionary Approach*. London: Longman.
- De Cat, C. (2007). *French Dislocation. Interpretation, Syntax, Acquisition*. Oxford: OUP (Oxford Studies in Theoretical Linguistics 17).
- Demuth, K. & Tremblay, A. (2008). Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language* 35, 99-127.
- Deshaies, D., Guibault C., & Paradis, C. (1993). Prosodie et dislocation à gauche par anaphore en français québécois spontané. In Crochetière, A., Boulanger, J.-C., & Ouellon, C., (éds.), Actes du XV^e Congrès international des linguistes vol.II. Québec: P.U.L., 31-34.
- Dietiker, S. (1978). *En bonne forme: Révision de grammaire française*. Heath, Lexington (Mass.).
- Doetjes, J., Delais-Roussarie, E., & Sleeman, P. (2002). The Prosody of Left Detached Constituents in French. *Proceedings of the Speech Prosody 2002 conference, Aix-en-Provence*.
- Durand, Jacques, Bernard Laks & Chantal Lyche (2002). La phonologie du français contemporain: usages, variétés et structure. In: C. Pusch & W. Raible (eds.) *Romanistische Korpuslinguistik- Korpora und gesprochene Sprache/Romance Corpus Linguistics - Corpora and Spoken Language*. Tübingen: Gunter Narr Verlag, pp. 93-106.
- Durand, Jacques, Bernard Laks & Chantal Lyche (2005). Un corpus numérisé pour la phonologie du français. In G. Williams (ed.) *La linguistique de corpus*. Rennes: Presses Universitaires de Rennes. pp. 205-217. Actes du colloque 'La linguistique de corpus', Lorient, 12-14 septembre 2002.
- Fonseca-Greber, B., & Waugh, L. (2003). The subject clitics of Conversational European French. In Núñez-Cedeño, R., López, L., & Cameron, R. (éds.), *A Romance Perspective on Language Knowledge and Use, Selected papers from the 31st Linguistic Symposium on Romance Languages (LSRL), Chicago*.
- Fuß (2005). *The Rise of Agreement. A Formal Approach to the Syntax and Grammaticalization of Verbal Inflection*. Amsterdam: J. Benjamins.
- Kayne, R. (1975). *French Syntax : the Transformational Cycle*. Cambridge, MA : MIT.

- Klausenburger, J. (1999). Grammaticalization. *Studies in Latin and Romance Morphosyntax*. Amsterdam: J. Benjamins.
- Kroch, T. (2000). Syntactic change. In M. Baltin and C. Collins (eds.), *The Handbook of contemporary syntactic theory*. Oxford: Blackwell Publishers, 699–729.
- Lightfoot, D. (1999). *The Development of Language: Acquisition, Change, and Evolution*. New York: Blackwell.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.
- Marandin, Jean-Marie. 2004. Formatage de l'information : focus et contexte. In F. Corblin & C. Gardent (éds.), *Interpréter en contexte*. Paris: Hermes. 31-80.
- Miller, P. & Monachesi, P. (2003). Les pronoms clitiques dans les langues romanes. In Godard, G. (éd) *Les langues romanes : Problèmes de la phase simple*. Paris : CNRS éditions, 67-123.
- Nadasdi, T. (1995). NP doubling, matching, and minority French. *Language Variation and Change* 7, 1-14.
- Poletto, C. (2000). *The Higher Functional Field*. Evidence from Northern Italian Dialects. New York : Oxford.
- Rizzi, L. (1986). Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17, 501– 557.
- Roberge, Y. (1990). *The Syntactic Recoverability of Null Argument*. Montreal : McGill-Queen's University Press.
- Rossi, M. (1999). *L'intonation. Le système du français: description et modélisation*. Paris : Ophrys.
- Samek-Lodovici, V. (2005). Prosody-Syntax Interaction in the Expression of Focus. *Natural Language & Linguistic Theory* 23, 687–755.
- Sportiche, D. (1992). Clitic Constructions. In Zaring, L. & Rooryck, J. (éds), *Phrase Structure and the Lexicon*, Dordrecht: Kluwer Academic Publishers, 213–276.
- Suñer, M. (1988). The Role of Agreement in Clitic-Doubled Constructions. *Natural Language and Linguistic Theory* 6:391-434.
- Suñer, M. (1992). Subject clitics in the Northern Italian vernaculars and the matching hypothesis. *Natural Language and Linguistic Theory* 10, 641-672.
- Zanuttini (1997). *Negation and Clausal Structure. A Comparative Study of Romance Languages*. New York : Oxford.
- Zribi-Hertz, A. (1994). The Syntax of nominative clitics in Standard and Advanced French. In Cinque, G., Koster, J. Polluck, J.-Y., Rizzi, L. and Zanuttini, R. (eds). *Paths Towards Universal Grammar*. Washington, D.C.: Georgetown University Press, 453-472.
- Zwickey, A. & Pullum, G. (1983). Cliticization vs. Inflection: English n't. *Language* 59, 502-513.

¹ Cette étude a été menée dans le cadre d'une bourse de recherche doctorale NSF accordée à Jenny Culbertson et une bourse NSF IGERT accordée au département de sciences cognitives à Johns Hopkins. Nous tenons à remercier Thierry Nazzi, Isabelle Barrière, Bob Frank, Paul Smolensky, et Colin Wilson pour leur conseils et commentaires constructifs tout au cours de sa réalisation.

² Nous laissons de côté les traits supplémentaires des variables Q qui sont nécessaires pour les différencier des syntagmes quantifiés, et donc les séparer dans la hiérarchie en (10). Voir Poletto (2000) et Suñer (1992) pour plus de détails.

³ L'usage du redoublement avec les sujets non-quantifiés est noté, entre autres, dans Côté (2000) et De Cat (2007).

⁴ Voir Auger (1994) pour une discussion du traitement de *ça* en tant qu'affixe flexionnel avec des traits différents.