

## Étude morphosyntaxique des pronoms sujets à partir d'un corpus de maternelle<sup>1</sup>

**Katerina Palasis**

Université Nice Sophia Antipolis, CNRS, BCL, UMR 7320

palasis@unice.fr

**Résumé :** Les syntacticiens débattent depuis des décennies du statut morphosyntaxique des pronoms personnels sujets en français contemporain : ces pronoms sont-ils manipulés en syntaxe comme des morphèmes libres ou entrent-ils dans la morphologie du verbe sous la forme de marqueurs préverbaux de personne ? La seconde analyse implique que le français contemporain est, comme les autres langues romanes, une langue à « sujet nul ». La fréquence, la position, la forme phonologique et les contextes d'utilisation des pronoms sujets dans un corpus de maternelle (trente-sept enfants entre 2;3 et 4;11) apportent un éclairage développemental nouveau à ce débat. La très grande majorité des données indique un traitement morphologique de ces constituants. Néanmoins, l'analyse longitudinale met également en avant l'émergence progressive de plusieurs phénomènes cooccurrents, dont l'utilisation d'un paradigme de pronoms syntaxiques. Ces faits nous invitent à suivre l'hypothèse de la diglossie selon laquelle les locuteurs francophones manipulent deux grammaires, apparentées mais bien distinctes.

**Abstract:** This contribution addresses the long-standing question surrounding the morpho-syntactic status of subject clitics in contemporary French: Do French speakers process these items as free morphemes or agreement markers? The latter implies that contemporary French is a “null-subject” language, on a par with other Romance languages. Frequency, placement, phonological forms and contexts of subject clitics are examined in a kindergarten corpus (thirty-seven children between 2;3 and 4;11), and the results, considered in a developmental perspective, shed new light on the debate. The data indicate that children seem to process subject clitics as preverbal agreement markers. Nevertheless, the longitudinal perspective also highlights a progressive and concomitant emergence of several phenomena, one of which is the emergence of a paradigm of proper syntactic pronouns. These facts support the diglossic hypothesis, which poses that French speakers handle two cognate nevertheless distinct grammars.

La notion de sujet grammatical interpelle les grammairiens et les syntacticiens depuis l'Antiquité. En effet, celle-ci est extrêmement paradoxale. Dans l'histoire de la grammaire, la fonction sujet, à la différence d'autres fonctions telles que celle de complément d'objet, est ancienne et stable, et ce depuis la réflexion sur les parties du discours menée par les philosophes grecs dont Platon. Cette notion appartient également au fonds commun des connaissances de tout écolier dans la mesure où elle est décrite à des fins d'analyse grammaticale dès les premières années à l'école élémentaire. Néanmoins, le consensus et la simplicité ne sont qu'apparents car de nombreux questionnements et points de vue divergents existent également autour de cette notion. La prégnance et la stabilité de la notion contrastent alors avec son aspect complexe comme en témoigne la préconisation de Grevisse et Goosse (2008) que « Le plus sage est sans doute de considérer la notion de sujet comme une espèce de postulat, et de fournir simplement des moyens de l'identifier ».

Comme d'autres chercheurs, nous souhaitons néanmoins continuer à développer la connaissance que la communauté scientifique a déjà de cette notion. Nous proposons donc de contribuer à la réflexion sur la notion de sujet syntaxique et de son absence dans les langues à travers l'étude de la nature des éléments potentiellement sujets en français contemporain oral

---

<sup>1</sup> Recherches effectuées dans le cadre du projet ANR-DFG « Données et analyses dialectologiques, acquisitionnelles et diachroniques des pronoms sujets en Gallo-Roman » (DADDIPRO, 2012-2015, n° ANR 11 FRAL 007 01) sous la direction de Michèle Olivieri (Université Nice Sophia Antipolis, CNRS, BCL, UMR 7320) et Georg Kaiser (Université de Constance, Allemagne).

enfantin. En Section 1, nous exposons la problématique sur la nature du sujet à l’oral et nous l’illustrons de données extraites de corpus adultes. La Section 2 détaille ensuite les données enfantines sur lesquelles nous nous appuyons dans nos travaux. La Section 3 rapporte nos analyses des pronoms clitiques enfantins, en termes de fréquence dans les énoncés, position par rapport au verbe fléchi, contextes syntaxiques et formes phonologiques. L’analyse de ces résultats nous permet alors en Section 4 de proposer une discussion sur le statut morphosyntaxique des pronoms sujets en français oral enfantin. Nous constatons que les clitiques des petits francophones possèdent de nombreuses caractéristiques communes avec les marqueurs d’accord présents dans les dialectes du nord de l’Italie. Nous terminons notre discussion en l’élargissant à plusieurs questions d’ordre plus général : le statut du français enfantin contemporain comparé au français standard, l’hypothèse d’une diglossie en France et les aspects diachroniques de ces deux questions.

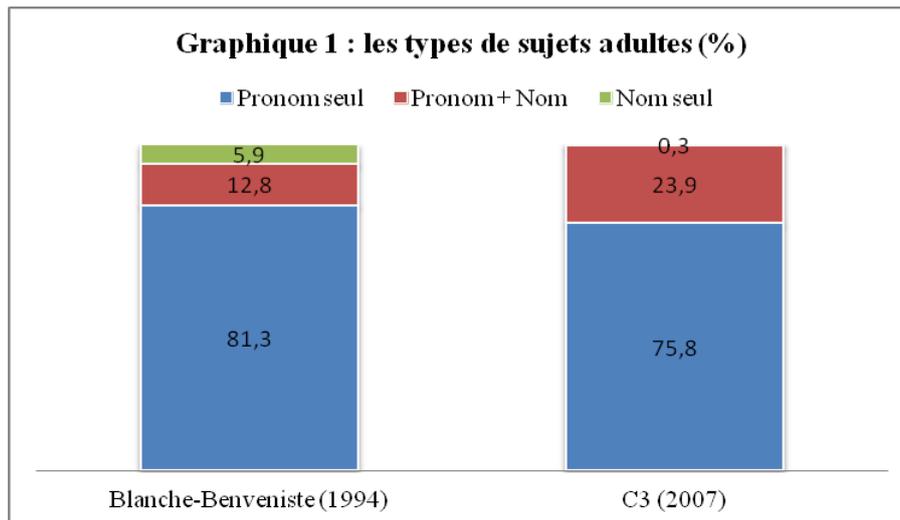
## 1. La problématique

Les constituants traditionnellement décrits comme sujets en français sont, la plupart du temps, soit nominaux soit pronominaux, comme illustré par Grevisse et Goosse (2008) dans des exemples tels que « le professeur écrit au tableau » ou « je travaille ». Or, les corpus oraux montrent que les locuteurs francophones associent souvent ces deux types de constituants au sein d’un même énoncé. Ainsi, Blanche-Benveniste (2000) rapporte l’inquiétude des grammairiens dès le XVII<sup>e</sup> siècle face à ces cooccurrences avec l’exemple du grammairien Chifflet (1659) qui recommande de dire « mon père est malade » au lieu de « mon père il est malade ». Les corpus oraux actuels affichent également des taux de présence élevés des pronoms atones. Ainsi, une des caractéristiques qui distingue très nettement le français oral du français standard est la rareté des syntagmes nominaux seuls à l’oral dans la mesure où les locuteurs insèrent la plupart du temps un pronom atone devant chacun de leurs verbes fléchis, comme le montrent l’exemple (1) et les pourcentages du Graphique 1 ci-après.<sup>2</sup> L’étude de corpus oraux indique par ailleurs que les pronoms atones émergent associés non seulement à des syntagmes définis mais également indéfinis<sup>3</sup>, dans des propositions coordonnées, sans inversion dans les énoncés interrogatifs et que le pronom *il* est éliidé devant consonne (cf. exemples (2) et (3)). Ces caractéristiques sont généralement répertoriées comme distinguant le français oral du français standard (Kayne, 1975, Morin, 1979, Lambrecht, 1981, Ashby, 1984, Zribi-Hertz, 1994, Blanche-Benveniste, 2000, Fonseca-Greber et Waugh, 2003).

- (1) je crois que Lou elle va te montrer où il est le début. (C3, adulte)
- (2) tous les pirates i[z] ont des tatouages. (C3, adulte)
- (3) mais il est triste et i met les pieds dans l’eau ? (C3, adulte)

<sup>2</sup> Le contenu du corpus nommé C3 est détaillé en Section 2.

<sup>3</sup> Sont considérés comme indéfinis : (i) les syntagmes dont le déterminant porte uniquement une valeur quantifiante (par exemple : *un, une, des, tous, chaque, aucun*) et (ii) les pronoms indéfinis (par exemple : *personne, rien*) (cf. Grevisse et Goosse, 2008 : 738).



Les cooccurrences « Pronom + Nom » portent un grand nombre de noms différents dans les grammaires. On trouve ainsi les termes de « reprise », « détachement », « redondance » (Grevisse et Goosse, 2008), « apposition » (Bonnard, 1997), « redoublement » et « dislocation » (Denis et Sancier-Chateau, 1994). La diversité des appellations traduit une diversité des approches et, très schématiquement, trois analyses sont envisagées : (i) le pronom est le sujet syntaxique et l'autre élément est relégué en dehors de la valence verbale (Kayne, 1975, De Cat, 2005) ; (ii) le pronom est un marqueur de personne et le véritable sujet grammatical est l'autre constituant (les pronoms portent alors le nom de « indices pronominaux » (Creissels, 1995) ou de « marqueurs d'accord affixaux » (Culbertson, 2010)) ; (iii) les deux éléments forment ensemble le sujet (Bauche, 1928). Cette absence de consensus provient en partie du statut phonologique des pronoms personnels du français moderne qui sont des pronoms sans accent propre (autrement dit atones ou clitiques), ce qui autorise les analyses mentionnées en (ii) et (iii) ci-dessus. Un tel débat n'existe pas dans les langues comme l'anglais par exemple dans lesquelles les pronoms personnels sont toniques. De plus, l'analyse traditionnelle du français (en (i) ci-dessus) implique que le français est une langue à part dans le groupe des langues romanes dans la mesure où les autres langues de ce groupe, telles que le portugais, l'espagnol ou l'italien, sont des langues dites « à sujet nul », c'est-à-dire des langues dans lesquelles la présence d'un sujet n'est pas obligatoire contrairement au français standard qui est dite « à sujet non-nul » (hors cas particuliers tels que les énoncés à l'impératif). *A contrario*, une analyse telle que celle proposée en (ii) ci-dessus implique de classer le français comme langue « à sujet nul ». De plus, les trois analyses mentionnées ci-dessus renvoient à un débat récurrent aussi bien en grammaire traditionnelle qu'en syntaxe moderne concernant le statut morphosyntaxique du pronom atone préverbal en français : pronom au sens étymologique du terme ou marqueur d'accord (Heap et Roberge, 2001 pour un historique de la question) ? Les partisans actuels des deux camps ont chacun des arguments de valeur (De Cat, 2005 vs Culbertson, 2010 par exemple). Nous proposons donc de reprendre ces arguments et de les mettre à l'épreuve de nouvelles données enfantines.

## 2. Le corpus

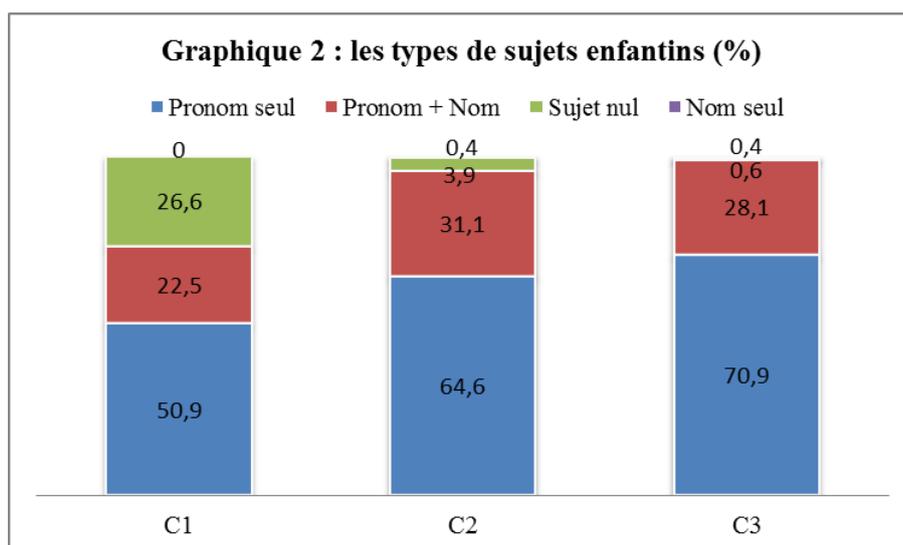
Nos recherches sur l'acquisition du français contemporain reposent sur un ensemble de données recueilli auprès de deux classes d'enfants de maternelle. Le premier groupe se compose de dix-sept enfants âgés de 2;3 à 3;1. Ils ont été enregistrés à l'occasion d'échanges spontanés avec un chercheur pendant cinq séances au cours de leur premier trimestre en toute petite section de maternelle. Ce corpus (dorénavant « C1 ») comporte 1.072 énoncés verbaux (Palasis, 2005). La seconde classe comprend vingt petits élèves. Ces enfants ont été

enregistrés à raison de dix à treize séances par an pendant leurs trois années de maternelle (soit un total de soixante-cinq heures d'enregistrements). Ces données audio et vidéo sont traitées par année scolaire. La première année (dorénavant « C2 ») comprend 15.992 énoncés provenant des vingt enfants alors âgés de 2;5 à 4;0 ainsi que 12.891 énoncés produits par le chercheur qui a conduit les séances (Palasis, 2010a, b).<sup>4</sup> Le traitement des données de la seconde année (dorénavant « C3 ») est quasiment achevé (neuf séances sur dix ont été transcrites et codées). Nous disposons ainsi actuellement de 12.969 énoncés enfantins (dix-neuf enfants entre 3;6 et 4;11) et de 9.042 énoncés adultes. Nous espérons pouvoir traiter les données de la troisième année (« C4 » : 18 enfants entre 4;5 et 5;11) très prochainement. Nos analyses reposent donc sur des données enfantines à la fois transversales (avec un total de trente-sept enfants) et longitudinales (pour le moment, deux années scolaires consécutives), ainsi que sur des données produites par un adulte francophone au cours d'interactions informelles.

### 3. Analyse des sujets enfantins

#### 3.1. Les pronoms atones

L'observation de C1-C3 confirme un certain nombre de faits qui avaient été mis en avant dans d'autres corpus francophones (cf. les travaux de Morgenstern, 2006 sur le développement du sujet et Prévost, 2009 pour une revue de quelques autres corpus). Les jeunes enfants n'associent quasiment jamais les pronoms atones à un verbe non-fléchi par exemple. Nos données permettent également de formuler des observations supplémentaires d'ordre développemental (entre 2;5 et 4;10 pour le moment). Le Graphique 2 illustre les faits suivants. Tout d'abord, nous constatons que les pronoms atones sont présents dans la quasi-totalité des énoncés verbaux enfantins. En effet, si nous ajoutons les catégories « Pronom seul » et « Pronom + Nom » (illustrées en (4)-(7) ci-dessous), nous constatons deux faits : les totaux sont très élevés dans les trois corpus (73,4 %, 95,7 %, 99 %) et ils augmentent avec l'âge des enfants. Cette omniprésence des pronoms atones implique alors que ceux-ci sont répétés lorsque plusieurs verbes fléchis sont juxtaposés et/ou coordonnés dans le même énoncé (cf. (7)).



<sup>4</sup> Ces données (enregistrements audio et/ou vidéo alignés à une transcription orthographique, ainsi qu'un codage morphosyntaxique (%mor) et un codage des occurrences non-cibles (%err) ; Palasis, 2012b) sont disponibles sur le site de CHILDES (MacWhinney, 2000a,b) : <http://childes.talkbank.org/browser/index.php?url=Romance/French/Palasis/>.

- (4) et après il est content de retrouver ses amis. (C3, Carla, 4;7)
- (5) moi j'ai des feutres de Mickey à ma maison. (C3, Massimo, 3;6)
- (6) le nounours il arrive pas à mettre son manteau. (C3, Nina, 3;10)
- (7) elle joue elle danse elle met la musique et elle dort. (C3, Lucille, 3;10)

Nous constatons également que ces pronoms sont quasiment toujours préverbaux, quel que soit le corpus. En effet, C1 et C3 ne comportent aucun pronom postverbal et C2 n'en compte que six (cf. (8)). Les jeunes enfants n'utilisent donc globalement pas l'inversion lorsqu'ils forment des interrogations (cf. (9)). Pourtant, la structure de leurs énoncés interrogatifs se différencie de celle de leurs phrases déclaratives dans la mesure où le syntagme nominal n'apparaît pas au même endroit dans les deux types de structures. Nous illustrons cette différence par la paire d'exemples en (10) et (11) produite par le même enfant au cours de la même séance. Nous constatons ainsi une forte latéralisation du syntagme à gauche dans les déclaratives (91,6 %) et une tendance à la latéralisation du syntagme à droite dans les interrogatives (62,3 %).

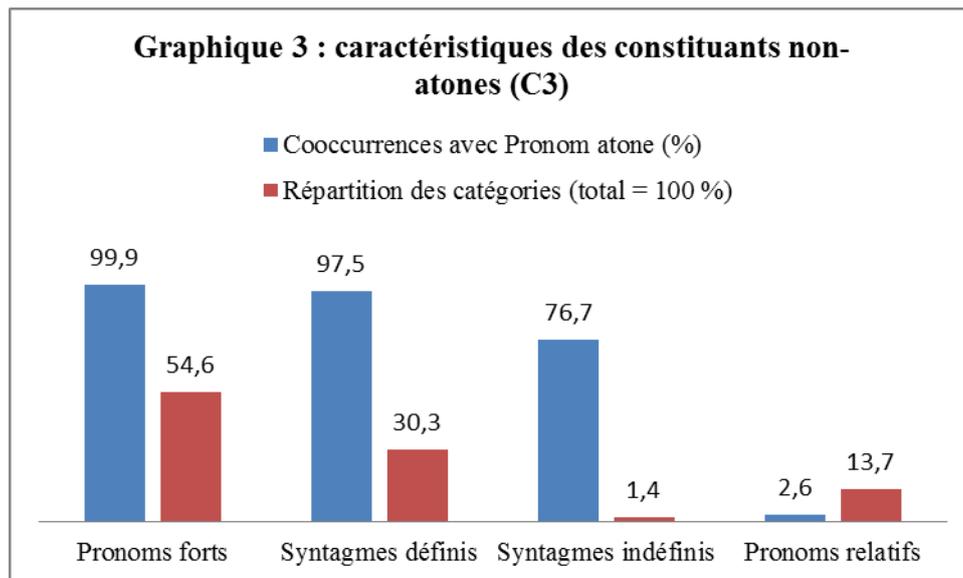
- (8) où es-tu Maman Ours ? (C2, Carla, 3;5)
- (9) où elle est ta maison ? (C3, Matteo, 3;11)
- (10) Sanbarbe il essaie de l'attraper. (C3, Matteo, 4;4)
- (11) et qu'est-ce qu'il fait Sanbarbe là ? (C3, Matteo, 4;4)

La catégorie non-cible « Sujet nul », illustrée en (12), diminue fortement et de façon attendue avec l'âge des enfants (26,6 %, 3,9 %, 0,6 %) (Palasis, 2012a). Finalement, la catégorie « Nom seul », illustrée en (13), est très faiblement représentée chez les jeunes enfants et n'augmente guère pendant la période étudiée (0 %, 0,4 %, 0,4 %). La baisse des Sujets nuls, l'augmentation concomitante des Pronoms et la stabilité basse des Noms semblent alors nous indiquer que ce sont des Pronoms et non des Noms qui prennent petit à petit la place des Sujets nuls dans le discours enfantin. La nature des Sujets nuls enfantins semble donc pronominale, non nominale et le discours enfantin est donc très loin de ressembler au français écrit qui, d'après les données de Blanche-Benveniste (1994), comporte entre 40 % et 80 % de syntagmes nominaux sujets. *A contrario* et en toute logique, ce discours s'apparente bien davantage aux productions adultes orales lorsque ceux-ci s'adressent à des enfants (cf. les corpus mentionnés en Section 1).

- (12) peux chanter ? (C1, Raphaël, 2;6)
- (13) Juliette est en vacances. (C3, Elena, 4;9)

### 3.2. Les autres constituants

Les exemples (5) et (6) ci-dessus illustrent deux types de cooccurrences possibles : le pronom atone émerge soit avec un pronom fort soit avec un syntagme nominal défini. Le Graphique 3 ci-dessous, réalisé à partir des données de C3, nous informe d'une part que, en effet, ces deux catégories émergent très rarement seules (dans 99,9 % et 97,5 % des cas, elles sont accompagnées d'un pronom atone) et, d'autre part, qu'elles couvrent 84,9 % des sujets du corpus (54,6 % + 30,3 %). Les dix-neuf enfants de C3 produisent ces deux types d'énoncés. Nous constatons ainsi que la plupart des cooccurrences est donc grammaticalement définie (*moi, le nounours*, etc.). L'étude des syntagmes nominaux réalisés seuls dans le même corpus (les 0,6 % dans le Graphique 2, soit  $n=35$ , cf. (13) ci-dessus) révèle la même caractéristique dans 71,4 % des cas, donc dans une proportion légèrement moins forte.



A *contrario*, les syntagmes nominaux indéfinis, qu'ils soient seuls ou accompagnés d'un pronom atone (cf. les exemples (14) et (15) ci-dessous), sont donc extrêmement peu présents en fonction sujet (1,4 %). Douze enfants sur les dix-neuf du groupe produisent ces syntagmes et, à la différence des constituants définis dont les taux d'association sont très homogènes d'un enfant à un autre (entre 88 et 100 %), les cooccurrences avec un pronom atone dans cette catégorie présentent des taux extrêmement hétérogènes (2 enfants : 0 %, 2 : 50 %, 1 : 66,7 %, 1 : 100 %). Notons également que l'un des deux enfants qui ne combine jamais un syntagme nominal indéfini avec un pronom atone dans ce corpus est le même et seul enfant à produire un pronom fort seul (cf. exemple (16)). Néanmoins, le Graphique 3 met en avant le fait que, lorsqu'ils émergent à cette place, les syntagmes nominaux indéfinis sont majoritairement accompagnés d'un pronom atone (76,7 %).

(14) des écureuils vit dans les arbres. (C3, Kelsang, 4;8)

(15) et même qu'un Pokemon il a des crocodiles. (C3, Massimo, 3;9)

(16) mais lui n'est pas d'accord. (C3, Kelsang, 4;7)

Terminons ce tour d'horizon par un type de cooccurrence particulièrement intéressant dans la mesure où il est considéré comme non-cible et donc inexistant dans les productions adultes (Labelle, 1988). Il s'agit de l'association d'un pronom relatif et d'un pronom atone, comme illustré en (17) et (18) ci-dessous. Bien qu'absentes du langage adressé aux enfants, ces cooccurrences représentent 2,6 % des pronoms relatifs et concernent quatre enfants dans C3 avec des taux relativement homogènes entre 9,1 % et 20 %. Nous nous sommes également intéressée à la forme du pronom relatif sujet *qui* dans les productions enfantines dans la mesure où la non-élision du [i] devant voyelle est interprétée comme un argument en faveur de l'analyse morphologique de ce [i] en français colloquial du Québec (Auger, 1994). Dix-sept enfants sur dix-neuf produisent le pronom relatif *qui* devant une voyelle dans C3 ( $n=92$ ) et seuls 13 % de ces occurrences sont élidés. Globalement, les enfants n'élident donc pas *qui* devant voyelle et onze enfants ne l'élident jamais ( $n=52$ ), comme illustré en (19).

(17) le petit renard qui il a faim. (C2, Mathilde, 3;2)

(18) ma mère qui il a dit qu'y aura des médailles. (C3, Massimo, 3;9)

(19) ah c'est le tee-shirt qui est trop grand ! (C3, Dylan, 3;8)

### 3.3. La forme du pronom *il*

De nombreux travaux ont montré que la forme phonologique des pronoms atones varie en fonction du contexte en français oral (Morin, 1979, Ashby, 1984). Ainsi, le pronom *il* est présent sous sa forme pleine devant voyelle et est élide devant consonne. D'après Pope (1934), cette réduction phonologique est attestée en français populaire depuis le XII<sup>e</sup> siècle et Chervel (2011) nous informe que l'école primaire et la pression de la norme ont restauré la prononciation du [i] devant consonne au XIX<sup>e</sup> siècle. Cette alternance dans les formes du pronom *il* est donc bien connue et attestée, que ce soit dans les corpus adultes ou enfantins (Grégoire, 1947, P. Guillaume, 1927). C2 et C3 ne dérogent pas à la règle puisque 99,4 % et 99,3 % des pronoms *il* dans ces deux corpus (totaux respectifs :  $n=1688$  et  $n=2208$ ) respectent l'alternance *i/il*, comme illustré en (20) et (21) ci-dessous.

(20) i veut sa maman parc'que il est tout seul. (C2, Nina, 3;3)

(21) et après i sort et après il habite ici. (C3, Dylan, 3;8)

L'élément nouveau que nous souhaitons mettre en avant est le contexte dans lequel l'alternance ne s'applique pas dans C2 (0,6 %,  $n=11$ ) et C3 (0,7 %,  $n=15$ ). Les données de C2 sont particulièrement tranchées à cet égard puisque dix énoncés sur onze présentent un contexte récurrent : la présence de la particule préverbale négative *ne*, comme illustré en (22). Notons que l'utilisation de la négation discontinue est très rare chez les enfants de cet âge puisqu'elle n'est présente que dans 1,2 % des énoncés verbaux négatifs de C2 (soit  $n=16$ ). Dans C2, nous observons alors une corrélation systématique entre le type de négation (simple vs discontinue) et la forme de *il* (élide vs pleine) : C2 ne comporte ainsi aucun énoncé dans lequel la négation simple et la forme pleine de *il* ou la négation discontinue et la forme élide de *il* sont cooccurrents. Tous les énoncés négatifs respectent donc les associations suivantes : négation discontinue et forme pleine (cf. (22)) ou négation simple et forme élide (cf. (23)) (Palasis, 2013).

(22) il ne chante plus. (C2, Mathilde, 3;2)

(23) i nous mangera pas. (C2, Lucille, 3;1)

Qu'en est-il dans C3 ? Quinze énoncés sur 2208 (0,7 %) comportent une forme non-élide de *il* devant consonne dans ce corpus. L'étude des contextes dans ces énoncés révèle que sept d'entre eux sont négatifs et que la négation y est systématiquement représentée par le morphème discontinu *ne...pas*, comme dans C2. La corrélation décrite d'après les données de C2 entre le type de négation et la forme du pronom *il* existe donc également dans C3, comme illustré par la paire d'énoncés en (24) et (25) ci-dessous, produite par la même enfant, au cours de la même séance, avec le même verbe. La forme non-élide de *il* devant consonne peut donc être expliquée par la présence du morphème négatif *ne* dans 99,94 % des cas dans C2 et 99,64 % des cas dans C3. Nous reviendrons sur les 0,06 % ( $n=1$ ) et 0,36 % ( $n=8$ ) restants dans notre discussion.

(24) mais lui il ne voulait toujours pas. (C3, Lou, 4;9)

(25) mais i voulait pas que j' le nourrisse. (C3, Lou, 4;9)

## 4. Discussion

### 4.1. Sur le statut morphosyntaxique des clitiques

L'observation de nos données enfantines permet de confirmer et d'affiner les descriptions existantes sur les sujets oraux enfantins. Tout d'abord, nous avons noté que, non seulement les pronoms atones sont majoritairement présents dans les propositions indépendantes et principales, mais que leur présence augmente avec l'âge dans ces contextes (de 73,4 % dans C1 à 99 % dans C3). La proposition subordonnée relative représente alors le seul contexte

dans lequel le clitique est majoritairement absent (97,4 %). Cette dernière observation est intéressante à deux titres.

Tout d'abord, le taux d'absence très élevé n'est pas surprenant dans la mesure où le type de cooccurrence « Pronom relatif sujet + Clitique » n'existe pas dans les productions adultes (Labelle, 1988). Nous faisons alors l'hypothèse que ces cooccurrences chez les quatre enfants qui les produisent (total  $n=11$ ) ne sont pas le résultat d'une répétition mais d'une computation morphosyntaxique<sup>5</sup> au cours de laquelle l'enfant considère le clitique comme un marqueur morphologique, ce qui lui permet de l'associer à un pronom relatif sujet. Cette proposition nous renvoie alors à l'analyse de type (ii) mentionnée dans la Section 1 (Creissels, 1995, Culbertson, 2010). Par extension, nous pouvons alors supposer que ces enfants utilisent tous leurs clitiques comme des marqueurs de personne, non seulement dans ce contexte avec un pronom relatif mais également dans les trois autres contextes étudiés (syntagmes nominaux indéfinis, syntagmes nominaux définis et pronoms toniques). Nous nous retrouvons alors dans une situation qui existe par ailleurs dans d'autres langues. En effet, les dialectes du nord de l'Italie sont bien connus dans la littérature pour deux raisons : ces dialectes doublent leurs pronoms toniques, leurs syntagmes nominaux définis et indéfinis et leurs pronoms relatifs avec un clitique et ce clitique est analysé de façon consensuelle comme un marqueur d'accord (Rizzi, 1986, Brandi et Cordin, 1989, Poletto, 2000), comme illustré en (26) ci-dessous. Malgré la présence constante du clitique, ces dialectes sont donc considérés comme des langues à sujet nul, au même titre que l'Italien standard par exemple. Néanmoins, le phénomène du doublement du sujet n'est pas un phénomène homogène dans cette aire dialectale. Poletto (1993, 2000) notamment a étudié cette variation. L'auteure a ainsi non seulement proposé un classement des dialectes en fonction du nombre de contextes d'application du doublement du sujet (de un à quatre) mais a également établi une liste d'implications entre les quatre contextes (cf. (27)-(29) ci-dessous).

(26) El Gianni el magna. « Jean il mange » (dialecte trentin, extrait de Rizzi, 1986)

(27) Si les syntagmes nominaux définis sont doublés par un clitique dans un dialecte, alors tous les pronoms toniques le sont également dans cette même langue (en trentin par exemple).

(28) Si les syntagmes nominaux indéfinis sont doublés par un clitique, alors tous les syntagmes nominaux définis et tous les pronoms toniques le sont également (en milanais par exemple).

(29) Si les variables dans les contextes *wh*- (c'est-à-dire les structures relatives, interrogatives et clivées) sont doublées par un clitique, alors le doublement est obligatoire avec tous les autres types de sujets (en ligurien par exemple).

Il est alors extrêmement intéressant de retrouver le même type d'implications dans nos corpus de français enfantin. En effet, dans C2, les trois enfants qui doublent les pronoms relatifs doublent également les syntagmes indéfinis, définis et les pronoms toniques. Dans C3, deux des quatre enfants qui doublent les pronoms relatifs doublent également les trois autres types de sujets (les deux autres enfants ne produisant aucun syntagme indéfini, l'implication ne peut être vérifiée à ce niveau). Ainsi, la présence quasi-obligatoire de ces clitiques en position préverbale uniquement et les similitudes avec le doublement dans les dialectes du nord de l'Italie sont autant d'arguments qui nous conduisent à proposer une analyse de ces clitiques en tant que marqueurs de personne dans les corpus enfantins étudiés. Nous proposons ainsi que dans un énoncé tel que *et après i sait* (Elena, 4;4) : (i) le clitique est utilisé comme un marqueur préverbal de personne ; (ii) il n'y a pas de sujet grammatical phonétiquement réalisé ; (iii) le français oral fait partie du même groupe que les autres langues romanes dites « à sujet nul » dans la mesure où la réalisation d'un sujet grammatical n'est pas obligatoire.

<sup>5</sup> Nous suivons le type de raisonnement appliqué en psycholinguistique depuis les travaux fondateurs de Berko (1958) sur l'acquisition de la morphologie.

Néanmoins, nous avons également vu que certaines occurrences représenteraient alors des contrexemples à cette analyse morphologique. *Quid* par exemple des énoncés dans lesquels le sujet n'est pas doublé par un clitique (C1 : 0 %, C2 et C3 : 0,4 %, soit  $n=39$  et  $n=35$ ) ? La forme du clitique *il* lorsqu'il n'est pas élidé devant consonne nous semble alors apporter un éclairage supplémentaire à la discussion sur le statut morphosyntaxique des clitics en français et ouvre notre champ d'horizon à une discussion plus générale sur le statut du français étudié dans ces corpus enfants.

#### 4.2. Sur le statut du français enfantin contemporain

L'hypothèse morphologique défendue ici fait donc les prédictions que le clitique est : obligatoire, préverbal et élidé devant consonne. Les énoncés avec un sujet non doublé par un clitique ou avec un clitique postverbal ou non-élidé devant consonne sont donc incompatibles avec cette analyse. Nous pouvons néanmoins rendre compte de leur existence et de leur fonctionnement en suivant les propositions de Massot (2008, 2010), Barra-Jover (2010, 2013), Zribi-Hertz (2011) et Massot et Rowlett (2013) sur l'hypothèse d'une diglossie en France (d'après les travaux de Ferguson, 1959). En effet, si, comme ces auteurs, nous postulons l'existence de deux grammaires du français, c'est-à-dire une grammaire acquise spontanément dès le plus jeune âge et généralement utilisée dans des contextes plutôt informels (appelons-la  $G_1$ ) et une seconde grammaire apprise plus tard, souvent par le biais de l'école et utilisée dans des contextes plus formels ( $G_2$ ), alors nous pouvons rendre compte de la variation entre les deux types de comportements du clitique *il* en postulant l'existence de deux paradigmes distincts de clitics. Les clitics de  $G_1$  seraient des marqueurs de personne (obligatoires, préverbaux, élidés devant consonne) et les clitics de  $G_2$  seraient des pronoms syntaxiques (pronoms au sens étymologique du terme, pré- ou postverbaux, de forme fixe quel que soit le contexte).

Les données sur la négation dans C2 et C3 étayaient cette hypothèse dans la mesure où elles montrent que les distributions de l'éllision et de la négation ne sont pas aléatoires mais que, au contraire, elles corréllent dans les énoncés, comme illustré avec les exemples (24) et (25), repris en (30) et (31) ci-dessous. La négation discontinue (en (31)), les clitics postverbaux ((8), repris en (32)), les clitics non-élidés devant consonne (en (31) et (33)), les pronoms forts et les syntagmes nominaux non-doublés (en (34) et (35)) relèvent alors de  $G_2$ , les enfants utilisant les deux grammaires de façon différenciée mais inconsciente (un des principes de la diglossie). Le faible taux de représentation de  $G_2$  dans nos corpus est cohérent avec la chronologie des acquisitions dans une situation diglossique et nous nous attendons à une augmentation des données de  $G_2$  dans notre futur C4, en termes de proportion des énoncés par rapport à  $G_1$  et du nombre d'enfants qui pratiquent  $G_2$  (pour la négation discontinue par exemple, les 1,2 % des énoncés verbaux négatifs sont répartis sur quatre enfants dans C2 et les 1,8 % des énoncés verbaux négatifs sont répartis sur dix enfants dans C3).

(30) mais i voulait pas que j' le nourrisse. (C3, Lou, 4;9) →  $G_1$

(31) mais lui il ne voulait toujours pas. (C3, Lou, 4;9) →  $G_2$

(32) où es-tu Maman Ours ? (C2, Carla, 3;5) →  $G_2$

(33) il va mettre les gens en prison (C3, Eliza, 4;5) →  $G_2$

(34) mais lui n'est pas d'accord (C3, Kelsang, 4;7) →  $G_2$

(35) mais des fois Lisa va à la cantine. (C3, Matteo, 4;2) →  $G_2$

#### 4.3. Sur l'aspect diachronique de la diglossie

La remarque de Hirschbühler et Labelle (2004), traduite ci-dessous, sur les données enfantines du futur roi Louis XIII (Foisil, 1989) nous incite à élargir notre champ

d'investigations à ce corpus enfantin du XVII<sup>e</sup> siècle, décrit comme très précieux pour l'étude du français (Morgenstern, 2009).

Louis XIII a deux grammaires en compétition, une avec *ne* et une sans *ne*. Si ceci est correct, l'omission de *ne* à l'oral, qui est également une caractéristique du français oral contemporain, remonte aussi loin que le début du XVII<sup>e</sup> siècle. (Hirschbühler et Labelle, 2004 : 196)

Nos investigations, actuellement en cours, portent pour le moment sur les énoncés verbaux négatifs de Louis entre 2;0 et 3;3 ( $n=67$ ). Elles mettent en avant les résultats provisoires suivants. Tout d'abord, Louis, comme nos petits locuteurs contemporains, associe un clitique à quasiment chacun de ses verbes fléchis dans ce petit corpus (94 % des énoncés).<sup>6</sup> Les quatre énoncés négatifs qui ne comportent pas de clitiques sont deux énoncés à sujet nul (exemple en (36)) et deux énoncés avec un syntagme nominal sujet seul (exemple en (37)). Nous constatons alors, dans un second temps, que les deux énoncés avec un syntagme nominal seul comportent la forme discontinue de la négation. Ces deux énoncés enfantins du XVII<sup>e</sup> siècle combinent ainsi deux caractéristiques majeures qui ont été corrélées d'après l'observation de données du XXI<sup>e</sup> siècle. Selon notre hypothèse, ces deux énoncés présentent alors deux caractéristiques cohérentes pour être classés comme relevant de G<sub>2</sub>. Cette observation intéressante mais uniquement liminaire sera complétée prochainement par l'étude de l'alternance *i/il* dans ces mêmes énoncés puis par une observation étendue à l'ensemble des énoncés avant d'élargir l'investigation à une tranche d'âge plus importante.

(36) fai pa beau saboté. « on n'est pas beau en sabots » (Louis, 2;6)

(37) que le Roy ne veu pa descende.<sup>7</sup> (Louis, 2;11)

#### 4.4. Sur l'aspect diachronique du statut morphosyntaxique des clitiques

Nous avons mentionné en Section 1 que l'une des raisons pour lesquelles le débat sur le statut morphosyntaxique du pronom existe en français et non dans d'autres langues telles que l'anglais est son statut phonologique faible. En français, le pronom personnel est atone. Il ne peut donc pas être utilisé seul (en réponse à une question par exemple), ni être coordonné avec un autre pronom clitique, ni être séparé du verbe fléchi par autre chose qu'un autre clitique (Kayne, 1975), contrairement aux pronoms de l'ancien français (exemples extraits de Kaiser, 2009). En français moderne, les pronoms toniques sont utilisés dans ces trois contextes, comme illustré dans le Tableau 1.

	Ancien français	Français moderne
Isolation	jo	*je / moi
Coordination	tu é li	*tu et il / toi et lui
Séparation	je meismes preu i avrai <sup>8</sup>	*il, souvent, mange / lui

Tableau 1 : évolution du statut des pronoms personnels en français

Nous avons donc dans un premier temps, entre le latin et l'ancien français, l'émergence d'un paradigme de pronoms préverbaux et indépendants du verbe. Le processus diachronique évoqué, lorsqu'un lien est postulé avec l'affaiblissement du caractère distinctif des suffixes verbaux depuis le latin, est la déflexivité, c'est-à-dire le processus par lequel un morphème flexionnel acquiert le statut de morphème libre, généralement par antéposition (cf. par exemple Moignet (1965), à la suite des travaux de G. Guillaume (2003 [1950-1959])). Puis, dans un deuxième temps, à partir du Moyen français, ces pronoms perdent de leur

<sup>6</sup> Dufter et Stark (2007) ont noté la même proportion de clitiques dans les propositions principales négatives de Louis entre 4;0 et 5;0.

<sup>7</sup> En réponse à la question : « Que dit-on ? ».

<sup>8</sup> « je moi-même en tirerai un avantage ».

indépendance et deviennent clitiques. Ce processus d'affaiblissement est fréquemment qualifié de grammaticalisation (Hopper et Traugott, 1993) et est généralement distingué de la déflexivité (Bottineau, 2010). C'est à cette période qu'émergent les pronoms toniques et l'utilisation concomitante des deux séries de pronoms avec un même verbe fléchi.

La question qui était posée dans le cadre de cette étude était de savoir si les clitiques du français ont franchi la dernière étape de la grammaticalisation, celle qui consiste à passer du statut de clitique phonologique à celui d'affixe flexionnel. Nous avons mentionné que la question était largement débattue chez les syntacticiens et que les deux réponses possibles reposaient chacune sur des arguments de valeurs (De Cat, 2005 vs Culbertson, 2010). Nous suggérons que les partisans des deux camps ont finalement raison dans la mesure où ils décrivent chacun une langue différente : le français oral (G<sub>1</sub>) dans lequel les clitiques sont manipulés comme des marqueurs préverbaux de personne et le français standard (G<sub>2</sub>) dans lequel les clitiques sont manipulés en syntaxe au même titre que les syntagmes nominaux.

## Bibliographie

- Ashby William J., 1984, The elision of /l/ in French clitic pronouns and articles, in Ernst Pulgram (éd.), *Romanitas: Studies in Romance Linguistics*, Ann Arbor, University of Michigan Press, p. 1-16.
- Auger Julie, 1994, *Pronominal Clitics in Québec Colloquial French: A Morphological Analysis*, PhD Dissertation, University of Pennsylvania.
- Barra-Jover Mario, 2010, « Le » français ou ce qui arrive lorsqu'un état de choses est observé comme une entité, *Langue française* N° 168, p. 3-18.
- Barra-Jover Mario, 2013, Linguistique et école primaire : en quoi l'approche diglossique est-elle la meilleure façon d'apprendre le français académique ?, *Journal of French Language Studies* N° 23, p. 87-108.
- Bauche Henri, 1928, *Le langage populaire*, Paris, Éditions Payot.
- Berko Jean, 1958, The Child's Learning of English Morphology, *Word* N° 14, p. 150-177.
- Blanche-Benveniste Claire, 1994, Quelques caractéristiques grammaticales des « sujets » employés dans le français parlé des conversations, in Marina Yaguello (éd.), *Subjecthood and Subjectivity. Proceedings of the Colloquium "The Status of the Subject in Linguistic Theory"*, London, Éditions Ophrys, p. 77-107.
- Blanche-Benveniste Claire, 2000, *Approches de la langue parlée en français*, Paris, Éditions Ophrys.
- Bonnard Henri, 1997, *Grammaire française à l'usage de tous*, Paris, Éditions Magnard.
- Bottineau Didier, 2010, Introduction, *Langages* N° 178, p. 3-10.
- Brandi Luciana et Patrizia Cordin, 1989, Two Italian Dialects and the Null Subject Parameter, in Osvaldo Jaeggli et Kenneth Safir (éds.), *The Null Subject Parameter*, Dordrecht, Kluwer Academic Publishers, p. 111-142.
- Chervel André, 2011, Histoire de l'orthographe française : réformes, enseignement, niveau, in Olivier Bertrand et Isabelle Schaffner (éds.), *Variétés, variations & formes du français*, Palaiseau, Les Éditions de l'École Polytechnique, p. 195-210.
- Chifflet Laurent, 1659, *Essay d'une parfaite grammaire de la langue française*, Anvers, Jacques van Meurs.
- Creissels Denis, 1995, *Éléments de syntaxe générale*, Paris, Presses Universitaires de France.
- Culbertson Jennifer, 2010, Convergent Evidence for Categorical Change in French: From Subject Clitic to Agreement Marker, *Language* N° 86, p. 85-132.
- De Cat Cécile, 2005, French Subject Clitics are not Agreement Markers, *Lingua* N° 115, p. 1195-1219.
- Denis Delphine et Anne Sancier-Chateau, 1994, *Grammaire du français*, Paris, Le Livre de Poche.
- Dufter Andreas et Elisabeth Stark, 2007, La linguistique variationnelle et les changements linguistiques 'mal compris' : le cas du *ne* de négation, in Bernard Combettes et Christiane Marchello-Nizia (éds.), *Études sur le changement linguistique en français*, Nancy, Presses Universitaires de Nancy, p. 115-128.
- Ferguson Charles A., 1959, Diglossia, *Word* N° 15, p. 325-340.
- Foisil Madeleine (éd.), 1989, *Journal de Jean Héroard Médecin de Louis XIII*, Paris, Fayard.
- Fonseca-Greber Bonnie et Linda R. Waugh, 2003, The Subject Clitics of Conversational European French: Morphologization, Grammatical Change, Semantic Change, and Change in Progress, in Rafael Núñez-Cedeño, Luis López et Richard Cameron (éds.), *A Romance Perspective on Language Knowledge and Use*, Amsterdam, John Benjamins, p. 99-117.
- Grégoire Antoine, 1947, *L'apprentissage du langage : la troisième année et les années suivantes*, Paris & Liège, E. Droz & Faculté de Philosophie et Lettres.
- Grevisse Maurice et André Goosse, 2008, *Le bon usage*, Bruxelles, De Boeck & Duculot (14ème édition).
- Guillaume Gustave, 2003 [1950-1959], *Prolégomènes à la linguistique structurale I*, Québec, Les Presses de l'Université Laval.
- Guillaume Paul, 1927, Le développement des éléments formels dans le langage de l'enfant, *Journal de Psychologie* N° 24, p. 203-229.

- Heap David et Yves Roberge, 2001, Cliticisation et théorie syntaxique, 1971-2001, *Revue québécoise de linguistique* N° 30, p. 63-90.
- Hirschbühler Paul et Marie Labelle, 2004, Enclisis and proclisis of object pronouns at the turn of the 17th century: The speech of the future Louis XIIIth, in Randall S. Gess et Deborah Arteaga (éds.), *Historical Romance Linguistics: Retrospective and perspectives*, Amsterdam, John Benjamins, p. 187-208.
- Hopper Paul J. et Elizabeth Closs Traugott, 1993, *Grammaticalization*, Cambridge, Cambridge University Press.
- Kaiser Georg A., 2009, Losing the Null Subject. A Contrastive Study of (Brazilian) Portuguese and (Medieval) French, in Georg A. Kaiser et Eva-Maria Remberger (éds.), *Null-subjects, Expletives, and Locatives in Romance*, Konstanz, Konstanz Working Papers of the Department of Linguistics, p. 131-156.
- Kayne Richard S., 1975, *French Syntax. The Transformational Cycle*, Cambridge, MA, The MIT Press.
- Labelle Marie, 1988, *Prédication et mouvement : le développement de la construction relative chez les enfants francophones*, Thèse de Doctorat, Université d'Ottawa.
- Lambrecht Knud, 1981, *Topic, Antitopic and Verb Agreement in Non-standard French*, Amsterdam, John Benjamins.
- MacWhinney Brian, 2000a, *The CHILDES Project. Tools for Analyzing Talk: The Database*, Mahwah, NJ, Lawrence Erlbaum Associates, <http://childes.talkbank.org/>.
- MacWhinney Brian, 2000b, *The CHILDES Project. Tools for Analyzing Talk: Transcription Format and Programs*, Mahwah, NJ, Lawrence Erlbaum Associates, <http://childes.talkbank.org/>.
- Massot Benjamin, 2008, *Français et diglossie. Décrire la situation linguistique française contemporaine comme une diglossie : arguments morphosyntaxiques*, Thèse, Université Paris 8.
- Massot Benjamin, 2010, Le patron diglossique de variation grammaticale en français, *Langue française* N° 168, p. 87-106.
- Massot Benjamin et Paul Rowlett (éds.), 2013, L'hypothèse d'une diglossie en France, *Journal of French Language Studies* N° 23.
- Moignet Gérard, 1965, *Le pronom personnel français. Essai de psycho-systématique historique*, Paris, Klincksieck.
- Morgenstern Aliyah, 2006, *Un « je » en construction : genèse de l'auto-désignation chez le jeune enfant*, Paris, Éditions Ophrys.
- Morgenstern Aliyah, 2009, *L'enfant dans la langue*, Paris, Presses Sorbonne Nouvelle.
- Morin Yves-Charles, 1979, La morphophonologie des pronoms clitiques en français populaire, *Cahier de linguistique* N° 9, p. 1-36.
- Palasis Katerina, 2005, *Problèmes d'acquisition et le Paramètre du Sujet Nul*, Mémoire de Master, Université Nice Sophia Antipolis.
- Palasis Katerina, 2010a, Introducing New French Child Data: Thoughts on their Gathering and Coding, *Corpus* N° 9 *La syntaxe de corpus*, p. 33-51, <http://corpus.revues.org/index1801.html>.
- Palasis Katerina, 2010b, *Syntaxe générative et acquisition : le sujet dans le développement du système linguistique du jeune enfant*, Villeneuve d'Ascq, ANRT Diffusion.
- Palasis Katerina, 2012a, Economy Principles within Child Speech: When the Nominative Clitic does not Surface, in Sandrine Ferré, Philippe Prévost, Laurice Tuller et Rasha Zebib (éds.), *Selected Proceedings of the Romance Turn IV Workshop on the Acquisition of Romance Languages*, Newcastle, Cambridge Scholars Publishing, p. 261-283.
- Palasis Katerina, 2012b, *Romance French Child Corpus 1 (2;5-4;0)*, CHILDES, <http://childes.talkbank.org/browser/index.php?url=Romance/French/Palasis/>.
- Palasis Katerina, 2013, The case for diglossia: Describing the emergence of two grammars in the early acquisition of Metropolitan French, *Journal of French Language Studies* N° 23, p. 17-35.
- Platon, [1920-1964], *Le Sophiste, Oeuvres complètes*, Paris, Les Belles Lettres ; traduction de Diès.
- Poletto Cecilia, 1993, *La sintassi del soggetto nei dialetti italiani settentrionali*, Padova, Unipress.
- Poletto Cecilia, 2000, *The Higher Functional Field: Evidence from Northern Italian Dialects*, New York & Oxford, Oxford University Press.
- Pope Mildred K., 1934, *From Latin to Modern French*, Manchester, Manchester University Press.
- Prévost Philippe, 2009, *The Acquisition of French: the Development of Inflectional Morphology and Syntax in L1 Acquisition, Bilingualism, and L2 Acquisition*, Amsterdam, John Benjamins.
- Rizzi Luigi, 1986, On the Status of Subject Clitics in Romance, in Osvaldo Jaeggli et Carmen Silva-Corvalan (éds.), *Studies in Romance Linguistics*, Dordrecht, Foris, p. 391-419.
- Zribi-Hertz Anne, 1994, La syntaxe des clitiques nominatifs en français standard et en français avancé, *Travaux de Linguistique et de Philologie* N° 32, p. 131-148.
- Zribi-Hertz Anne, 2011, Pour un modèle diglossique de description du français : quelques implications théoriques, didactiques et méthodologiques, *Journal of French Language Studies* N° 21, p. 231-256.