



HAL
open science

**Le flux de l'information, aspects syntaxiques et
discursifs : une étude fonctionnaliste et
développementale**

Audrey Mazur-Palandre

► **To cite this version:**

Audrey Mazur-Palandre. Le flux de l'information, aspects syntaxiques et discursifs : une étude fonctionnaliste et développementale. Linguistique. Université Lumière - Lyon II, 2009. Français. NNT : 2009LYO20053 . tel-01543943

HAL Id: tel-01543943

<https://theses.hal.science/tel-01543943>

Submitted on 21 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Lumière Lyon 2

École doctorale : Lettres, langues, linguistique, arts

Équipe de recherche : *Dynamique Du Langage*

**Le flux de l'information,
aspects syntaxiques et
discursifs**

Par Audrey MAZUR-PALANDRE

*Une étude fonctionnaliste et
développementale*

Thèse de doctorat en Sciences du langage

Sous la direction d'Harriet JISA et Michel FAYOL

Présentée et soutenue publiquement le 18/09/2009

Composition du jury :

Harriet JISA, Professeur des universités, Université Lyon 2

Michel FAYOL, Professeur des universités, Université Clermont-Ferrand 2

Jean-Pierre CHEVROT, Professeur des universités, Université Grenoble 3

Judy REILLY, Professeur d'université, University of California

Michel CHAROLLES, Professeur des université, Université Paris 3

Contrat de diffusion

Ce document est diffusé sous le contrat *Creative Commons* « [Paternité – pas d'utilisation commerciale - pas de modification](#) » : vous êtes libre de le reproduire, de le distribuer et de le communiquer au public à condition d'en mentionner le nom de l'auteur et de ne pas le modifier, le transformer, l'adapter ni l'utiliser à des fins commerciales.

À ma famille,

À mon mari,

Remerciements

Je tiens tout d'abord à exprimer ma gratitude envers mes deux directeurs de thèse, Harriet Jisa et Michel Fayol.

Merci au Professeur Harriet Jisa pour m'avoir fait découvrir le monde de l'acquisition et du développement du langage. Voilà maintenant six ans qu'elle a accepté de m'encadrer, d'abord pour un mémoire de maîtrise et ensuite pour une thèse de doctorat, six années de formation durant lesquels elle m'a toujours soutenue.

Merci de m'avoir donné envie de faire de la psycholinguistique,

Merci de m'avoir fait connaître le monde scientifique,

Merci de m'avoir formée,

Merci d'avoir eu confiance en moi,

Merci pour votre humanité et votre bienveillance.

Merci au Professeur Michel Fayol d'avoir accepté de codiriger mon travail de recherche,

Merci pour l'expertise que vous m'avez accordée,

Merci d'avoir eu confiance en moi,

Merci pour votre soutien,

Merci pour votre gentillesse.

Je remercie Michel Charolles ainsi que Jean-Pierre Chevrot et Judy Reilly, mes rapporteurs, d'avoir accepté de faire partie de mon jury,

Merci d'avoir porté un intérêt à mon travail de recherche,

Merci de l'avoir jugé,

Merci de votre disponibilité durant cette période de vacances.

Cette étude n'aurait pu se faire sans l'équipe pédagogique de trois écoles primaires et d'un collège de l'est lyonnais,

Merci pour leur accueil et leur confiance,

Merci à tous les enfants volontaires d'avoir pris les expérimentations si au sérieux,

Un grand merci tout particulièrement au Principal et à la CPE du collège.

Je suis reconnaissante à tous les membres du Laboratoire Dynamique Du Langage, et en son nom à François Pellegrino, directeur du laboratoire, qui met tout en œuvre pour que chaque membre du laboratoire trouve sa place et réussisse sa mission.

Je remercie les DDLers pour leur disponibilité : *Nathalie Bedoin*, *Mahé Ben Hamed* et *Fanny Meunier* pour avoir répondu à mes questions de statistiques ; *Linda Brendlin* pour m'avoir facilité la vie pour toutes les questions administratives ; *Florence Chenu* pour tous ses conseils avisés et ses relectures ; *Emmanuel Ferragne* pour les diverses discussions ; *Christian Fressard* pour tous les coups de pouce informatiques et *les séances Psy gratuites* ; *Frédérique Gayraud* pour ses conseils, relectures, et pour toutes ses discussions concernant le mouvement de grève des universitaires ☺ ; *Sophie Kern* pour sa disponibilité, ses bons conseils, ses relectures et sa gentillesse ; *Égidio Marsico* pour ses divers conseils (informatiques, bibliographiques, etc.) ; mon consultant personnel en « mise en page Word », *Françoise Rose* ; mon professeur et relecteur particulier d'anglais, *Caroline Rossi*.

Un grand merci tout particulièrement aux *Caro* (-*Cau*, -*Ja* et -*Ro*), *Françoise*, *Laëtitia*, *Ristian*, *Seb*, *Sophie*, *Vinz*, pour leurs conseils, leur soutien moral, leur bonne humeur, leurs blagues, pour les pauses et soirées jeux, et pour avoir supporté mes angoisses...

Je remercie également toutes les personnes avec qui j'ai travaillé ces dernières années à l'université et les membres du Projet ReFlex (Projet n° NT 05-2_41686 Renforcer la flexibilité lexicale et syntaxique dans l'expression orale et écrite des adolescents) : Michel Bert, Bernard Lété, Cécile Lux, Séverine Maggio, Sophie Manus, Soraya Mokrani, Sandra Teston, Lolke Van der Veen, *etc.*

Je ne peux également oublier tous mes amis, qui en dehors du laboratoire, m'ont accompagnée tout le long de cette expérience, et notamment : Bibi&Dadou, Delphine&Guillaume, Jess&Lolo, Jen, Jojo&Isa, Mag, Nayomi, Poulaine.

Merci pour leur amitié,

Merci pour leur soutien,

Merci pour leur ouverture d'esprit,

Merci pour ne pas avoir posé trop souvent la douloureuse question : « alors, tu finis quand ??? »,

Merci à Nayomi pour ses relectures, ses conseils, son hôtel, ...

Bien entendu, ces remerciements ne seraient complets sans la mention de ma famille,
Merci à mes parents pour m'avoir donné le goût des études, pour m'avoir encouragé, et donné
cette exigence de moi-même indispensable dans le milieu scientifique,
Merci à mon petit frère Rémy et à ma petite sœur Carol-Anne pour avoir joué le rôle d'ainé
durant les moments les plus difficiles,
Merci à mes grands-parents pour la fierté qu'on peut lire dans leurs yeux,
Merci à ma belle-famille.
Enfin, pour terminer, je tiens à remercier mon mari, Damien, pour sa présence, sa confiance,
son soutien, sa patience, pour avoir su trouver les mots justes dans les moments les plus
difficiles, et bien entendu pour son amour ...

[Résumés]

Le flux de l'information, aspects syntaxiques et discursifs : Une étude fonctionnaliste et développementale

Le but de ce travail de thèse, qui fait partie du projet ANR ReFlex (*n° NT 05-2_41686 Renforcer la flexibilité lexicale et syntaxique dans l'expression orale et écrite des adolescents*), est l'étude des capacités linguistiques d'enfants et d'adolescents appartenant à une CSP défavorisée et scolarisés en ZEP dans la région de Lyon. Partant du constat que l'adolescence était une période clef du développement de l'expression orale et écrite en situation monologique (Berman, 2006), nous étudions les productions de trois groupes d'enfants monolingues français (scolarisés en CM2 (âge moyen : 10;9), 5^{ème} (âge moyen : 12;7) et 3^{ème} (âge moyen : 15;2)). Les textes ont été produits dans divers contextes de production : à l'oral et à l'écrit, en exposition et en narration. Nous étudions alors le passage de l'enfant du statut de « *native speaker* » à celui de « *proficient speaker* » (Berman, 2004:9, 2008:2). L'enfant, après avoir acquis des unités et des règles linguistiques, doit apprendre lors de l'encodage d'un message à s'adapter aux différents contextes communicationnels, *etc.* : il doit devenir flexible. Des analyses *off line* ont été effectuées, et mon travail de thèse se concentre essentiellement sur une thématique spécifique : les syntagmes nominaux dans une perspective syntaxique, développementale et discursive. Notre analyse des syntagmes nominaux est fortement inspirée par l'hypothèse de la *Preferred Argument Structure* (Du Bois 1987) et examine la cohésion référentielle de textes produits dans différents contextes (oral/écrit ; narratif/expositif). Notre but est ainsi de démontrer : (1) que les choix syntaxiques sont gouvernés par les contraintes communicationnelles et (2) que ces contraintes communicationnelles varient selon la modalité de production (oral ou écrite).

Mot clef : Acquisition, contexte de production, développement, fonctionnalisme, langage, littéracie, production orale et écrite, psycholinguistique, transmission de l'information, syntaxe

Information flow, syntactic and discursive aspects: A developmental and functionalist approach

The main aim of this research, which is apart of a large project (the ANR project: ReFlex project, n° NT 05-2_41686 Renforcer la flexibilité lexicale et syntaxique dans l'expression

orale et écrite des adolescents), is the study of linguistic capacities in lower-class children attending an educational priority school in the suburbs of Lyon. Much development in the ability to produce a monologue text occurs during childhood and adolescence (Berman, 2006). For this reason, I have studied the productions of three groups of French monolingual children (mean age: 10;9 (CM2), 12;7 (5ème) and 15;2 (3ème)). Each participant produced two text types (expository/narrative) in two modalities (written/spoken). The primary focus is on how children grow from “native speaker” to “proficient speaker” (Berman, 2004:9, 2008:2). After having acquired linguistic forms and rules children must learn to adapt their verbal message to the different communicational contexts, etc.: this requires that children become expressively flexible. Off line analyses were conducted, and my PhD thesis tackles a specific theme: noun phrases in a syntactic, developmental and discursive perspective. Our analysis of noun phrases is inspired by the hypothesis of Preferred Argument Structure (Du Bois 1987) and examines referential cohesion in texts produced in different contexts (spoken/written, expository text/narrative text). Our aim is to demonstrate: (1) that language production is governed by communicative constraints; and (2) that those communicative constraints differ for written and spoken production.

Keywords: Acquisition, context of production, development, functionalism, language, literacy, spoken and written production, psycholinguistic, information flow, syntax

Liste des abréviations

A	Amplitude
Anc.	Ancienne information
Au	Position autre
CSP	Classe Socio-Professionnelle
E1	Niveau explicite 1 du Modèle RR
E2	Niveau explicite 2 du Modèle RR
E3	Niveau explicite 3 du Modèle RR
Expo.	Texte expositif
I	Niveau implicite du Modèle RR
INSEE	Institut National des Statistiques et des Études économiques
LEP	Lycée d'Enseignement Professionnel
MDT	Mémoire de travail
MLT	Mémoire à long Terme
Narr.	Texte narratif
Nou.	Nouvelle information
NS	Non significatif
O	Position objet
PAS	Preferred Argument Structure
Pro.	Pronom
REP	Réseau d'Éducation Prioritaire
(Modèle) RR	Modèle de la redescription des représentations
S	Position Sujet
SI	Position sujet intransitif
SC	Position sujet de copules
SN	Syntagme nominal
SNL	Syntagme nominal lexical
ST	Position sujet transitif
SVO	Sujet-Verbe-Objet
UT	Unité Texte
VD	Variable Dépendante
VI	Variable Indépendante
ZEP	Zone d'Éducation Prioritaire
σ	Écart-type

La production langagière au-delà de 10 ans Introduction générale

i. Informations liminaires

La production langagière écrite comme orale est au centre des préoccupations gouvernementales et notamment des professionnels de l'éducation et de la santé (Demont, 2007). La maîtrise des outils linguistiques et leur adaptation aux divers contextes de production semblent en effet essentielles afin d'assurer la réussite scolaire, professionnelle et sociale (Verdelhan-Bourgade, 2002). L'intérêt des chercheurs pour la production de l'écrit et de l'oral dans une optique cognitive et psycholinguistique apparaît dans les années 80 avec Hayes et Flower (1980) et Levelt (1989).

L'étude de la production langagière en psycholinguistique mobilise deux types de méthodologies : (a) une méthode *off line*¹ et (b) une méthode *on line*.² L'analyse *on line* correspond à la prise en considération des processus engagés en temps réel dans l'activité langagière en se penchant sur divers types d'indicateurs tels que les pauses, les débits et les protocoles verbaux (Fayol, 1997). L'analyse *off line*, qui est la méthode la plus ancienne et qui a permis l'élaboration de la plupart des modèles de production (Fayol, 1997), se concentre sur les produits finis de la production langagière.

Notre travail de recherche, qui s'inscrit dans le cadre du projet ANR ReFlex,³ constitue une analyse *off line* de productions langagières de populations appartenant à des classes sociales

¹ Entre autres : Argerich et Tolchinsky (2000), Berman (2004, 2008), Berman et Verhoeven (2002), Jisa (2004), Jisa, Reilly, Verhoeven, Baruch et Rosado (2002), Jisa et Strömquist (2002), Jisa et Viguié (2005), Ravid et Tolchinsky (2002), Ravid (2000, 2002, 2005), Ravid, Van Hell, Rosado et Zamora (2002), Strömquist, Johansson, Kriz, Ragnarsdottir, Aisenman et Ravid (2002), Tolchinsky, Perera, Argerich, Aparici (1999).

² Entre autres : Chanquoy, Foulin et Fayol (1990), Duez (1991, 2001), Fayol (1997), Flower et Hayes (1981), Foulin (1995, 1998), Hayes et Flower (1980), Matsuhashi (1981), Schilperoord (1996), Strömquist (2000), Strömquist, Holmqvist, Johansson, Karlsson, et Wenglin (2006), Zellner (1994).

³ Projet n° NT 05-2_41686 Renforcer la flexibilité lexicale et syntaxique dans l'expression orale et écrite des adolescents ; porteur du projet, Harriet Jisa (Laboratoire DDL, UMR 5596, Université de Lyon) ; coordinateur, Florence Chenu (Laboratoire DDL, UMR 5596, Université de Lyon) ; collaborateur principal, Michel Fayol (Laboratoire LAPSCO, UMR 6024, Université Blaise Pascal).

défavorisées. Nous nous concentrons sur les collégiens, étant donné que les travaux antérieurs établissent que l'adolescence est une période clef du développement de l'expression orale et écrite (entre autres : Berman, 1997, 2006 ; Berman et Slobin, 1994 ; Ravid, 2005). Leur répertoire lexical se développerait considérablement, il deviendrait plus complexe et plus diversifié (entre autres : Aisenmann et Berman, 2000 ; Argerich et Tolchinsky, 2000 ; Gayraud, 2000 ; Johannsson, 1999 ; Katzenberger, 2004). Leur syntaxe connaîtrait la même évolution : les connexions syntaxiques, la structure des syntagmes verbaux et lexicaux, entre autres, deviendraient également plus complexes et plus diversifiés (entre autres : Berman, 1999, 1995, 2000, 2008 ; Berman et Katzenberger, 2004 ; Berman et Slobin, 1994 ; Berman et Verhoeven, 2002 ; Jisa, 1998, 2000, 2004 ; Gayraud, Jisa et Viguié, 1999 ; Ravid, 2000, 2002, 2004). Pour mener à bien ce projet, l'équipe s'est inspirée d'une étude internationale (le projet Spencer)⁴ qui avait pour but l'étude du développement de l'expression orale et écrite, dans une perspective translinguistique, de quatre populations de classe moyenne supérieure : des enfants de 9/10 ans, des pré-adolescents de 12/13 ans, des adolescents de 15/16 ans et des adultes BAC+5.⁵

Une opposition majeure organise les théories linguistiques depuis plusieurs décennies. Deux courants théoriques principaux sont alors distingués (Caron, 1989/2001). Selon la première approche, une théorie linguistique a pour but d'explicitier les contraintes formelles qui déterminent les règles de combinaison des signes des langues (Chomsky, 1957, par exemple). Selon la deuxième approche, une théorie linguistique a comme objectif de rendre compte de la façon dont les différentes langues (ou des locuteurs différents au sein d'une même communauté linguistique) mettent en relation des formes linguistiques et des significations. Cette seconde approche, dans laquelle nous nous inscrivons, répond à la perspective fonctionnaliste. Quatre caractéristiques principales peuvent définir le fonctionnalisme et nous verrons dans quelle mesure chacune d'entre elles justifie notre position.

En premier lieu, le langage apparaît comme un instrument de communication, fonction première et essentielle du langage (entre autres : Bassano, Maillochon, Klampfer et Dressler, 2001 ; Bates, McNew, MacWhinney, Devescovi et Smith, 1982 ; Bates et MacWhinney,

⁴ Ce projet regroupait sept pays, l'Espagne, la France, l'Islande, Israël, les Pays-Bas, la Suède et les États-Unis d'Amérique. Cette étude, dont le responsable scientifique est Ruth Berman, était financée par la fondation Spencer (USA) et intitulée « Developing literacy in different contexts and different languages ».

⁵ Entre autres : Berman (1995, 1999a, 1999b, 2000, 2006), Jisa et Viguié (2005), Jisa (2004), Jisa, Reilly, Verhoeven, Baruch et Rosado (2002), Jisa et Strömqvist (2002), Ravid (2000, 2005, 2006) ; Ravid, Van Hell, Rosado et Zamora (2002), Strömqvist, Johannsson, Kriz, Ragnarsdottir, Aisenman et Ravid (2002), Tolchinsky, Perera, Argerich, Aparici (1999). Pour une analyse de la connectivité intra-textuelle du français, se reporter à la thèse de Viguié-Simon (2001).

1989 ; Hickmann, 2000 ; Karmiloff-Smith, 1979a ; Matthey, 2003 ; etc.). « *L'idée de base de l'approche fonctionnaliste est ainsi que les formes des langues naturelles sont créées, gouvernées, contraintes, acquises et utilisées au service des fonctions de communication* » (Bassano *et al.*, 2001:84). Cet instrument de la communication possède deux propriétés essentielles : sa plurifonctionnalité (une forme peut avoir plusieurs fonctions et une fonction plusieurs formes) et son ancrage dans le contexte d'énonciation (un message ne peut être interprété en isolation de son contexte de production) (entre autres : Hickmann, 2000, 2002, 2003). Dans notre protocole expérimental, nous avons choisi de contrôler au maximum le contexte de production afin que ce dernier soit une variable indépendante dans l'analyse. Les participants ont en effet produit quatre textes : un expositif oral, un expositif écrit, un narratif oral et un narratif écrit. Ainsi, nous nous attendons à observer des différences de choix linguistiques selon les contextes de production.

En second lieu, trois des niveaux de la sphère linguistique – la syntaxe, la sémantique et la pragmatique – apparaissent comme des niveaux interdépendants. Cette perspective prend également en compte le niveau d'organisation du discours, en s'intéressant à des thématiques telles que le statut informationnel des éléments (à savoir le continuum de la nouvelle à l'ancienne information) qui est au cœur de nos préoccupations (*Cf.* Chapitre 1, p. 32).

En troisième lieu, le langage ne peut être traité sans faire « *référence aux caractéristiques générales de la cognition humaine* » (Matthey, 2001:25) qui, comme nous le verrons, sont impliquées dans la production textuelle ; nous pensons notamment aux types de planification mobilisée dans une tâche de production (*Cf.* Chapitre 2, p. 79).

En quatrième lieu, la forme linguistique est au service de la fonction. Selon une approche fonctionnaliste, le développement du langage implique le passage graduel (a) d'une série de marqueurs juxtaposés, monofonctionnels et de processus de traitement à (b) l'organisation intralinguistique de systèmes plurifonctionnels d'options permettant de faire varier le sens (Karmiloff-Smith, 1979a). Cette conception de la forme linguistique nous intéresse tout particulièrement dans la mesure où nous nous concentrons sur l'usage des formes pronominales *versus* lexicales (*Cf.* Chapitre 1, p. 40).

Un des modèles les plus intéressants s'inscrivant dans la perspective fonctionnaliste est celui de Karmiloff-Smith (1979, 1981, 1986, 1990, 1993/1995). La chercheuse développe une conception permettant de réconcilier la théorie purement innéiste (Chomsky, 1957) ; Fodor, 1983, 1983/1986) et la théorie constructiviste (Piaget, 1923). L'être humain naît avec des

prédispositions qui vont évoluer tout au long du développement du langage, ce qui est explicité par le *Modèle RR* de Karmiloff-Smith. Selon ce modèle, le traitement du langage passe par deux processus au cours du développement : celui de *la modularisation progressive* et celui de la *redescription des représentations*. L'enfant possède au départ des connaissances générales et se spécialise peu à peu en se construisant en interaction avec l'environnement. L'existence de connaissances prélinguistiques innées est reconnue mais la pertinence du développement du langage également. Tout au long de la vie, la langue du locuteur évolue.

ii. Qu'est-ce que le développement du langage ?

Tous les locuteurs/scripteurs⁶ apprennent leur langue maternelle, de l'état de fœtus à celui d'adulte en passant par l'enfance et l'adolescence (Karmiloff et Karmiloff-Smith, 2001). Dans ce domaine est distinguée l'acquisition précoce (jusqu'à environ cinq ans) de l'acquisition tardive (au-delà de cinq ans). L'intérêt de notre étude se trouve être l'évolution de l'individu locuteur natif (de plus de cinq ans) au locuteur performant ayant les capacités de gérer des situations langagières diverses et complexes l'obligeant à manipuler ses connaissances linguistiques, cognitives, discursives et culturelles. Nous nous intéressons à l'usage de la langue dans divers contextes de production, et donc à l'adaptabilité de l'individu. Nous partons du postulat que le langage continue de se développer après cinq ans et durant les années de scolarisation (entre autres, Berman, 2006). Nous étudions ainsi le passage chez l'enfant du statut de *native speaker* à celui de *proficient speaker* (Berman, 2004, 2006, 2008).

Le locuteur passerait par cinq phases principales lors de son développement (Berman, 2004:14). La première phase (*rote-knowledge*) correspond à une phase initiale de l'acquisition du langage durant laquelle les enfants apprennent des formes linguistiques sans les analyser. Lors de la seconde phase (*initial alternations*), les individus utilisent ces formes dans des paradigmes et des constructions. Lors de la troisième phase (*interim schemata*), les enfants tentent d'appliquer certaines règles de la langue et tentent de faire des généralisations. Lors de la quatrième phase (*rule knowledge*), le locuteur devient capable de suivre les règles mais des lacunes demeurent. Ces quatre premières phases caractérisent le développement de l'enfant

⁶ Par locuteurs/scripteurs, nous entendons un ensemble de personnes parlant et écrivant.

avant 5 ans. Après 5 ans, les enfants entrent dans la dernière phase (*proficient integration of knowledge and use*) qui se définit comme l'usage mature de la langue. Lors de l'encodage d'un message, l'enfant, après avoir acquis des unités et des règles linguistiques, doit apprendre à s'adapter aux normes communicationnelles de sa communauté et aux différents contextes communicationnels. Il doit devenir flexible. Dans ce travail de thèse, nous nous intéressons à cette dernière phase de l'acquisition du langage en nous concentrant sur un type spécifique de langue, *la langue de scolarisation*.

iii. La langue de scolarisation et la littéracie linguistique

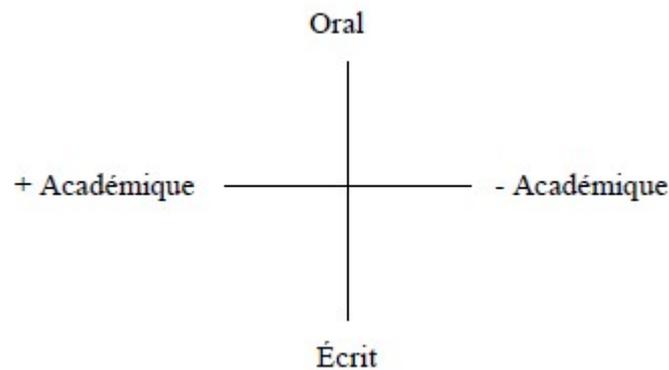
Il est communément admis que, contrairement à l'oral, l'écrit est, prototypiquement, un état plus académique, plus détaché et mieux planifié de la langue (Biber, 1988 ; Ravid et Berman, sous presse ; Ravid et Tolchinsky, 2001). En effet, à l'écrit, comme nous l'évoquons dans la Section 2.2 (p. 54), le rythme de production est lent, le temps de planification peut alors être plus conséquent. Ce temps de planification peut permettre, entre autres, la mobilisation et la production de formes définies comme plus coûteuses, telles que les subordinations non finies, le passif ou encore les *sujets lourds*. Ces formes caractérisent un état soutenu, formel ou encore académique de la langue, acquis par le biais de *la langue de scolarisation*.

Cependant, il existe une interaction entre le registre (informel/formel) et la modalité (oral/écrit) : ainsi pouvons-nous trouver des oraux académiques comme les présentations scientifiques, et des écrits moins conventionnels, tels les *sms* (Gregory et Carroll, 1978 ; Gregory, 1967). Si l'écrit est, prototypiquement, perçu comme plus académique que l'oral, ces deux modalités ne peuvent être considérées comme deux langues différentes. En effet :

« Produire du langage écrit, c'est produire du langage mais dans des conditions, qui, pour certaines au moins, diffèrent parfois de manière considérable des conditions usuelles de la production orale. Pourtant comme le montrent les travaux ayant comparé les modalités orales et écrites, l'opposition oral/écrit n'est pas absolue. Elle relève plutôt d'un continuum allant d'activités à dominante informelle [...] à d'autres activités formellement plus contraintes [...] » (Fayol, 1997:4-5).

La Figure 1 sous-entend que les modalités ont des sous types plus ou moins éloignés de la situation prototypique.

Figure 1 : Le continuum oral/écrit



Les situations telles que la production écrite en milieu scolaire font appel à un registre plutôt académique, il s'agit d'une activité à *dominante formelle*. Durant ses années de scolarisation, l'individu doit apprendre une langue de prestige, que Verdelhan-Bourgade (2002) appelle *la langue de scolarisation*, et qui joue trois grands rôles dans le milieu scolaire. Premièrement, elle est objet d'enseignement, c'est-à-dire que les enfants apprennent à travailler sur cette *langue de scolarisation* en tant que discipline à part entière. Elle apparaît alors comme un objet d'étude (apprentissage de la grammaire). Deuxièmement, elle a un rôle de médiation, c'est-à-dire que sa maîtrise et son utilisation permettent l'apprentissage d'autres disciplines (comme l'histoire, les mathématiques ou encore la biologie). Troisièmement, elle constitue une langue de communication scolaire, qui conditionne l'insertion dans le système générant la réussite scolaire.

Les tâches que nous avons fait réaliser aux participants de notre étude peuvent faire appel à la *langue de scolarisation*. Selon le type de texte à produire, les enfants utilisent davantage ce code, notamment en production expositive écrite. Les individus de notre étude produisent dans un contexte bien spécifique. En production écrite comme orale, ils sont en situation monologique ; ceci a du sens dans une étude ayant pour objet la *langue de scolarisation*. En effet, les tâches requises correspondent à des exercices à réussir à l'école. Néanmoins, à l'oral, les participants étaient face à un expérimentateur qui, sans nul doute, pouvait transmettre des messages non verbaux. Ainsi, les écrits demandés sont des écrits assez prototypiques, faisant alors appel à la *langue de scolarisation* ; les oraux sont un peu moins prototypiques dans la mesure où les individus doivent produire un monologue.⁷

⁷ Les critères de prototypie des deux modalités sont définis dans la Section 2.2 (p. 54).

Dans notre société occidentale, une place toute particulière est accordée à l'écrit, qui est considéré comme une norme.⁸ La gestion de l'écrit prototypique et normatif, à savoir *la langue de scolarisation*, « accroît les chances d'intégration, voire d'ascension dans l'échelle sociale » (Guérin, 2006:61). Cette norme renvoie à un certain idéal, qui ne doit pas être transgressé, et qui peut être conçu comme « *la synthèse des usages observables dans les écrits des auteurs dits « bons »* » (Guérin, 2006:73). Cet idéal est enseigné à l'école à travers, notamment, la grammaire scolaire. Lorsque nous parlons de registre académique ou encore formel, sans émettre aucun jugement de valeur, nous faisons référence à cet état de la langue, certes artificiel et surévalué, mais normé, dont la maîtrise est exigée dans notre société. Nous disons alors que l'écrit est plus planifié que l'oral.

La langue de scolarisation peut être abordée à travers le concept de littéracie linguistique (Ravid et Tolchinsky, 2000, 2002). La littéracie linguistique sous-entend l'idée que les individus, avec les années de scolarisation, développent leurs connaissances sur la langue écrite et orale (Berman, 1998, 2006 ; Ravid et Tolchinsky, 2000, 2002 ; Tolchinsky, 2006). Être linguistiquement mature suppose la possession d'un répertoire linguistique, qui englobe un large éventail de registres et de genres, et la capacité de s'adapter aux contextes communicationnels (Berman, 2006 ; Berman et Ravid, 2009 ; Ravid et Tolchinsky, 2000, 2002 ; Tolchinsky, 2006). Cette littéracie est liée à la flexibilité et à l'adaptabilité rhétorique : les locuteurs/scripteurs ont besoin de capter l'attention des destinataires par le biais de structures linguistiques adaptées (Ravid et Tolchinsky, 2000, 2002).

La littéracie linguistique consiste en un trait spécifique : *le contrôle des variations linguistiques* (Ravid et Tolchinsky, 2000, 2002). *Le contrôle des variations linguistiques* et l'encodage des messages par le biais de structures linguistiques adaptées à divers contextes apparaissent comme un des objectifs principaux des systèmes d'éducation.

La conception de la littéracie linguistique inclut un processus spécifique : *le métalangage* (Berman, 2006 ; Ravid et Tolchinsky, 2000, 2002). Si les compétences linguistiques sont essentiellement implicites, les locuteurs/scripteurs développent, par le biais de l'expérience et de la scolarisation, une connaissance plus explicite et analytique de la langue. Ils déploient la capacité de contrôler leur répertoire linguistique et de réorganiser leurs représentations mentales afin de devenir plus performants.

⁸ Par norme, nous ne référons pas ici à la *norme normal* et donc à une moyenne mais à la norme comme caractère normatif renvoyant à de l'évaluatif et à du prescriptif (Guérin, 2006 ; Rey, 1972).

La littéracie linguistique implique également une condition, *la familiarité et maîtrise de l'action d'écrire et du langage écrit*. Ceci exige la maîtrise de deux aspects concernant la modalité écrite (Berman, 2006 ; Ravid et Tolchinsky, 2000, 2002 ; Tolchinsky, 2006) : (a) les individus doivent appréhender la production écrite comme un discours spécifique en le distinguant notamment de la production orale (ces deux types de productions impliquent des contraintes différentes et donc des choix linguistiques différents)⁹ ; (b) les individus doivent apprendre à considérer l'écrit comme un système notationnel spécifique (gestion du système orthographique et de la ponctuation).

Cette vision de la littéracie est totalement adaptée à notre étude dans la mesure où elle propose d'analyser son impact sur le développement linguistique de l'enfant. En effet, nous testons l'adaptabilité des participants de notre étude face aux différents contextes de production (narratif écrit, narratif oral, expositif écrit, expositif oral) imposés par le protocole expérimental.

iv. Notre étude

Notre protocole expérimental a été conçu dans le but de contrôler l'éventuel impact de diverses variables indépendantes : (a) le NIVEAU SCOLAIRE (CM2 (âge moyen = 10;9) / 5^{ème} (âge moyen = 12;7) / 3^{ème} (âge moyen = 15;2)) ; (b) la MODALITÉ (orale/écrite) ; (c) le TYPE DE TEXTE (expositif/narratif) ; et (d) l'ORDRE DE PRODUCTION (l'écrit avant l'oral et l'oral avant l'écrit).

Notre travail de thèse se concentre sur l'analyse des produits finis de trois groupes d'enfants en tentant d'étudier en profondeur une thématique spécifique : les syntagmes nominaux (SN) dans une perspective syntaxique, développementale et discursive. L'étude du syntagme nominal se justifie par différents aspects.

Premièrement, la langue française recourt à des formes linguistiques différentes pour encoder un référent selon (a) son statut informationnel dans le discours (1) et (b) sa position syntaxique (1) et (2). Dans (1), la nouvelle information, *une petite fille*, est introduite sous forme lexicale en position non sujet. Puis le référent est ensuite maintenu par le biais d'un pronom, *elle*, en position sujet. Dans (2), en position non sujet, nous pouvons trouver des syntagmes nominaux lexicaux complétés par des modifieurs (*une gentille petite fille*). En revanche, en position sujet, les syntagmes nominaux lexicaux sont plus simples (*elle, une*

⁹ Les différences entre la production orale et la production écrite sont développées dans la Section 2.2 (p. 54).

filles). Il s'agit de préférences linguistiques reflétant notamment la transmission de l'information en production orale.

(1) *Il y a une petite fille, elle s'est perdue. Elle pleure.*

(2) *Il y a une gentille petite fille, elle s'est perdue. / Une fille s'est perdue.*

Deuxièmement, l'intérêt pour le syntagme nominal vient également du fait qu'il est également une unité intermédiaire, plus petite que la phrase mais plus importante que le mot, qui implique une organisation en lien avec ces deux unités extrêmes (Ravid et Berman, sous presse).

Troisièmement, le syntagme nominal est également intéressant par le rôle qu'il joue dans le système référentiel d'une langue : il peut introduire une information et également maintenir des référents. Le statut informationnel appartient au système référentiel de la langue qui se développe avec l'âge et qui est une des exigences de la cohérence d'un texte (Hickmann, 2000, 2002, 2004).

Quatrièmement, nous nous concentrons davantage sur les syntagmes nominaux lexicaux (SNL) dans la mesure où ils mettent en jeu des items lexicaux qui jouent un rôle essentiel dans l'acquisition du langage (Dockrell et Messer, 2004 ; Ravid 2005, 2006 ; Ravid et Berman, sous presse) ainsi que dans la construction de la structure du discours (Francis, 1986 ; Ravid et Berman, sous presse).

Enfin, cinquièmement, le nom constituant le SNL a également un rôle important dans le développement de la syntaxe ; les constructions syntaxiques dépendraient en partie des noms utilisés (Ravid, 2006 ; Ravid et Berman, sous presse). Le lexique et la syntaxe sont fortement liés et, plus les textes contiennent de SNL, plus les constructions syntaxiques les réalisant sont complexes (Ravid, 2006).

Aussi allons-nous observer la production des unités pronominales et lexicales en prenant en considération les positions syntaxiques, les statuts informationnels et le contexte de production. Le statut informationnel, à savoir la différenciation et la manipulation de l'information nouvelle et ancienne, fait partie du développement discursif de l'enfant. Il « constitue l'une des facettes les plus importantes du langage », notamment de l'acquisition tardive de l'enfant (Hickmann, 2002:181). L'enfant doit alors acquérir les principes pragmatiques qui gouvernent le flux informationnel et donc, organiser l'information à travers les phrases afin de construire un discours cohésif (Hickmann, 1995). C'est dans cette mesure

que nous nous concentrons sur le SN qui peut soit introduire une nouvelle information soit maintenir une ancienne.

Afin d'observer au mieux cette unité linguistique et de compléter les recherches dans ce domaine, notre protocole expérimental présente cinq originalités.

Dans un premier temps, les enfants ayant participé à notre étude sont tous issus d'un même réseau ZEP/REP, ainsi nous pouvons imaginer que nos individus de CM2 sont les futurs 5^{ème} et ces derniers les futurs 3^{ème} de notre étude. Ceci nous permet de parler d'étude développementale transversale.

Dans un second temps, la thématique des textes a été contrôlée. Les enfants devaient tous produire autour du thème, *les problèmes entre les gens*. Une vidéo concernant les problèmes entre les gens a été projetée au début de la collecte des données. Le fait d'orienter le thème des productions permet alors d'obtenir des textes comparables de par leur contenu. De plus, les enfants ont produit des textes monologiques ; l'observation de monologues plutôt que de conversations, par exemple, est un des meilleurs moyens d'évaluer et de comprendre le développement langagier tardif (Berman, 2006 ; Berman et Slobin, 1994 ; Ravid et Berman, sous presse). En situation monologique, l'enfant doit faire appel à des connaissances pragmatiques et discursives différant de celles de la situation conversationnelle, acquises très tôt (Berman, 2006).

Dans un troisième temps, le protocole expérimental a été conçu de sorte à obtenir pour chaque individu une version orale et une version écrite d'une production ainsi que deux genres textuels. Chaque participant a en effet produit quatre textes : un narratif oral, un narratif écrit, un expositif oral et un expositif écrit. Les enfants devaient raconter la même histoire à l'oral et à l'écrit ; et ils devaient également rapporter un exposé équivalent dans les deux modalités. La comparaison entre la version orale et écrite, ou encore entre le type de texte expositif et narratif, est alors possible pour un même individu. Ceci n'est pas le cas de toutes les études se consacrant à la comparaison des deux modalités du langage ou encore des différents types de textes (par exemple, les travaux de Blanche-Benveniste, 1995).¹⁰ Nous pouvons alors observer le comportement linguistique d'un même individu dans des contextes différents. Une place toute particulière est accordée au contexte de production, l'intérêt étant de voir son impact sur les locuteurs/scripteurs, notamment sur leurs choix linguistiques. De plus, nous voulons

¹⁰Dans ces travaux, Blanche-Benveniste utilise la conversation spontanée pour la modalité orale et des articles de presse pour la modalité écrite. Ainsi, la variation observée n'est pas due seulement à la modalité de production mais également à la thématique, au but communicatif ou encore à l'âge des locuteurs/scripteurs ayant produit.

étudier la relation que peuvent entretenir la modalité de production (oral/écrit), impliquant des différences importantes de processus, et le type de texte (narratif/expositif), engendrant des styles de discours différents (Berman et Ravid, 2009).

Dans un quatrième temps, le protocole a été prévu de sorte à contrôler l'ordre de passation. Nous avons pu observer si le fait de produire dans un premier temps un texte écrit et dans un second temps un texte oral, ou l'inverse, avait des conséquences quant aux choix des structures linguistiques. Ce contrôle semble essentiel dans la mesure où des études pilotes ont révélé un effet significatif de l'ordre de production pour certains critères d'analyse (Fayol, Jisa et Mazur-Palandre, 2007, 2008 ; Jisa, 2004 ; Jisa et Mazur, 2006a, 2006b ; Gayraud *et al.*, 1999, 2001 ; Mazur-Palandre, 2007a, 2007b, 2008a, 2008b).

Dans un cinquième temps, notre recherche s'inscrivant dans une perspective fonctionnaliste, nous avons effectué nos analyses en tenant compte des différents niveaux linguistiques ; nous étudions certes la syntaxe mais ne l'isolons pas de la pragmatique et de la sémantique, d'où notre intérêt pour le statut informationnel des unités observées.

Enfin, pour finir, l'équipe de recherche a pris la décision de se pencher sur une population de CSP défavorisée scolarisée dans des établissements ZEP/REP. Dans l'académie de Lyon, l'éducation prioritaire concerne **un** élève sur **huit**,¹¹ ce qui n'est pas négligeable, d'où la nécessité de s'intéresser à des individus scolarisés en ZEP/REP.

v. La question de langue et classe sociale

Notre choix d'une population socio-économiquement défavorisée nous met face à la problématique de la relation entre langue et classe sociale ; nous tenons alors à éclaircir notre position. Dans les années 50, des études trouvent une corrélation faible entre la proportion des subordonnées et la classe socio-professionnelle des locuteurs mais ne concluent pas à un lien entre le niveau de langue et les classes socio-professionnelles (Hurell, 1956). Néanmoins, à la même époque, Sampson (1956) conclut d'une étude sur de jeunes enfants que le contexte social influence le développement du discours. Que penser de l'hypothèse de l'existence d'un lien entre compétences langagières et classes sociales ?

¹¹ Source : <http://www.ac-lyon.fr/specificites-academie-lyon.html>.

L'important n'est pas forcément une différence de classe sociale, mais plutôt la prise en considération, dans son analyse du langage, du contexte. Labov (1978/1993) est l'un des premiers à considérer la langue dans son contexte social et à montrer que, selon les contextes sociaux, nous pouvons observer des productions linguistiques différentes. Bernstein (1971, 1973, 1975a, 1975b, 2000, 2007) parle de deux codes principaux de la langue, le code élaboré et le code restreint (ou commun). Néanmoins, la méthodologie de Bernstein, basée principalement sur des questionnaires (qui encourageraient davantage la population favorisée), a été très critiquée et controversée (Léger, non publié). Les individus des classes ouvrières sont tout aussi capables de produire des textes cohérents et riches en information, même en utilisant le langage élaboré, si la production langagière se fait dans un contexte adéquat dans lequel l'enfant se sent à l'aise (Labov, 1978/1993). Ainsi, Labov montre que, face aux locuteurs du *vernaculaire noir-américain*, il faut manipuler le contexte pour que les locuteurs produisent de la façon la plus naturelle possible : Bereiter (cité par Labov, 1978/1993), lui, n'avait obtenu que des réponses monosyllabiques alors que Clarence Robins (collaborateur de Labov), grâce à un contexte modifié et adapté, a relevé des productions de meilleure qualité. En effet, des modifications importantes ont été apportées au contexte de production, comme : (a) permettre aux locuteurs d'être plusieurs lors de la tâche linguistique ; (b) le fait que l'expérimentateur appartienne lui-même à la communauté à laquelle les locuteurs appartiennent, ce qui a également des conséquences sur les productions telles que l'évocation de sujets tabous ; et (c) se mettre physiquement au même niveau que le sujet, en s'asseyant par terre comme lui. Les travaux de Labov (1978/1993) suggèrent que, du point de vue du texte narratif, du raisonnement, *etc.*, les membres de la *working-class* apparaissent « *comme des locuteurs bien plus efficaces que beaucoup de membres de la middle-class* » (Labov, 1978/1993:294).

Les différences langagières entre les classes sociales intéressent les chercheurs depuis plusieurs décennies et ont été à l'origine de nombreuses polémiques. Il existe certainement différents niveaux de maîtrise de la *langue de scolarisation* et l'enfant doit sans doute passer par un développement pour atteindre certaines capacités. Avoir la capacité de manier la *langue de scolarisation* est, de plus, essentiel à la bonne intégration scolaire, sociale puis professionnelle des individus. Se prononcer sur l'éventuelle existence de deux langues (codes élaboré *versus* code restreint) est difficile et sans doute non pertinent.

C'est en partant de telles conclusions que nous voulons établir, dans notre travail de thèse, le profil linguistique d'une population défavorisée sur un domaine sémantico-syntaxique spécifique du français. Le protocole a été élaboré de sorte que les enfants se sentent le mieux possible : ainsi, (a) nous avons fait produire les individus dans leur cadre scolaire (un des lieux où ils sont le plus habitués à produire *la langue de scolarisation*) ; (b) les individus connaissaient tous une des expérimentatrices¹² ; (c) des liens sociaux amicaux existaient entre cette dernière et les participants.

vi. La composition du travail de recherche

Ce travail de thèse s'articule en trois parties principales, la première théorique, la deuxième méthodologique et la dernière regroupant les résultats.

La première partie, constituée de trois chapitres, présente le cadre théorique dans lequel s'inscrit cette recherche. Le Chapitre 1 (p. 21) fait le point sur les concepts de cohésion/cohérence. Dans ce même chapitre, sont ensuite décrites les unités observées dans cette recherche, à savoir les formes pronominales et lexicales. Puis, la notion de statut informationnel est présentée avec une focalisation sur la *Preferred Argument Structure* (PAS) de Du Bois (1987). Dans le Chapitre 2 (p. 51), nous nous concentrons sur la production langagière ; deux thématiques principales sont abordées. Dans un premier temps, après avoir discuté des deux modalités orale et écrite du langage, nous situons notre travail de thèse parmi les modèles de production orale et écrite. Dans un second temps, la notion de type textuel est étudiée en nous focalisant sur deux genres particuliers : le texte expositif et le texte narratif. Enfin, dans le Chapitre 3 (p. 91), sont exposées nos questions et hypothèses de recherche.

La seconde partie, apparaissant sous le titre de *Méthodologie*, est constituée de trois chapitres. Le Chapitre 4 (p. 113) fait le point sur la collecte des données et le protocole expérimental en fournissant des informations sur : (a) les aspects expérimentaux de la recherche, (b) la sélection des lieux et des individus et (c) le choix du matériel. Le Chapitre 5 (p. 131) se concentre essentiellement sur l'exploitation des données avec une présentation succincte des logiciels utilisés lors de la transcription et du codage des productions. Le Chapitre 6 (p. 141)

¹² Cette personne, l'auteur de cette thèse, travaillait depuis deux ans, comme surveillante, dans le collège dans lequel étaient scolarisés les deux tiers des participants de cette étude.

précise certains aspects du codage des unités linguistiques observées dans ce travail de thèse, les SNL et les pronoms.

Dans la troisième partie de ce travail, intitulée *Résultats*, sont exposées les analyses conduites dans le cadre de cette thèse. Dans un premier temps (Chapitre 7, p. 161) sont présentés des résultats concernant des indicateurs généraux, principalement de longueur de textes. Les proportions de deux unités linguistiques, les pronoms et les SNL, sont l'objet des préoccupations du Chapitre 8 (p. 175). Suit le Chapitre 9 (p. 205) dans lequel, en partant du postulat que la production de SNL est un indice de maturité linguistique, nous nous concentrons sur le degré de lexicalité des SNL. Notre analyse de la production verbale est complétée par une étude sur les rôles informationnels de ces deux unités. Dans le Chapitre 10 (p. 229) sont alors mis en relation les formes linguistiques et les statuts informationnels, dans la mesure où le statut informationnel influence les choix linguistiques des individus (entre autres : Ariel, 1990 et 1996 ; Du Bois, 1987, 2003 ; Khorounjaia et Tolchinsky, 2004 ; Levelt, 1989 ; Sanford et Garrod, 1981a, 1981b). Cette partie résultat se poursuit avec le Chapitre 11 (p. 259) relatif à l'utilisation non conventionnelle des pronoms (*référent manquant, référents multiples, réintroductions pronominales, problème d'accord*). Si les individus de notre étude, et même les plus jeunes, ont une faible proportion d'utilisation non conventionnelle des pronoms, c'est sans doute qu'ils savent que le SNL doit être utilisé pour introduire/réintroduire une information ou pour éviter les ambiguïtés. Nous décidons alors de nous pencher davantage sur le statut de l'ancienne information. En effet, ce statut référentiel peut être encodé soit par une forme pronominale soit par une forme lexicale. Le choix le plus naturel, et obéissant aux principes de transmission de l'information de la langue orale, est d'utiliser un pronom. Mais, est-ce que ce principe résiste au NIVEAU SCOLAIRE, au TYPE DE TEXTE, à la MODALITÉ et à l'ORDRE DE PRODUCTION ? C'est ainsi que le Chapitre 12 (p. 267) est consacré aux SN maintenant une information et notamment à leurs formes.

Enfin, des conclusions et une mise en perspective achèvent ce travail de recherche (p. 295).

Partie I

La production textuelle : principes et modèles théoriques

Chapitre 1

Cohérence/Cohésion : concepts et marques

Ce chapitre présente les concepts de cohésion et de cohérence essentiels pour une étude prenant en considération des unités linguistiques interprétables à partir du texte dans sa globalité (Section 1.1, p. 22). La notion de statut informationnel et de *Preferred argument structure* (PAS) (Du Bois, 1987) sont par la suite définis dans la Section 1.2 (p. 32). Suit une présentation des deux unités linguistiques étudiées dans ce travail, à savoir les formes pronominales et les formes lexicales (Section 1.3, p. 40). Ce chapitre est complété par un état des lieux des recherches menées en acquisition tardive concernant le SN, la continuité référentielle et la PAS.

1.1. À la recherche des définitions de la cohésion et de la cohérence

1.1.1. Cohérence et cohésion : différenciation

L'étude psycholinguistique du texte dans sa globalité trouve sa pertinence dans la mesure où certains domaines syntaxiques, tels que le SN, un des centres de préoccupations de cette

recherche, ne peuvent s'expliquer seulement par l'observation de phrases isolées mais exigent la prise en considération du texte tout entier. Dans cette optique, nous travaillons sur le texte dans sa globalité et, comme le disent Kail et Fayol (2000) en parlant des travaux de Hickmann (2000), nous mettons en relation la phrase et le discours. Il semble alors essentiel de travailler les notions de *cohérence* et de *cohésion*, dans une optique n'intégrant pas seulement l'organisation du texte narratif comme cela l'a beaucoup été (Berman et Slobin, 1994 ; Hickmann, 2002, 2004) mais également d'autres types de textes, tels que l'expositif. En effet, nous verrons à la Section 2.4 (p. 70), dans quelles mesures les types de textes peuvent être différents et avoir des exigences variées.

La structuration du texte fait appel aux notions de *cohésion* et de *cohérence*, distinction ayant émergé des années 70-80. Définir ces termes n'a pas été chose aisée. Ces deux concepts ont posé problème dans la mesure où la littérature regorge de positions et de définitions différentes (Hickmann, 1995, 2003 ; Salles, 2006) et où ces deux notions ont été amalgamées (Hickmann, 1995, 2003). La section qui suit cherche à définir ces concepts.

La *cohésion* renvoie aux outils linguistiques et à leur organisation (Bosch, 1985 ; Charolles, 1988 ; Charolles et Ehrlich, 1991 ; Favart, 2005 ; Halliday et Hasan, 1976/1989 ; Hickmann, 1995, 2000, 2002, 2003, 2004 ; Karmiloff-Smith, 1993/1995 ; Reinhart, 1980 ; Zammuner, 1986) et « *concerne les manifestations grammaticales exprimant les mises en relation relevant de la cohérence* » (Favart, 2005:306). Halliday et Hasan (1976/1989) définissent la *cohésion* comme le phénomène impliqué lorsque l'interprétation d'un élément du discours dépend de celle d'un autre si bien que le présupposant et le présupposé sont potentiellement intégrés dans un texte. Ces marques de cohésion ont « *des qualités fonctionnelles dans la structuration du texte* » (Favart, 2005:306). Pour réguler le flux de l'information et réussir un message, les locuteurs/scripteurs sont face à de nombreuses contraintes. Premièrement, ils doivent assurer l'interprétation du discours, à savoir le statut de l'information, en gérant les paramètres personnels et spatio-temporels (Favart, 2005 ; Hickmann, 2000, 2004). Les locuteurs/scripteurs doivent appréhender les statuts informationnels des référents (ancien/nouveau) pour utiliser les outils cohésifs adaptés ; ainsi pour une information connue, les individus ne font pas appel aux mêmes outils linguistiques que pour introduire un nouveau référent. Deuxièmement, ils doivent prendre en considération les connaissances des destinataires du message (c'est-à-dire procéder à une évaluation des connaissances mutuelles) (Hickmann, 2000, 2004). Enfin, troisièmement, ils doivent différencier les informations

essentielles des informations secondaires (*foreground* versus *background*) (Favart, 2005 ; Hickmann, 2000, 2004). Notre travail de recherche se penche principalement sur le premier principe en se focalisant sur le statut informationnel des SNL et des pronoms.

Si la cohésion concerne les marques de relation entre les énoncés, la cohérence a plus « *à voir avec l'interprétabilité des textes* » (Charolles, 1988:53 ; Charolles et Ehrlich, 1991) et elle est en fort lien avec les représentations cognitives associées à la structuration des textes (Favart, 2005). Lors de la production d'un texte, le locuteur/scripteur fait référence à des situations, des personnages, des lieux, *etc.* qu'il doit encoder de manière à permettre au destinataire de se créer une représentation mentale du message – un modèle mental (Johnson-Laird, 1980, 1983), un modèle de situation (Van Dijk et Kintsch, 1983) ou encore un scénario (Sanford et Garrod, 1981a, 1981b) – (Fayol, 1997). En effet, « *comprendre un récit, c'est construire un modèle mental de la situation décrite* » (Fayol, 2002:184, selon Bower et Morrow, 1990). Le locuteur/scripteur sélectionne des informations selon l'objectif de son message et les organise selon le type de production qu'il a à réaliser. Ce modèle mental a la caractéristique d'être multidimensionnel (Fayol, 2002). Lorsque le locuteur/scripteur souhaite émettre un message, il doit passer d'un modèle mental multidimensionnel à un message linéaire. Pour cela, il a recours, entre autres, aux outils cohésifs que nous présentons dans la partie suivante.

1.1.2. Les marques de cohésion

Halliday et Hasan (1976/1989), qui sont des précurseurs dans les travaux sur la cohésion et les marques cohésives, proposent de distinguer cinq catégories d'outils cohésifs : la référence, la substitution, l'ellipse, la conjonction et la cohésion lexicale ; sont à ajouter les signes de ponctuation à l'écrit¹³ ou l'intonation à l'oral. Parmi ces outils cohésifs, nous nous sommes essentiellement intéressée à ceux de la référence et de la cohésion lexicale.

La *référence* est l'outil cohésif permettant de lier deux éléments linguistiques évoquant une même entité. Ainsi, dans (3), le pronom personnel *ils* réfère au syntagme nominal lexical, *des enfants*, référent introduit précédemment.

(3) *Il y a des enfants dans la cours de récréation.*

Ils jouent au chat et à la souris.

¹³ Nous ne nous penchons pas plus longuement sur la ponctuation ou sur l'intonation dans la mesure où nous ne nous concentrons pas sur ces aspects dans les analyses présentées dans ce travail de recherche. Pour une étude de la ponctuation dans l'optique d'observer sa contribution à la cohérence, se reporter, entre autres, à Fayol (1989, 1997) et Fayol et Abdi (1988, 1990).

Trois types de référence existent : la référence personnelle, démonstrative et comparative (Halliday et Hasan, 1976/1989). La référence personnelle renvoie aux marqueurs de personne (comme les pronoms personnels et possessifs). La référence démonstrative regroupe les marques démonstratives (comme les pronoms ou déterminants démonstratifs). La référence comparative met en jeu des formes impliquant un sémantisme de comparaison (adjectifs et adverbes de comparaison). Sans entrer dans de telles distinctions, la référence pronominale a été un des centres de préoccupations de ce travail de recherche.

Le maintien de référent peut se réaliser soit sous une forme pronominale soit sous une forme lexicale. Le maintien d'une information par le biais d'un SNL correspond à la *cohésion lexicale*.

Deux types de *cohésion lexicale* sont distingués: la collocation et la répétition (Halliday et Hasan, 1976/1989). La répétition est le phénomène de cohésion lexicale nous intéressant. Il s'agit du maintien d'un référent sous forme lexicale qui implique soit (a) la répétition (4),¹⁴ soit (b) la synonymie (4), soit (c) l'hyponymie (5).

(4) *Il y avait un gros **champignon** poussant près d'elle, arrivant à peu près à sa hauteur [...]. Elle se mit sur la pointe des pieds et jeta un coup d'œil par-dessus le **champignon**.*

(5) *J'ai lu le **livre** de Marc Levy. **Le bouquin** est très sympa !*

(6) *J'ai pu caresser des **tigres**. Ces **félins** sont très impressionnants.*

Si nous nous sommes intéressée à la cohésion lexicale et plus particulièrement à la répétition, nous n'avons cependant pas réalisé d'analyse quantitative concernant les différents types de répétition. Ce type de maintien, présent dans nos données, a pour autant un nombre d'utilisations trop peu important pour permettre des analyses statistiques.

1.1.3. Règles cohésives et principes de cohérence

Dans l'exemple (7) qui suit, l'enfant introduit la nouvelle information par le biais d'une forme lexicale. Dans (7), *la violence*, puis *la triche* et enfin *le vol*, sont trois nouveaux référents introduits mais ni maintenus ni développés.

(7) ***La violence** n'est pas bien du tout. **La triche** en classe ne sert à rien. **Le vol** est très mal.*

Ce message est-il satisfaisant ? L'individu renouvelle certes l'apport de nouvelle information dans son texte mais il ne maintient pas ces référents.

¹⁴ Exemple inspiré de Halliday et Hasan (1976/1989:278).

En effet, un message doit comporter deux types d'informations : de nouveaux et d'anciens référents (Clark et Haviland, 1974, 1977 ; Khorounjaia et Tolchinsky, 2004) qui ont des dispositions préférentielles dans la clause. Clark et Haviland (1974, 1977) proposent la terminologie de *given/new contract*. L'ancienne information apparaîtrait avant la nouvelle information dans la clause (Clark et Haviland, 1974, 1977 ; Branigan, McLean et Reeve, 2003). Deux principales explications à ce principe sont proposées (Branigan *et al.*, 2003). Premièrement, le locuteur respecterait cet ordre (ancienne/nouvelle dans la clause) car il serait conscient qu'il facilite la transmission de l'information ; ainsi par souci de collaboration, il tendrait à concrétiser le *given-new contract*. Deuxièmement, cet ordre serait en lien avec le degré d'accessibilité de l'information. Ainsi, au plus une information serait accessible, au plus elle serait destinée à être en position sujet.

Les locuteurs/scripteurs doivent alors effectuer un *travail de conservation/progression* (Schneuwly, 1988:41) des syntagmes nominaux pour produire un texte cohérent, ce qui implique deux analyses : (a) celle de l'anaphore et (b) celle de l'organisation de l'information. Le premier aspect suppose trois phénomènes : l'introduction de nouveaux référents, la reprise de ces référents et enfin la différenciation de ces référents. Le second aspect concerne l'organisation de l'information, à savoir le statut informationnel.

Ainsi, produire un message cohérent ne résulte pas seulement de l'introduction de nouveaux éléments. Un locuteur expert introduit une nouvelle information, la développe en la maintenant, fait des prédications puis passe à une nouvelle information, qu'il développe, *etc.*. Le fait de maintenir dans le but de développer son texte est une aptitude à acquérir. Dans tout texte, le locuteur doit jongler entre apport de nouvelles informations et maintien d'informations connues à partir desquelles des prédications sont faites.

La cohérence textuelle est basée que quatre règles principales : la *méta-règle de répétition* (Bellert, 1970 ; Charolles, 1978 ; Reinhart, 1980), la *méta-règle de progression* (Charolles, 1978), la *méta-règle de non-contradiction* (Charolles, 1978 ; Reinhart, 1980) et la *méta-règle de relation* (Charolles, 1978 ; Reinhart, 1980). Ces règles ont une visée micro/macrostructurelle et relèvent à la fois du domaine de la linguistique et de celui de la pragmatique.

Pour qu'un texte soit cohérent, la *méta-règle de répétition* doit être respectée. Le texte doit comporter « *dans son développement linéaire des éléments à récurrence stricte* » (Charolles, 1978:14). Afin de respecter cette première règle, la langue comporte différents moyens, divers

outils cohésifs, vus précédemment. La pronominalisation peut permettre la répétition comme l'illustre l'exemple qui suit (8) (de Charolles, 1978:15).

(8) *Une vieille femme a été assassinée la semaine dernière à Besançon. Elle a été retrouvée étranglée dans sa baignoire.*

La *métra-règle de progression* implique que pour assurer la cohérence d'un texte « *il faut que son développement s'accompagne d'un apport sémantique constamment renouvelé* » (Charolles, 1978:20). Ainsi la première règle ne se suffit pas à elle-même ; un texte ne concernant qu'un seul référent sans cesse répété avec des prédications de même sémantisme ne remplit pas les conditions d'un texte cohérent (9).

(9) *Les veuves ne reçoivent que la moitié de la retraite de leur feu mari. Les femmes non mariées perçoivent une pension égale à la moitié de celle que percevait leur mari défunt. Elles n'ont que cinquante pour cent des indemnités que touchait leur mari quand il était vivant. Du temps qu'il était en retraite les épouses des retraités partageaient avec leur mari la totalité de leur pension.*

Dans cet exemple, non seulement un seul référent est introduit, mais en plus les informations le concernant sont toujours le même. Produire un texte cohérent exige alors un équilibre entre « *une continuité thématique et (une) progression sémantique (...) une telle performance exige donc que soient conjointement maîtrisées les métra-règles de répétition et de progression* » (Charolles, 1978:21).

Néanmoins, l'introduction de nouveaux référents ne peut se faire aléatoirement, ce qui implique la nécessité de la troisième métra-règle, celle de *non-contradiction*. En effet, pour que le texte produit soit cohérent, « *il faut que son développement n'introduise aucun élément sémantique contredisant un contenu posé ou présupposé par une occurrence antérieure ou déductible de celle-ci par inférence.* » (Charolles, 1978:22), comme l'illustrent les exemples qui suivent (empruntés à Charolles, 1978:24).

(10) *Ma tante est veuve. Son mari collectionne les machines à coudre.*

(11) *Pierre n'a pas de voiture. Il vend la sienne pour en acheter une neuve.*

Pour finir, une dernière règle est énoncée : la *métra-règle de relation*, introduisant l'idée que pour avoir un texte cohérent, « *il faut que les faits qu'ils dénotent dans le monde représenté soient reliés* » (Charolles, 1978:31). Charolles (1978:31) schématise en disant qu'« *on dira que, dans un monde représenté « M », deux états de choses « p » et « q », sont congruents si*

et seulement si « p » est pertinent pour « q » ». Un énoncé comme (12) ne respecte pas cette règle.

(12) Marie a acheté une armoire Louis XV. Elle a des migraines épouvantables.

Introduire, maintenir, développer, ne pas se contredire, *etc.* : le locuteur/scripteur, pour réaliser ce travail, doit également déterminer et prendre en considération les connaissances que son (ses) destinataire(s) possède(nt), et *vice versa*. Dans l'optique du langage en tant qu'action partagée (Clark, 1996), pour assurer la cohérence et pour qu'il y ait intercompréhension entre un locuteur/scripteur et son (ses) destinataire(s), les individus doivent partager des connaissances communes. C'est ce que nous appelons le *common ground* (Clark, 1996 ; Clark et Murphy, 1982 ; Clark et Schaefer, 1989 ; Clark et Wilkes-Gibbs, 1986 ; Schober et Clark, 1989).¹⁵ Un contrat implicite est établi entre le locuteur/scripteur et les interlocuteurs ; la communication implique un effort coopératif constant entre les participants à une conversation (Haviland et Clark, 1974). Le locuteur/scripteur doit énoncer de nouvelles informations que le destinataire extrait du message et intègre dans sa mémoire comme étant des informations anciennes. Le locuteur/scripteur peut alors les considérer comme telles et utiliser les formes linguistiques relatives à ce statut d'ancienne information. Selon l'estimation du locuteur/scripteur du degré de connaissances communes, les formes linguistiques sont différentes. Ainsi, avant et pendant la conception de son message, le destinataire se doit d'identifier son interlocuteur (*visée vers l'auditoire*, Clark et Murphy, 1982 ; Clark et Wilkes-Gibbs, 1986).¹⁶ Les auteurs proposent trois sources principales permettant au locuteur/scripteur d'un message d'évaluer ce *common ground* : la coprésence physique (la connaissance partagée de l'environnement et du contexte) ; la coprésence linguistique (la connaissance partagée des messages émis préalablement) ; l'appartenance à une même communauté (une connaissance commune d'un ensemble de savoirs).

Clark, Traxler et Gernsbacher (1995) émettent l'idée que la communication orale propose un cadre plus collaboratif que la communication écrite. À l'oral, les locuteurs peuvent signaler un manque de connaissance partagée (Hickmann, Kail et Roland, 1995) et agir en conséquence. Plus les locuteurs collaborent et possèdent un haut degré de *common ground*,

¹⁵Notion introduite pour la première fois par Stalnaker (1978).

¹⁶De la traduction « *audience design* ». (Clark et Murphy, 1982 ; Clark et Wilkes-Gibbs, 1986).

plus la communication est facilitée et, donc, plus la cohérence est assurée (Traxler et Gernsbacher, 1995).

1.1.4. La cohérence et la cohésion : rétrospection en développement

La notion de *cohérence* implique un phénomène cognitif permettant aux destinataires d'un message de construire l'image mentale que le locuteur/scripteur souhaite faire passer (Gernsbacher et Givón, 1995 ; Givón, 1995 ; Traxler et Gernsbacher, 1995). Ainsi, les locuteurs/scripteurs, que ce soit lors d'une tâche de production ou d'une tâche de compréhension, à l'écrite ou à l'orale, collaborent avec les destinataires dans le but spécifique d'être cohérents en négociant les connaissances communes (Gernsbacher et Givón, 1995). Pour réaliser un texte et assurer la bonne construction de la représentation mentale du destinataire, les locuteurs/scripteurs disposent d'outils linguistiques telles les formes pronominales ou lexicales (Charolles, 1978 ; Givón, 1995 ; Halliday et Hasan, 1976/1989). Ces outils syntaxiques assurent une adéquation entre la représentation mentale du locuteur/scripteur et le texte, et permettent alors au(x) destinataire(s) du message d'accéder à cette représentation mentale (Givón, 1995).

Assurer une telle cohérence triangulaire (représentations mentales du destinataire / texte / représentations mentales du locuteur) se développe avec l'âge (Traxler et Gernsbacher, 1995). Les enfants apprendraient à remédier à l'*égocentrisme cognitif* (Hickmann, 2003, selon Piaget, 1923) et à concevoir que les destinataires de leur message aient des représentations mentales différentes et non partagées. L'appréciation et la prise en considération du non-partage de la référence avec un destinataire ne seraient effectives qu'à partir de neuf ans et continueraient d'évoluer au moins jusqu'à onze ans (Favart et Passerault, 1996 ; Hickmann, Kail et Roland, 1992). Les enfants doivent apprendre à décontextualiser leurs représentations de leur perception immédiate et à les recontextualiser (Hickmann, 2003). Au cours de leur développement, les jeunes locuteurs acquièrent « *la capacité à ancrer la parole dans le co-texte et donc à se servir du langage pour contextualiser le langage* » (Hickmann, 1987:239). Ce travail est facilité en production orale car il y a un partage de certaines connaissances, en revanche, l'écrit complique la situation : l'enfant n'a pas de destinataire direct et les connaissances partagées sont moindres.

La cohésion, et donc indirectement la cohérence, d'un texte est atteinte lorsque les utilisations et les représentations des outils linguistiques sont interdépendantes (Chafe, 1976 ; Halliday et Hasan, 1976/1989 ; Hickmann, 2003). Dès 6 ans, les enfants entrent dans une période durant laquelle ils apprennent à réguler le flux de l'information dans une vision de niveau supérieur à la phrase, à savoir dans une perspective textuelle. L'organisation discursive « *constitue l'une des facettes les plus importantes du langage chez l'enfant après six ans* » (Hickmann, 2002:181). Ils doivent alors gérer aussi bien le niveau phrastique (relations syntaxiques à l'intérieur d'une clause) que le niveau discursif (organisation des énoncés avec le contexte). L'enfant doit mettre en relation les formes du langage avec les différentes fonctions qu'elles peuvent avoir tout en faisant cas du contexte de production. Durant cette période, les enfants apprennent à maîtriser un système complexe. Ce système doit mettre en relation les formes et les fonctions des unités linguistiques tout en prenant en considération le contexte et en évaluant la connaissance mutuelle. Les enfants doivent alors intégrer deux types de connaissances : celles relatives aux règles du contenu d'un message (soit la cohérence) et celles relatives aux règles gouvernant le flux de l'information entre les clauses d'une production (soit la cohésion). Nous avons vu, dans la Section 1.1.1 (p. 22), que le locuteur/scripteur, pour réguler l'information dans un discours, doit gérer trois aspects : (a) le statut informationnel des éléments ; (b) les connaissances des destinataires ; (c) la différenciation des informations essentielles et secondaires. Un des outils cohésifs permettant la gestion de ces trois points est la référence, et plus précisément l'anaphore.

L'utilisation et la gestion du système référentiel est une acquisition tardive (Hickmann, 2003). Les enfants sont capables, en compréhension, de distinguer la continuité et la discontinuité référentielle d'une production assez tôt (Hickmann et Hendriks, 1999). Cependant, la production des outils relatifs à la référence relève d'une complexe interaction entre les fonctions syntaxiques, sémantiques et pragmatiques des formes. Les analyses révèlent que, si les référents sont très clairement établis, alors la cohésion est assez bien gérée dès l'âge de 9/10 ans (Pellegrini, Galda et Rubin, 1984). Néanmoins, en cas de connaissance partagée limitée, la gestion des marques concernant le statut informationnel est entravée (Bartlett, 1984 ; Hickmann *et al.*, 1995 ; Kail et Hickmann, 1992). L'acquisition de la référence, à l'écrit comme à l'oral, se prolongerait au moins jusqu'à 10/11 ans (Favart et Passerault, 1996 ; Hickmann, 2000, 2003 ; Hickmann *et al.*, 1996 ; Hickmann et Hendricks, 1999 ; Karmiloff-Smith, 1993).

L'utilisation des outils cohésifs relatifs à la référence diffère selon l'âge (De Weck, 1991 ; De Weck et Schneuwly, 1994 ; Hickmann, 2000, 2003 ; Hickmann *et al.*, 1996 ; Hickmann et Hendricks, 1999 ; Schneuwly, 1988 ; Schneuwly, Rosat et Dolz, 1989 ; Pellegrini, *et al.*, 1984). La production de ces outils semble également être corrélée au contexte de production. En effet, selon le genre textuel, la densité anaphorique¹⁷ diminue avec l'âge comme dans les textes narratifs (histoire fictive) mais augmente avec l'âge dans les textes explicatifs et argumentatifs (De Weck et Schneuwly, 1994). De plus, les plus jeunes enfants (environ 7 ans) ont des textes plus cohésifs à l'oral qu'à l'écrit (Pellegrini, *et al.* 1984). En revanche, les enfants plus âgés (environ 9 et 11 ans) utilisent plus d'outils cohésifs dans leurs textes écrits que dans leurs textes oraux (Pellegrini, *et al.* 1984).

Des travaux soutiennent l'idée que les différents types de textes ont un schéma spécifique quant à l'utilisation des outils cohésifs (De Weck, 1991 ; De Weck et Schneuwly, 1994 ; Schneuwly, 1988 ; Schneuwly et Bronckart, 1986). Le texte argumentatif, par exemple, se caractériserait comme ayant plus de pronoms personnels que d'anaphores nominales. À l'inverse, la narration fictive aurait une proportion d'anaphores nominales supérieure à celle des pronoms personnels. Aussi, les textes informatifs possèdent un plus grand nombre d'unités anaphoriques que les textes argumentatifs.

Toutes ces études montrent que l'observation de la cohésion référentielle, et notamment pour nous des statuts informationnels, doit se réaliser en intégrant dans le protocole des enfants d'âges différents, divers types de textes et modalités. Nous ne pouvons proposer une conclusion que partielle sur cette thématique si nous ne prenons en considération le contexte de production et l'âge des enfants. Notre étude a l'avantage de pouvoir observer le comportement d'enfants d'âges divers dans des contextes de production spécifiques. Chaque enfant a en effet produit un texte narratif écrit, narratif oral, expositif écrit et expositif oral. Deux types de textes sont alors pris en considération dans les deux modalités du langage. Pour finir, nos analyses vont venir compléter ces études dans la mesure où nous travaillons sur les productions d'enfants plus âgés.

1.1.5. Cohérence/cohésion : conclusion

Cette partie a deux buts principaux. Premièrement, clarifier les définitions des notions de *cohésion* et *cohérence*. Une question émane des études concernant ces notions : quel lien

¹⁷ La densité anaphorique correspond au nombre total d'anaphores pour 100 verbes.

entretiennent la cohésion et la cohérence ? Plusieurs réponses ont été données. Une première perspective considère que la *cohérence* et la *cohésion* fonctionneraient en interdépendance de manière parallèle. Une seconde envisage que la *cohérence* déterminerait la *cohésion*. Une troisième option peut être proposée selon laquelle ces deux concepts seraient sans cesse en interaction (Berman et Slobin, 1994 ; Hickmann, 2003, 2004). La cohérence influence la cohésion comme cette dernière influence la cohérence. Des textes sans outils cohésifs peuvent être compréhensibles comme des textes en contenant à outrance peuvent être incompréhensibles (Charolles et Ehrlich, 1991 ; Salles, 2006). L'utilisation d'outils cohésifs est nécessaire à l'interprétabilité des textes et à la construction de la cohérence. Néanmoins, la cohésion ne constitue pas le seul facteur de textualité (Halliday et Hasan, 1976/1989 ; Salles, 2006). L'exemple qui suit (Charolles, 1978 et repris par Salles, 2006) est un texte très cohésif mais incohérent de par le non respect d'une des règles de cohérence, celle de progression.

(13) Le forgeron est vêtu d'un pantalon noir et d'un chapeau clair et d'une veste grise et marron foncé. Il tient à la main le piquet de la pioche et il tape dessus avec un marteau sur le piquet de la pioche. Les gestes qu'il a faits : il a le piquet de la pioche et avec son marteau il tape sur le piquet de la pioche. Le bout de cet outil qui s'appelle le piquet de la pioche est pointu et l'autre bout est carré. Pour le faire devenir rouge avec le piquet de la pioche, il l'a mis dans le feu et les mains sont rouges.

En effet, dans cet exemple, notons « *la flagrante circularité du propos* » (Charolles, 1978:21). Ce texte est cohérent, si nous nous en tenons à la méta-règle de répétition mais la progression sémantique est limitée.

Faire un bon usage de ces outils cohésifs tout en respectant les règles de cohérence est un bon compromis ; ceci souligne « l'étroite solidarité entre la cohésion et la cohérence et c'est sûrement, aussi, laisser entrevoir le rôle primordial de la cohérence – et, plus précisément, des relations de cohérence – pour la cohésion, pour le choix des expressions référentielles et pour leur interprétation » (Salles, 2006:communication en ligne).

Deuxièmement, il est utile de souligner l'importance de la gestion des statuts informationnels dans la capacité de produire un texte cohérent. Introduire, maintenir et distinguer les différents statuts référentiels des éléments d'un énoncé est primordial (Clark et Haviland, 1974, 1977 ; Favart, 2005 ; Hickmann, 2000, 2004 ; Khorounjaia et Tolchinsky, 2004 ; Schnewly, 1988). Les locuteurs/scripteurs doivent également réaliser que le destinataire de la production a certainement des représentations mentales différentes de celles de l'auteur. La prise en

considération de ce paramètre lors de l'émission du message se voit à travers la gestion du statut informationnel des référents (Section 1.2, p. 32) et la manipulation du syntagme nominal (SN) comme marque de cohésion (Section 1.3, p. 40), aspects centraux de ce travail de thèse. La Section 1.2 qui suit se penche alors sur le statut informationnel.

1.2. Le statut informationnel

1.2.1. Nouvelle et ancienne information : définitions

Produire un message implique l'énonciation de diverses informations qui peuvent être nouvelles ou anciennes (Khorounjaia et Tolchinsky, 2004). Tout acte communicationnel exige des locuteurs la capacité de manipuler ces deux types d'information. La nouvelle information se caractérise comme un référent qui n'a pas été mentionné jusque-là dans le discours et qui n'est pas accessible par l'interlocuteur dans la situation d'énonciation (Allen, Skarabela et Hughes, 2008 ; Chafe, 1994 ; Halliday et Hasan, 1976/1989). Ce référent est alors réalisé par une forme de type SNL (« *high information form* », Allen *et al.*, 2008:108-109).

L'ancienne information correspond à une information retrouvable par l'interlocuteur dans le discours ou dans le contexte (Allen *et al.*, 2008 ; Halliday et Hasan, 1976 ; Chafe, 1994). Le continuum nouvelle/ancienne information implique des choix de structures linguistiques spécifiques. Selon le statut référentiel d'un référent, le locuteur/scripteur fait des choix spécifiques. Levelt place le statut informationnel parmi les cinq aspects majeurs de la « *microplanification* » (Levelt, 1989:144). Un locuteur coopératif¹⁸ marque syntaxiquement l'accessibilité¹⁹ d'un référent ; par exemple, l'individu a tendance à pronominaliser un ancien référent, à savoir un référent accessible pour le destinataire.

Entre la nouvelle et l'ancienne information, nous pouvons avoir un statut informationnel intermédiaire : l'information dite accessible (Chafe, 1976, 1994 ; Du Bois, 1987) ou inferrable (Prince, 1981, 1992). Il s'agit d'un référent activé et rendu accessible alors qu'il était précédemment semi-activé (Chafe, 1976, 1994) que nous pouvons retrouver également sous la dénomination d'*information accessible dans le discours* (Levelt, 1989). Levelt (1989) va au-delà en proposant quatre degrés d'accessibilité d'un référent : il peut être inaccessible,

¹⁸ De la traduction de l'expression de Levelt (1989:99) « *cooperative speaker* ».

¹⁹ Dans le sens de "givenness" (Levelt, 1989:99).

accessible, accessible dans le discours et accessible en tant que focus (Levelt, 1989:145). Ces informations sont schématisées dans la Figure 2.

Figure 2 : Les statuts d'accessibilité des référents (Levelt, 1989:145)



L'information inaccessible est l'équivalent de la nouvelle information. L'information accessible mais hors du discours est une information inférée par le biais de l'information générale et qui se base donc sur les connaissances communes d'un groupe de pairs. L'information accessible à partir du discours représente un référent accessible par le biais d'un référent déjà introduit précédemment et qui lui est étroitement lié. Enfin, l'information « *in focus* » correspond au référent prototypiquement ancien et activé à un moment donné.

Pour introduire et maintenir une information, les locuteurs/scripteurs ont recours à des outils linguistiques et cohésifs. Selon le statut informationnel d'une information, cette dernière est encodée préférentiellement sous différentes formes et distribuée différemment dans la clause (entre autres, Ariel, 1990, 1996 ; Khorounjaia et Tolchinsky, 2004 ; Levelt, 1989 ; Sanford et Garrod, 1981a, 1981b). Le statut informationnel aurait un fort impact sur les formes linguistiques choisies (Allen *et al.*, 2008). L'encodage d'un référent sous forme de SNL, qui est de faible accessibilité (« *low accessibility* » Khorounjaia et Tolchinsky, 2004:88, notion d'Ariel, 1990, 1996),²⁰ est utilisé pour introduire une nouvelle information (Allen *et al.*, 2008). En revanche, les pronoms, qui sont de haute accessibilité (« *high accessibility* », Khorounjaia et Tolchinsky, 2004:88), sont utilisés pour encoder des référents actifs donc anciens (Allen *et al.*, 2008). Comme le dit Fox (1987), en parlant du *pattern* basique de l'anaphore dans les textes expositifs écrits, un pronom est utilisé afin de référer à une

²⁰ The accessibility marking scale of Ariel (1990, 1996:21) : « zero < reflexives < agreement markers < cliticized pronouns < unstressed pronouns < stressed pronouns < gesture < proximal demonstrative (+NP) < distal demonstrative (+NP) < proximal demonstrative (+NP) + modifier < distal demonstrative (+NP) + modifier < first name < last name < short definite description < long definite description < full name < full name + modifier »

information mentionnée précédemment, dans les autres cas un SNL est préféré. L'anaphore, constituant un des outils assurant la cohésion d'un texte, est un aspect syntaxique dont nous nous sommes préoccupée en observant notamment les pronoms et SNL maintenant une information. Elle permet la continuité référentielle entre les clauses d'un texte, aspect essentiel pour assurer des représentations mentales cohérentes (Seigneuric et Megherbi, 2008). Le fait que le pronom encode préférentiellement l'ancienne information et le SNL une nouvelle information semble être un principe universel : il s'agit du traitement naturel de l'information.

Des principes universaux semblent régenter les langues. Le premier principe universel est celui selon lequel l'ancienne information se situerait préférentiellement au début d'une clause alors que la nouvelle information se situerait plutôt en fin de proposition (Du Bois, 1987 ; Hickmann, 1995 ; Jisa, 2000 ; Khorounjaia et Tolchinsky, 2004 ; Lambrecht, 1994). Le choix d'introduire la nouvelle information en position non sujet est gouverné par des contraintes discursives pesant sur les choix linguistiques (Du Bois, 1987, 2003 ; Khorounjaia et Tolchinsky, 2004). C'est dans cette optique qu'une des parties de nos analyses porte sur la *Preferred Argument structure* (PAS)²¹ de Du Bois (1987), qui intègre beaucoup de concepts abordés précédemment tels que les statuts référentiels mais également les positions syntaxiques. La PAS, comme nous le verrons plus tard, reflète les contraintes de la transmission de l'information de la langue orale. À cet universel une seconde tendance, la contrainte du sujet léger (*the light subject constraint*), s'ajoute (Chafe, 1994). Les locuteurs/scripteurs en production combleraient la position syntaxique sujet avec de l'ancienne information qui serait de plus encodée, le plus souvent, sous la forme d'un pronom.²²

La décision d'encoder un référent sous une forme particulière dans une position syntaxique spécifique relève de facteurs syntaxiques, pragmatiques, discursifs mais également cognitifs (Khorounjaia et Tolchinsky, 2004). En termes de cognition, la nouvelle information exige plus de ressources cognitives que l'ancienne information dans la mesure où cette dernière est active et donc plus ou moins immédiatement accessible (Chafe, 1994 ; Khorounjaia et Tolchinsky, 2004). C'est ainsi, étant donné le coût cognitif que mobilise l'introduction d'une nouvelle information, qu'une contrainte conversationnelle existe : la contrainte d'une seule

²¹ Le concept de la PAS est présenté dans la Section 1.2..3 (p. 36).

²² La contrainte du sujet léger est définie dans la Section 1.3 (p. 40).

nouvelle idée par clause (*the one new idea constraint*, Chafe, 1994:119). Les locuteurs/scripteurs ne transmettraient au maximum qu'une nouvelle information par clause (Du Bois, 1987, 2003 ; Chafe, 1994 ; Khorounjaia et Tolchinsky, 2004).

1.2.2. Statut informationnel : le cas du français

En français, la phrase s'analyse traditionnellement en deux parties : le thème et le propos. Dans une situation canonique et en règle générale, « *l'ordre linéaire de la phrase reflète l'ordre de l'information* » (Riegel *et al.*, 1994/2002:606). Le thème est placé en début de phrase alors que le propos se trouve en position post verbale. Le focus peut être porté sur le thème ou sur le propos. Par défaut, en français, le focus est porté sur le propos et se trouve donc en position post verbale. Cependant, certaines structures permettent d'inverser cet ordre et d'exprimer le focus en position autre que post verbale. Ceci implique une réorganisation de la phrase par des constructions spécifiques telles que la structure clivée ou la dislocation. Le français a alors recours à un réaménagement syntaxique des SN.

Pour ce travail de thèse, nous avons décidé d'analyser les textes dans une perspective liée à celle du thème/propos, celle concernant la dichotomie nouvelle/ancienne information. En français oral, les nouveaux référents apparaissent préférentiellement par le biais d'une structure clivée (14) (Gayraud *et al.*, 2001 ; Jisa, 2000a et 2000b) et dans une position post-verbale (15) (Jisa, 2000a, 2000b ; Lambrecht, 1994, Hickmann, Hendriks, Roland et Liang, 1996). Une fois introduit, ce nouveau référent peut être maintenu dans la clause. Le français parlé met à disposition quatre options pour réaliser ce maintien de l'information : la répétition du nom (16), la répétition du nom accompagné d'un pronom sujet coréférentiel (17), par un pronom zéro (ellipse du sujet) (18), l'utilisation d'un pronom anaphorique (19) ou relatif (20) (Gayraud *et al.*, 2001 ; Jisa, 2000a).

(14) **C'est un homme** que j'ai vu.²³

(15) Il y a **un homme**. / Hier j'ai vu **un homme**.

(16) Et **l'homme** m'a donné un bonbon.

(17) Et **l'homme il** m'a donné un bonbon.

(18) Et **Ø** m'a donné un bonbon.

(19) Et **il** m'a donné un bonbon.

²³ Les exemples de (14) à (20) sont de Jisa (2000).

(20) *Qui m'a donné un bonbon.*

Le français est considéré comme une langue ayant un ordre des mots de base SVO (Sujet-Verbe-Objet) (Hawkins, 1983). Dans ce type de langue, la position sujet est un indice important concernant la planification puisque cette position sert de point de départ à l'organisation phrastique (Gernsbacher, 1990 ; Gernsbacher et Hargreaves, 1988). Le début de la phrase est considéré comme le guide de l'interprétation de l'information et en français ce début de phrase se confond avec la position sujet (Gernsbacher, 1990 ; Gernsbacher et Hargreaves, 1988). Nous parlons alors de *point de départ* (*starting point*, Chafe, 1994:82-107) qui permet l'encodage d'une information déjà connue et préférentiellement sous une forme exprimant le haut degré d'accessibilité, à savoir le pronom. Dans une étude portant sur des conversations, Chafe (1994) montre que 81% des sujets maintiennent une information et l'ancienne information est encodée dans 98% des cas par un pronom. Le français parlé conforte ce pattern en préférant l'utilisation de clitiques en position sujet. L'utilisation de SNL en position de sujet pourrait donc être un indice important pour observer le développement du français écrit.

1.2.3. Statut informationnel, le français et la Preferred argument structure

La grande majorité des études développementales préalables sur le SN se centrait principalement sur l'adéquation référentielle (par ex. chaînes anaphoriques) (Bamberg 1986, 1987 ; Kail et Hickmann 1992 ; Karmiloff-Smith 1981 ; Wigglesworth 1991). Pour cette étude, nous nous basons aussi bien sur le statut informationnel des SN que sur leurs distributions, leurs organisations syntaxiques et leur constitution lexicale dans une optique intégrant syntaxe et pragmatique. C'est dans cette mesure que nous basons notre étude sur la notion de PAS de Du Bois (1987) qui prend en considération les différentes sphères de la linguistique, la syntaxe, la sémantique et la pragmatique. En travaillant sur le sacapultec Du Bois (1987) oriente ses travaux autour d'une question principale : parmi les différentes structures argumentales offertes par la langue, existe-t-il des structures que les locuteurs préfèrent ? À travers la PAS, le linguiste cherche à répondre à cette question et à explorer la connexion existant entre les formes et les fonctions. La PAS reflèterait les contraintes de la transmission de l'information et notamment de la langue orale. La transmission de l'information façonnerait le langage, la syntaxe émergerait du discours. De nombreuses

langues ont été observées : par exemple, Du Bois (1987) travaillant sur le sacapultec (première présentation en 1981) et les travaux de nombreux linguistes dont Du Bois énumère les études (1987).²⁴

Du Bois (1987) émet l'hypothèse que la PAS apparaîtrait comme un principe universel. La PAS correspond alors à une convergence entre le choix lexical, la forme référentielle et le rôle grammatical (Clancy, 2003). La PAS reprend des concepts essentiels du flux informationnel du discours tels que ceux abordés par Chafe (1994) (par exemple, *the one new idea per clause* ou le *light subject constraint*), Charolles (1978) ou encore Hickmann (1995) (par exemple, la nouvelle information en fin de phrase/l'ancienne information en début de phrase).

La PAS présente quatre contraintes (Du Bois, 1987 ; Du Bois, Kumpf et Ashby, 2003a et 2003b). La première contrainte A, la contrainte de l'argument unique (*the one lexical argument constraint*) sous-entend qu'**une clause est constituée au maximum d'un élément lexical** (21).²⁵

(21) a. *Il a frappé le chien* versus b. *L'homme a frappé le chien*

Les locuteurs préfèrent dire la version a. de (21) que la version b., cette dernière possibilité contenant plus d'un élément lexical.

Selon la contrainte B, la contrainte du sujet transitif non lexical (*the non-lexical A constraint*) **cet élément lexical n'apparaît pas en position sujet d'un verbe transitif** (22).

(22) a. *Il a frappé la femme* versus b. *L'homme a frappé la femme*.

Les locuteurs penchent également plus pour la version a. de (22) que pour la version b., étant donné que, dans la seconde proposition, un élément lexical se trouve en position sujet.

La contrainte C, la contrainte de l'argument unique introduisant une nouvelle information (*the one new argument constraint*), laisse entendre qu'**une clause est constituée d'un seul nouvel élément** (23).

(23) a. *Il a frappé une femme* versus b. *Un homme a frappé une femme*.

Les locuteurs encodent plus un message du type de (23) que (23); cette deuxième possibilité contenant plus d'un nouvel élément (contrainte C).

²⁴ Les études des langues maya mam (England, 1986) et rama (Craig, 1987), la langue australienne chamorro (Scancarrelli, 1985), le malay (Hopper, 1987), l'acehnese (Durie, 1985), le quetchua (Stewart, 1984), le papago (Payne, 1987), l'hébreu (Smith, 1987 et 1996), l'anglais (Iwasaki, 1985), l'allemand (Schuetze-Coburn, 1987), le portugais brésilien (Dutra, 1987), le japonais (Downing, 1985), le français (Lambrecht, 1987), *etc.*, montrent la pertinence de cette PAS (Du Bois, 1987:838-839).

²⁵ Les versions a. des exemples (21) à (24) sont, selon les contraintes de la PAS, les possibilités les plus naturelles et les moins coûteuses.

Selon la contrainte, D, la contrainte de l'ancienne information en position sujet transitif (*the given A constraint*), **l'ancienne information se trouve en position sujet de verbe transitif** (24).

(24) a. **Il** a frappé une femme versus b. L'homme a frappé une femme.

Produire la version a. de (24) au lieu de la version b. semble plus évident, étant donné que, dans la version a., la position sujet est réservée à l'ancienne information (sous forme pronominale).

La PAS présente donc quatre contraintes parmi lesquelles deux concernent la forme et deux autres la pragmatique (statut référentiel) (Du Bois, 1987 ; Du Bois, Kumpf et Ashby, 2003a, 2003b). De plus, deux des contraintes sont également en lien avec des notions syntaxiques. Ces quatre contraintes et leur interaction avec la forme, la pragmatique et la syntaxe sont présentées dans le Tableau 1.

[Tableau 1 : Les contraintes de la Preferred Argument Structure]

Tableau 1 : Les contraintes de la Preferred Argument Structure

Contraintes sur la forme	Contraintes pragmatiques
Un seul élément lexical par clause (contrainte A)	Un seul nouvel élément par clause (contrainte C)
Contraintes syntaxiques	
L'élément lexical en position non sujet (contrainte B)	L'ancienne information en position sujet (contrainte D)

1.2.4. La Preferred Argument Structure : rétrospection en développement

En 2003, Du Bois *et al.* font paraître un ouvrage regroupant des travaux prenant la PAS comme principal outil d'étude ; ces travaux se diversifient. Ainsi, le principe de la PAS considéré comme quasi universel est le centre d'intérêt : (a) d'études syntaxiques (Arnold, 2003 ; Corston-Oliver, 2003 ; Du Bois, 2003 ; Durie, 2003 ; England et Martin, 2003 ; Genetti et Crain, 2003 ; Hofling, 2003 ; Martin, 2003) ; (b) de travaux diachroniques (Ashby et Bentivoglio, 2003) ; (c) d'études en pathologie langagière (Kohn et Cragolino, 2003 ; Weber, 2003) ; (d) ou encore de recherches orientées en interaction et en acquisition (Allen et Schröder, 2003 ; Clancy, 2003) et en interaction pédagogique (Kumpf, 2003).

Dans un récent article, Allen *et al.* (2008) rapportent des études en acquisition précoce se concentrant sur l'accessibilité des référents et sur la réalisation argumentale et pour certaines sur la PAS. Sont cités des travaux portant sur des données spontanées dont ceux de Allen (1997, 2000, 2007), Allen et Schröder (2003) sur l'inuktitut, de Cho (2004) et de Clancy (1993, 1997, 2003) sur le coréen, de Guerreiro, Oshima-Takane et Kuriyama (2006) sur le japonais et l'anglais, de Hughes et Allen (2006) sur l'anglais, Mishina-Mori (2007) sur le japonais et le bilinguisme japonais/anglais, de Narasimhan, Budwig et Murty (2005) sur le hindi, de Paradis et Navarro (2003) sur l'espagnol, l'anglais et le bilinguisme espagnol/anglais, de Serratrice (2005) sur l'italien, de Serratrice, Sorace et Paoli (2004) sur l'italien, l'anglais et le bilinguisme italien/anglais, de Skarabela (2006, 2007) et Skarabela et Allen (2002) sur l'inuktitut.²⁶

Toutes ces études suggèrent que les jeunes enfants (d'âges allant de 1;05 à 4;07 ans) utilisent les formes linguistiques du type SNL, par exemple, pour encoder un référent de faible accessibilité et des formes de type pronom pour encoder un référent de haute accessibilité (Allen *et al.*, 2008). Certains aspects syntaxiques dont le choix des formes syntaxiques sont ainsi affectés par la connaissance qu'ont les enfants du flux informationnel du discours à savoir ici le degré d'accessibilité d'un référent (Allen *et al.*, 2008). Cette idée est confirmée par des études expérimentales en production, toujours en acquisition précoce, (Allen *et al.* (2008) citent notamment les travaux de Campbell, Brooks et Tomasello, 2000 ; Matthews, Lieven, Theakston et Tomasello, 2006 ; Wittek et Tomasello, 2005).

Concernant l'acquisition tardive, deux études principales ont retenu notre attention : une d'Ashby et Bentivoglio (1993) et une de Khorounjaia et Tolchinsky (2004). Ashby et Bentivoglio (1993) proposent une étude concernant le français et l'espagnol oral en prenant comme base la PAS. Pour chaque langue, les linguistes ont un corpus constitué de monologues oraux produits par douze locuteurs âgés de 16 à 30 ans. Khorounjaia et Tolchinsky (2004) utilisent l'outil de la PAS dans une étude en acquisition tardive et développementale, orientations similaires à celles de notre travail de thèse. Ont été analysées les productions orales narratives et expositives de locuteurs espagnols de quatre âges : 9/10, 12/13, 15/16 ans et étudiants à l'université. Les résultats de ces deux études révèlent que les locuteurs français (Ashby et Bentivoglio, 1993) et espagnols (Ashby et Bentivoglio, 1993 ; Khorounjaia et

²⁶ Notons l'absence d'étude sur le français.

Tolchinsky, 2004) ont une structure argumentale préférée coïncidant avec les règles énoncées par la PAS de Du Bois.

Ces travaux en acquisition prenant comme objet d'étude la PAS révèlent que, dans la plupart des cas, les individus observés (jeunes enfants, enfants, adolescents ou adultes) ont comme pattern préférentiel la PAS. La PAS apparaît comme un outil linguistique naturellement utilisé à l'oral qu'il n'est pas besoin d'acquérir.

1.2.5. Statut informationnel et Preferred Argument Structure : conclusion

Nous avons vu que, pour assurer la cohérence d'un message, les individus doivent gérer divers principes relatifs à la cohérence et à la cohésion (Section 1.1, p. 22). Un des aspects que l'enfant doit acquérir est le système référentiel de la langue, et donc la gestion et la différenciation de la nouvelle et de l'ancienne information dans les énoncés. Les individus doivent alors, pour faciliter au maximum la compréhension de l'interlocuteur, utiliser à bon escient la syntaxe, domaine auquel nous nous intéressons à travers l'étude du SN. Cette unité est une des marques cohésives des langues : elle permet, en effet, de réaliser la référence et la cohésion lexicale, dans le sens de Halliday et Hasan (1976/1989). La section 1.3 qui suit se consacre ainsi à la présentation du SN.

1.3. Le syntagme nominal

1.3.1. Les catégories du discours

Les différentes unités d'une langue sont organisées en classes syntaxiques. Un classement ayant pour base des critères distributionnels et morphologiques peut être proposé (Creissels, 2006a, 2006b ; Riegel, Pellat et Rioul, 1994/2002 ; Goosse, 1993/2004). Le critère distributionnel correspond aux positions que peut occuper une unité dans une phrase. Le critère morphologique concerne essentiellement la variabilité ou l'invariabilité de l'unité (Goosse, 2004 ; Creissels, 2006a, 2006b). Les mots sont alors classés selon leur distribution dans une proposition et selon leurs possibilités morphologiques.

Ainsi, les classes syntaxiques sont les suivantes : le nom, l'article, l'adjectif, le pronom, le verbe, l'adverbe, la préposition, la conjonction et l'interjection. Ces différentes parties du

discours peuvent être dissociées en deux grandes classes : les mots variables et les mots invariables.

La catégorie des mots invariables regroupe l'adverbe (25), la préposition (26), la conjonction de subordination (27), la conjonction de coordination (28) (Riegel *et al.*, 1994/2002), l'introducteur (29) et le mot-phrase (30) (Goosse, 1993/2004).

(25) *La fille était **très** énervée.*

(26) *J'ai copié sur la feuille **de** mon voisin.*

(27) *J'ai volé le sac **lorsqu'**elle a tourné le dos.*

(28) *Je me suis enfermée dans ma chambre **et** ai pleuré.*

(29) ***Voilà** mon cousin !*

(30) *Bravo !*

La catégorie des mots variables, qui nous intéressent davantage, regroupe le nom, l'adjectif, le déterminant, le pronom et le verbe. Le nom apparaît comme une unité du discours qui peut varier en genre et en nombre. Il est également accompagné d'un déterminant qui porte les marques du genre et du nombre du nom qu'il précède. Le nom peut occuper la fonction de sujet (31), de complément d'objet direct (32), *etc.* L'adjectif varie en genre et en nombre selon le nom auquel il est relié. Il peut apparaître comme épithète (33) et attribut (34). Le déterminant, comme l'adjectif, varie en genre et en nombre selon les caractéristiques que portent le nom auquel il se rapporte (35). Le pronom varie en genre et en nombre mais également, selon le type de pronom, en personne et selon leur fonction dans la phrase (36) ; il peut occuper les mêmes fonctions que le nom. Le verbe varie en personne, en nombre, en genre, en temps, en mode et en voix (37).

(31) ***La bagarre** ce n'est pas bien.*

(32) *J'achète **une** pomme.*

(33) *La **petite** fille s'est ensuite énervée.*

(34) *Elle est **méchante**.*

(35) ***Le** gars m'a ensuite frappé.*

(36) *Les gens regardaient mais **ils** ne sont pas intervenus.*

(37) *Elle **s'est enflammée** et après ils me **frappaient** tous.*

Parmi la catégorie des mots variables, deux types d'unités bien spécifiques nous ont particulièrement intéressée : le nom et le pronom. En effet, ces unités linguistiques entrent dans la construction du SN et dans la gestion du statut informationnel des référents dans un texte. Pour introduire et/ou pour maintenir une information, le locuteur/scripteur fait appel aux noms et aux pronoms.

1.3.2. Les unités linguistiques observées

Cette partie se consacre à la présentation des variables dépendantes de notre travail de thèse c'est-à-dire aux catégories d'analyses observées qui sont susceptibles de varier en fonction des variables indépendantes de notre étude.

Nous avons décidé de travailler sur la catégorie des SN en distinguant les SN pronominaux et les SNL, unités appartenant à la catégorie des mots variables. Le SN se définit comme un groupe d'unités consécutives de dimensions variables et constitue un niveau intermédiaire entre le niveau supérieur, la proposition et le niveau inférieur, les morphèmes. Le SN possède un élément central, qui « *peut être un nom, mais il peut également être constitué par toutes sortes de mots* » (Herslund, 2003:105), et un ou plusieurs éléments qui lui sont subordonnés.

Nous pouvons analyser le SN (a) d'un point de vue externe et également (b) d'un point de vue interne ou distributionnel (Ravid *et al.*, 2002 ; Riegel *et al.*, 1994/2002).

Du point de vue externe, le SN peut être constituant facultatif, constituant du verbe, constituant d'un syntagme prépositionnel, constituant facultatif détaché derrière un autre SN. Nous nous intéressons alors à sa fonction syntaxique.

Du point de vue interne, nous examinons la structure et la décomposition syntaxique du SN. Il peut être minimal (*déterminant+nom*) ou étendu (c'est-à-dire avec l'adjonction de modificateurs). Le SNL peut prendre différentes formes (Riegel *et al.*, 1994/2002). La forme canonique est bien sûr *déterminant+nom* à laquelle nous pouvons ajouter plusieurs modificateurs tels que des groupes prépositionnels, adjectivaux ou des subordonnées relatives (voire complétives dans certains cas). Cependant cette définition peut être assouplie et élargie si nous voulons introduire des éléments ne répondant pas aux critères du nom : les noms propres, les pronoms, les relatives substantives et les groupes infinitifs, qui n'ont pas forcément comme tête la structure prototypique *déterminant+nom* (Leeman-Bouix (1994 ; Riegel *et al.*, 1994/2002). En effet, les noms propres et les pronoms ne sont jamais précédés

de déterminants (38) et (39) ; les infinitifs peuvent l'être mais pas systématiquement comme dans l'exemple (40) ; les relatives substantives ne sont jamais précédées de déterminant (41).

(38) *Charles est sympa.*

(39) *Lui-même l'avoue.*

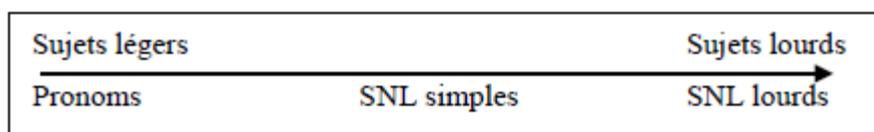
(40) *Se battre est complètement immature.*

(41) *Celui qui casse les verres les paie.* (Riegel et al., 2002 : 486)

Nous retenons la catégorisation de Ravid *et al.* (2002) ou encore de Riegel *et al.* (1994/2002) dans la mesure où elle prend en compte deux aspects intéressants pour notre étude. En effet, du point de vue externe, nous observons la fonction du SN (sujet, objet direct ou autre) et, du point de vue interne, nous avons examiné sa complexité syntaxique (nombre de nœuds lexicaux, nombre de modifieurs, nombre de modifieurs différents).

La considération et la distinction de la fonction syntaxique sont essentielles à la bonne compréhension de la production textuelle, notamment de la production des pronoms et des SNL. C'est dans cette mesure que nous introduisons la notion de *sujet léger* versus *sujet lourd*.²⁷ Les locuteurs/scripteurs combleraient la position syntaxique sujet, le plus souvent, sous la forme d'un pronom, un *sujet léger*, et laisseraient aux positions non sujet *les SN lourds*, à savoir les SNL. Parmi les SNL, certains sont plus lourds que d'autres. Ainsi, par exemple, ceux constitués de modifieurs (*un chat noir courant à travers le champ* versus *un chat*) et/ou de plus d'un nœud lexical qu'un (*Le problème soulevé par la presse* versus *le problème*) sont plus lourds. Nous pouvons alors parler d'un continuum (Figure 3).

Figure 3 : Continuum SN légers/SN lourds



François (1974) dans l'étude d'un long corpus de conversation entre les membres d'une famille appartenant à la classe ouvrière, avait trouvé un total de 1500 noms parmi lesquels seulement 46 (soit 3%) occupaient la position de sujet alors que tous les autres sujets étaient des clitiques (Lambrecht, 1987). Lambrecht (1987) parle même de *preferred clause pattern* en disant qu'il y a une préférence pour un type de construction spécifique soit un sujet

²⁷ Concept à lier à la notion de *light subject constraint* (Chafe, 1994), Cf. ci-dessous.

pronominal et un complément post verbal sous forme nominale (ceci n'est pas sans rappeler la *subject light constraint* de Chafe, 1994). Jeanjean (1980) confirme cette tendance de la langue française en trouvant, sur la base d'un corpus d'une conversation fortuite, une moyenne de seulement 11% de SNL en position sujet. Lors de la comparaison de différents types de discours issus d'un large éventail de locuteurs, Blanche-Benveniste (1990) révèle qu'une augmentation des SNL en position sujet est associée à un discours plus élaboré ; plus le registre est soutenu, plus les SNL sont nombreux. À la suite de cette observation, Blanche-Benveniste (1995) compare deux types différents d'histoires sur le thème des accidents, le premier consistant à raconter une histoire à l'oral, le second présentant des accidents rapportés dans la presse. Les deux types d'histoires contiennent des noms mais, dans la presse, il y avait plus de SNL (dont de nombreux types de nominalisations). Si les histoires orales se caractérisent par des verbes avec un simple argument comme par exemple, *il y a un homme, il est tombé et il est mort* les histoires écrites, quant à elles, peuvent encoder une information équivalente en une seule phrase telle que *un homme a fait une chute mortelle*.

Les formes pronominales et lexicales, sur lesquelles nous travaillons, sont présentées dans cette partie et sont mises en relation avec les concepts de statut informationnel et PAS. Après avoir travaillé ces notions, nous exposons une rétrospection des travaux menés dans ces différentes thématiques.

1.3.3. Les formes pronominales

Par son étymologie, le terme *pronom* signifie « à la place du nom » (Arrivé, Gadet et Galmiche, 1986 ; Goosse, 1993/2004 ; Mounin, 2004 ; Riegel *et al.*, 1994/2002). Le pronom peut remplacer des référents qui ont apparu (anaphore) (42) ou vont apparaître dans le contexte (cataphore) (43) ; il est alors diaphorique. Il s'utilise également souvent hors de toute relation contextuelle à un nom (44) ; il s'agit alors d'un pronom impersonnel non diaphorique.

(42) **La prof** était en colère, **elle** a collée tous les élèves.

(43) **Il** ressemblait à un fou, **le garçon** qui portait ce vêtement.

(44) **Il** arrive que les professeurs soient absents.

Creissels (2004, 2006a, 2006b) utilise le terme de *forme pronominale* afin de dénommer des mots grammaticaux « qui ont dans la phrase une distribution semblable aux constituants nominaux, et dont les référents seraient susceptibles d'être représentés par des constituants

nominaux [...] dont le sens lexical est fondamentalement relatif au contexte discursif [...] et ne met en jeu que secondairement les caractéristiques sémantiques intrinsèques de leur référent » (Creissels, 2006a). Ils peuvent également remplacer une proposition. Prototypiquement, le référent serait susceptible d'être récupéré grâce à un constituant nominal dans la production. Le pronom est un mot « *qui varie en genre et en nombre* » (Goosse, 1993/2004:955). Tout comme les constituants nominaux, ils peuvent aussi avoir différentes fonctions telles que sujet, objet, complément du nom ou de l'adjectif, etc.

La langue française compte divers pronoms : les pronoms personnels (*je, tu, il, me, te se, etc.*), démonstratifs (*celui, celle, etc.*), possessifs (*le mien, le tien, etc.*), interrogatifs (*qu(i) est-ce qui, préposition+ qui est-ce que(quoi), etc.*), relatifs (*qui, qui, etc.*), indéfinis (*nombreux, aucun, nul, etc.*) (Arrivé *et al.*, 1986 ; Riegel *et al.*, 1994/2002). Parmi ces pronoms, certains ont un rôle diaphorique (anaphorique/cataphorique) et d'autres non (entre autres, déictique / impersonnel). De plus, certains appartiennent à la catégorie des clitiques (ou formes conjointes), d'autres aux formes libres (formes disjointes).²⁸

Nous pouvons aborder des notions de l'ellipse du sujet, de la classification du *y* régional, et de l'allomorphe de *on, l'on*.²⁹

Si le français est une langue dans laquelle le sujet est préférentiellement mentionné par un SNL et/ou par un pronom (*versus* langue pro-drop telle que l'italien ou l'espagnol), il y a des cas où nous parlons de pronom zéro ou d'ellipse du pronom. Cinq situations dans lesquelles il y a ellipse du pronom sont recensées (Riegel *et al.*, 1994/2002). Il existe l'ellipse due à la modalité de l'oral qui se dispense de réaliser le sujet contraint et référentiellement vide de certains verbes impersonnels (45).

(45) *__faut pas dire ça, ce n'est pas bien.*

Ensuite, les impératifs ne requièrent pas un sujet explicite (46).

(46) *Ferme la porte ! fais tes devoirs ! Monte dans la chambre !*

Le troisième cas présenté est celui où l'expression du sujet d'un verbe n'est pas obligatoire s'il est coréférentiel avec un constituant précédent (47).

(47) *Jean ramassa ses affaires, (il) mit son chapeau, et (il) sortit.*

²⁸ Le clitique est lié strictement au verbe dont il ne peut être séparé seulement par un autre clitique (Riegel *et al.*, 1994/2002).

²⁹ Une réflexion a été menée sur la notion d'impersonnel, sur les différentes valeurs de *on* et de *ça*. Néanmoins, elle n'est pas présentée car elle n'a pas de lien direct avec le type de résultats présentés dans ce travail de thèse.

Le quatrième cas est celui présentant l'ellipse du sujet (engendrant celle du verbe) comme étant inclus dans une ellipse discursive (48).

(48) *Ou vas-tu ? (je vais) à paris.*

Lors du codage des données ces aspects ont été pris en considération ; ceci est abordé dans la Section 6.2 du Chapitre 6 (p. 143).

Les individus, qui ont participé à cette étude, habitent tous dans la région lyonnaise. Dans la mesure où nous avons remarqué l'utilisation du pronom *y* dans certaines structures, nous nous devons d'expliquer le phénomène du pronom *y* tel qu'il apparaît dans les structures des exemples 49.

(49) *Donne-moi ça, je vais y faire. – Donne-moi ça, je vais le faire.*

Cette structure, assez fréquente à l'oral, est très spécifique et fait partie intégrante des caractéristiques des locuteurs lyonnais comme de tous les locuteurs de la zone franco-provençale. Si cette tournure ne satisfait pas aux exigences morphosyntaxiques du français standard, Tuailleon estime à « *plusieurs millions (4 ou 5) le nombre de français qui emploient constamment cette forme de pronom* » (Tuailleon, 1983:368). La variété du français lyonnais aurait détourné l'usage standard du pronom complément d'objet indirect *y* pour l'utiliser comme pronom complément d'objet direct inanimé. Ainsi la dichotomie animé/inanimé neutralisée en français standard pour les compléments d'objets directs est rétablie. Dans le parler lyonnais, *y* est utilisé pour les compléments d'objets directs inanimés (50) par opposition au pronom *le* utilisé pour les compléments d'objets directs animés (51).

(50) *Rends-moi ça, j'y aime bien. – Rends-moi ça, je l'aime bien.*

(51) *Cet enfant, je l'aime bien. – *Cet enfant, j'y aime bien.*

Enfin, l'indice *l'* qui peut accompagner le pronom *on* est avant tout historique et non euphonique (Goosse, 1991, 1993/2004 ; Rey-Debove, 2001). Le pronom *on* vient du nom latin *homo*, *homme*. Étant à l'origine un nom, ce morphème était alors employé avec l'article défini *l'*, la forme élidée de l'article *le*. La forme *l'on* serait un vestige de l'inachèvement du processus de grammaticalisation du nom *homo* au pronom *on* (Coveney, 2004).

1.3.4. Les syntagmes nominaux lexicaux

Nous qualifions de SNL toutes constructions dont un des éléments fonctionne comme nom (Ravid et Berman, sous presse). Le nom, qui est l'unité de base du SNL, est nécessairement

déterminé. La forme minimale et prototypique du SNL est [*déterminant+nom*]. Il est donc constitué d'un déterminant et éventuellement modifié et/ou complété (Arrivé *et al.*, 1986 ; Ravid et Berman, sous presse ; Riegel *et al.*, 1994/2002). Les noms peuvent représenter des êtres, des choses, des actions, des sentiments, des qualités, ou d'autres phénomènes.

La catégorie du nom peut être décrite en opposition aux autres catégories du discours selon des critères morphosyntaxiques. Le nom connaît une variation en genre et en nombre mais, contrairement au verbe, il ne varie pas en temps. Du point de vue des variations, le nom, en français, s'éloigne de la catégorie du verbe, et se rapproche de celle des adjectifs, ces derniers variant en genre et en nombre par rapport aux noms qu'ils complètent. Afin de distinguer les noms et les adjectifs, nous pouvons avoir recours au critère sémantique. Ainsi le nom désigne et l'adjectif apporte des modifications ou informations au nom auquel il est rattaché. De plus, s'ajoutent des différences morphosyntaxiques. Premièrement, le nom prototypique est précédé immédiatement d'un déterminant mais pas l'adjectif (seulement si l'ordre du SNL est non prototypique [*déterminant+adjectif+nom*] et le nom est éliminé [*déterminant+adjectif*]). Deuxièmement, c'est le genre et le nombre du nom qui commandent la variation de l'adjectif, et non l'inverse.

Lorsque nous étudions le SN, nous pouvons également nous intéresser à son architecture interne. Nous pouvons parler de complexité linguistique, concept regroupant les notions de complexité lexicale et d'architecture syntaxique (Ravid, 2005 ; Ravid et Berman, sous presse). La complexité lexicale fait intervenir divers aspects dont la densité lexicale et la diversité syntaxique. La densité lexicale se définit comme le nombre d'items lexicaux dans une unité et la diversité, aux différents types d'unités utilisées dans une unité (Ravid, 2005). L'architecture syntaxique réfère à trois mesures : (a) la longueur en mot d'une unité ; (b) la profondeur, à savoir le nombre de nœud lexicaux dans une unité et (c) la diversité, à savoir les différents types d'unités syntaxiques par unité. Dans notre recherche, nous nous concentrons sur le SN, donc aux pronoms et aux SNL. Ainsi, la complexité des SNL se définit comme suit. Sont considérés comme complexes les SNL qui contiennent plus d'un item lexical et ceux étant constitués de modifieurs, de surcroît de types différents (Ravid, 2000 ; Ravid et Berman, sous presse ; Ravid, *et al.*, 2002). Les SNL ayant au moins deux items lexicaux (52) sont donc plus complexes que ceux n'en contenant qu'un seul (53).

(52) La **fil**le avec le **manteau** rouge m'a frappé.

(53) La **fil**le m'a frappé.

De même, les SNL contenant plusieurs types de modificateurs sont plus complexes (54) que ceux n'en contenant qu'un seul (55).

(54) *La petite fille avec le manteau rouge est punie.*

(55) *La fille est punie.*

Qualitativement, des différences de degré de complexité existent entre les modificateurs : ainsi les modificateurs du type subordonnée relative apparaissent comme des unités mobilisant plus de ressources cognitives que les adjectifs (Ravid, *et al.*, 2002).

1.3.5. Le syntagme nominal : rétrospection en développement

La manipulation des SNL semble être une des compétences requises permettant aux locuteurs/scripteurs d'être « *linguistically literate* » (Ravid et Tolchinsky, 2000:1 ; Ravid et Tolchinsky, 2002:420) ou « *proficient speakers* » (Berman, 2008:2). Un répertoire conséquent de ces formes lexicales ainsi que son utilisation pertinente – à savoir, (a) la variation des lexèmes, (b) leur adaptation aux contextes de production, (c) leur emploi dans un syntagme plus ou moins complexe syntaxiquement, et (c) leur présence en position sujet – permettent aux individus d'atteindre une certaine maturité langagière. Ainsi, ils accèdent à la *langue de scolarisation*, but visé par l'enseignement.

Dans des travaux antérieurs, il est montré que la proportion des SN varie selon l'âge et le contexte de production. Ainsi, la production de SN augmente avec l'âge mais la proportion des SN, d'une façon générale, est plus importante dans les textes narratifs que dans les textes expositifs (Ravid, 2000).

Lorsque la nature des SN est observée plus attentivement, les analyses révèlent des résultats différents selon le type de SN. Les textes, avec l'âge, contiennent de plus en plus de lexique et de moins en moins de pronoms (Jisa et Mazur, 2005 ; Mazur, 2005 ; Ravid, 2000, 2004, 2006 ; Ravid *et al.*, 2002 ; Strömquist *et al.*, 2002). Les enfants font de plus en plus preuve de flexibilité linguistique : leur panel de formes lexicales se développe considérablement et ils adaptent ces formes à la situation de production. De plus, les SN pronominaux ont une proportion plus importante dans les textes narratifs que dans les textes expositifs ; la situation s'inverse pour la proportion des SNL (Jisa et Mazur, 2005 ; Mazur, 2005 ; Ravid, 2000 ; Ravid *et al.*, 2002 ; Strömquist *et al.*, 2002).

D'autres travaux nuancent ces résultats en intégrant la position syntaxique des SN (Jisa et Mazur, 2005 ; Mazur, 2005 ; Ravid, 2000 ; Ravid *et al.*, 2002 ; Strömquist *et al.*, 2002). Ainsi, la production d'items lexicaux en position sujet fait partie des éléments qui se développent et qui s'acquièrent tardivement. En effet, les adultes utilisent plus de formes lexicales en position sujet que les enfants (entre autres, Argerich et Tolchinsky, 2000 ; Berman et Verhoeven, 2002 ; Jisa et Mazur, 2005 ; Mazur, 2005 ; Ravid, 2000, 2002, 2005 ; Ravid *et al.*, 2002 ; Strömquist *et al.*, 2002). Les enfants forment de plus en plus de SNL, notamment en position sujet, et diversifient leur répertoire avec l'âge (Berman et Verhoeven, 2002 ; Jisa et Mazur, 2005 ; Mazur, 2005 ; Ravid, 2000 ; Ravid *et al.*, 2002 ; Strömquist *et al.*, 2002). De plus, les SNL en position sujet sont plus importants dans les productions expositives que les productions narratives et sont plus fréquents dans les textes écrits que dans les textes oraux (Jisa et Mazur, 2005 ; Mazur, 2005 ; Ravid, 2000 ; Ravid *et al.*, 2002).

Enfin, avec l'âge, les individus produisent des SNL de plus en plus complexes (Jisa et Mazur, 2005 ; Mazur, 2005 ; Ravid, 2000 ; Ravid et Berman, sous presse ; Ravid *et al.*, 2002). En effet, les SNL des populations les plus âgées contiennent plus de items lexicaux et plus de modificateurs que ceux des populations les plus jeunes. Ce type de SNL caractérise également davantage les textes expositif et écrits (Jisa et Mazur, 2005 ; Mazur, 2005 ; Ravid, 2000 ; Ravid et Berman, sous presse ; Ravid *et al.*, 2002).

1.3.6. Le syntagme nominal : conclusion

Ainsi, les pronoms et les SNL sont les deux unités qui sont observées dans ce travail. Seulement, nous n'avons pas limité l'analyse uniquement à la nature et la position syntaxique de ces unités. Nous avons également pris en considération leur statut informationnel. Pour chacune de ces unités, nous nous sommes intéressée à leur statut informationnel dans le texte, à savoir : ancienne information ou nouvelle information. Ces unités présentées, nous pouvons alors nous concentrer sur un autre aspect de notre étude : le statut informationnel.

1.4. Cohésion/cohérence : concepts et marques, conclusion

Notre étude se concentre sur divers aspects de la syntaxe et du flux informationnel. Ce travail complète les études sur les SN dans le sens où nous souhaitons associer des observations sur

la forme au flux informationnel et surtout à la PAS de Du Bois. Ainsi, nous observons certes le rôle informationnel des pronoms et des SNL mais nous nous intéressons également à la place de l'information dans l'énoncé (position syntaxique) dans quatre contextes de production : expositif écrit, expositif oral, narratif écrit et narratif oral. Nous observons l'impact du type de texte sur l'utilisation de la langue mais également l'impact de la modalité de production, de l'ordre de passation, dans une perspective développementale.

Notre travail de thèse complète, ainsi, les études concernant le flux informationnel dans le sens où nous ajoutons des dimensions jusque-là non testées. Un des buts de ce travail de thèse est en effet de vérifier que les individus de notre population ont des préférences linguistiques qui reflètent la PAS.

L'outil universel qu'est la PAS est inséré dans une étude développementale du français. Partant du postulat que le développement langagier continue au-delà de cinq ans, nous nous attendons à d'importantes variations entre les différents groupes d'âge. Si nous savons que la PAS caractérise les productions orales de locuteurs français âgés de 16 à 30 ans (Ashby et Bentivoglio, 1993), nous ne savons pas ce qu'il en est des enfants plus jeunes et ne connaissons pas l'évolution de ces contraintes linguistiques à travers le développement. Nous avons des indications dans des langues assez différentes du français et surtout en acquisition précoce mais nous n'avons guère d'analyses sur les enfants et adolescents francophones natifs et encore moins d'un point de vue développemental.

Nous avons un corpus assez conséquent dans la mesure où notre population est composée de 132 individus ayant produit chacun quatre textes, à savoir un total de 528 textes soit 57 302 mots et 9 765 clauses. Les études sur la PAS réalisées jusque-là regroupent de plus petits corpus avec un nombre plus limité de participants.

Nous avons analysé des textes oraux, narratifs et expositifs, comme cela l'a été fait précédemment, mais nous avons également ajouté des textes écrits. Si nos résultats révèlent que les individus de notre étude ont un pattern préféré, qui de surcroît répond aux critères de la PAS, il peut être intéressant de voir si la PAS s'impose dans les deux modalités de productions, étant données les importantes différences entre l'oral et l'écrit.³⁰ Dans une même étude, nous pouvons comparer les choix linguistiques des enfants et des adolescents dans des textes oraux et écrits, dans des productions narratives et expositives.

³⁰ Les différences entre les deux modalités du langage (orale et écrite) sont présentées dans le Chapitre 2 qui suit.

Nous avons également observé l'impact d'une autre variable indépendante, celle de l'ordre de production. Comme nous l'expliquons dans la Partie II *Méthodologie*, nous avons, lors de la collecte des données, divisé notre population en deux grands groupes : (a) une première moitié de nos individus a produit d'abord un texte écrit puis un texte oral, et (b) une deuxième moitié a produit un texte oral puis un texte écrit. Il est dit communément que les enfants écrivent comme ils parlent. Est-ce que cette idée de l'influence d'une modalité du langage sur l'autre peut être confirmée ? Les individus peuvent-ils parler comme ils écrivent ? Nous pouvons nous demander si le fait de produire une première fois dans une des modalités (écrit/oral) a un impact sur la deuxième production, de modalité différente.

Chapitre 2

Produire dans différents contextes : à l'oral, à l'écrit, en narration, en exposition

Dans ce chapitre sont présentées les caractéristiques et la complexité des deux modalités du langage (orale/écrite) puis deux modèles de production langagière. S'ensuit une section définissant les deux types de textes, expositif et narratif, que les participants de notre étude ont produits. Ce chapitre s'achève avec une focalisation sur le lien entre contexte de production (MODALITÉ et TYPE DE TEXTE), développement et planification du message.

2.1. La production langagière : informations préliminaires

Divers niveaux caractérisent la production langagière : *cognitif, lexico-grammatical, lexico-phonologique et articulatoire ou grapho-moteur* (Coirier, Gaonac'h et Passerault, 1996).

Le *niveau cognitif* est commun aux deux types de production langagière, orale et écrite. Il permet l'activation, la sélection puis l'organisation dans la *mémoire à long terme* (MLT) des contenus à communiquer. La MLT est souvent mise en opposition à la *mémoire de travail*

(MDT).³¹ Le rôle principal de la MLT dans la production textuelle est de stocker l'ensemble des connaissances des locuteurs, des informations référentielles, pragmatiques aux linguistiques (Alamargot, Chanquoy et Chuy, 2005 ; Alamargot, Lambert et Chanquoy, 2005). Hayes (1998) rappelle l'importance de la MLT dans l'activité langagière. Concernant la production de l'écrit, il suggère que : « *Rédiger ne serait pas possible si les rédacteurs ne possédaient pas des mémoires à long terme dans lesquelles stocker leurs connaissances sur la grammaire, le vocabulaire, le genre, le thème, le destinataire, etc.* » (Hayes, 1998:96). Cette remarque est tout aussi valide pour la production langagière orale. La MLT est en lien étroit avec la composante de la planification et le processus de récupération et influence le type et la qualité du produit final de l'activité langagière (Alamargot, Chanquoy et Chuy, 2005). Divers types de connaissances caractérisent la MLT. Tijus (2001) parle, par exemple, de *connaissances déclaratives* et de *connaissances procédurales*. Les *connaissances déclaratives* « *recouvrent un vaste ensemble de savoirs qui portent aussi bien sur l'orthographe que sur les types de textes* » (Fayol et Heurley, 1995:28). Concernant la mise en texte, trois types de *connaissances déclaratives* peuvent être mobilisées : les connaissances du domaine, les connaissances rhétoriques, et celles regroupant les aspects pragmatico-linguistiques. Les *connaissances procédurales* recouvrent les processus (ou savoir-faire) (Alamargot, Lambert et Chanquoy, 2005 se basant notamment sur Anderson 1983 et 1993), dont celui de la planification, qui sont certes stockés en MLT, mais activés en MDT. La MDT se définit comme le composant du système cognitif qui, entre autres, stocke temporairement les informations (Kellogg, 1998).

Le niveau lexico-grammatical grammaticalise la sélection des informations et gère leur organisation. Des choix syntaxico-lexicaux définissant la planification linguistique sont effectués. Cette planification linguistique est cognitivement coûteuse tout comme la planification conceptuelle (Bloom, Miller et Hood, 1975 ; Foulin et Fayol, 1988 ; Kellogg, 1987). Ce niveau est davantage détaillé pour chacune des deux modalités du langage dans la Section 2.3 (p. 59) de notre travail de thèse.

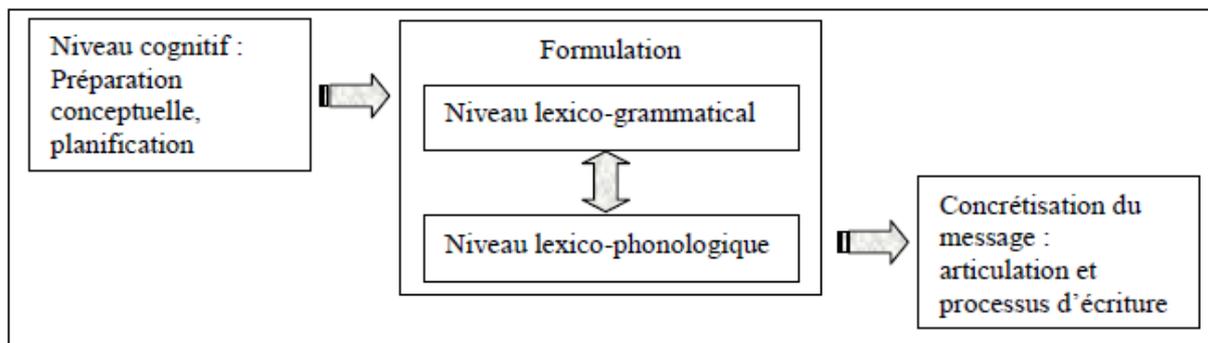
Les niveaux lexico-phonologique et articulatoire/grapho-moteur sont les dernières étapes. Ils permettent la construction des séquences sonores et des choix des morphèmes adéquats, à savoir la matérialisation du message. Ce serait au *niveau lexico-phonologique* que nous

³¹ Nous définissons ces deux concepts en nous basant notamment sur les travaux de Baddeley (entre autres, 1986, 1990/1992, 1992, 2000 et 2001 qui s'inspire du *modèle modal* d'Atkinson et Shiffrin (1968, 1971) ; Baddeley et Hitch, 1974 ; Gathercole et Baddeley, 1993). Cette conception est l'une des plus heuristiques à retenir (Olive et Piolat, 2005 ; Olive, 2004).

trouverions « *les marques linguistiques (lui) permettant de signaler la nature des relations entre les différentes informations (anaphores, connecteurs, ponctuation,...)* » (Coirier et al., 1996:147). Puis les structures sous-jacentes établies précédemment sont transformées en séquences articulatoires ou graphiques.

Les productions orales et écrites possèdent toutes deux ces composants essentiels. Nous pouvons schématiser les quatre grands niveaux de la production langagière en nous inspirant de Fayol (2002a, 2002b, 2002c) et Ferrand (2002) dans la Figure 4 qui suit.

Figure 4 : Schématisation des étapes de la production langagière



Si, lors de l'émission d'un message écrit ou oral, ces phases de la production textuelle sont engagées, les différences de ces deux modalités engendrent une gestion différente de ces niveaux. En effet, elles se différencient sur un certain nombre de points.³²

Aussi, allons-nous désormais étudier les caractéristiques des deux modalités du langage qui ont un impact sur le traitement linguistique. Faut-il alors considérer les deux modalités du langage comme dépendantes d'un seul système ou de deux systèmes distincts ? Dans ce cas, la séparation en deux modalités est-elle stricte ?

2.2. Le contexte de production des deux modalités

La modalité orale et la modalité écrite possèdent des caractéristiques propres que ce soit en production (Blanche-Benveniste, 1983, 1985, 1990, 1995, 2000 ; Fayol, 1990, 1997 ; Gadet, 1989, 1996 ; Gayraud, 2000 ; Péry-Woodley, 1993) ou en compréhension (par exemple, Ferreira et Anes, 1994).

³² Ces points sont développés dans la Section qui suit.

Si la production à l'oral et à l'écrit se réalise dans des circonstances bien différentes, il n'est pas question, ici, de parler de deux codes stricts d'une même langue ou encore de débattre sur la complexité supérieure d'une modalité sur l'autre. Il a été montré que cette discussion « *semble (donc) peu fertile* » (Gayraud, 2000:52, selon Biber, 1988). Ainsi, mieux vaut-il considérer ces deux modalités comme un continuum des situations d'énonciation et des variations (Biber, 1988 ; Fayol, 1990, 1997). L'opposition oral/écrit relèverait « *d'un continuum sur lequel se situent des activités plus ou moins formelles (rédaction scolaire ; lettre administrative mais aussi discours oral en public) ou peu contraintes (discussion entre familiers ; brouillon en vue d'une rédaction ; etc.)* » (Fayol, 1990:21).

Néanmoins, les situations prototypiques de l'oral et de l'écrit divergent sur divers points (Tableau 2) qui ont un impact sur le traitement linguistique.

[Tableau 2 : Principales différences entre un oral et un écrit prototypique]

Tableau 2 : Principales différences entre un oral et un écrit prototypique

	Modalité orale	Modalité écrite
Énonciation	Dialogique	Monologique
Conversation	Spontanée	Délibérément travaillé
Rythme de production	Rapide	Lent
Retour possible	Non	Oui
Type de pression	Communicative	Normes rédactionnelles
Partage du temps/espace	Oui	Non
Cadre collaboratif	Plus évident	Moins évident

Premièrement, si la modalité orale se caractérise par une situation d'énonciation prototypiquement interactive, la situation d'énonciation de la modalité écrite est monogérée. À l'oral, le locuteur régule son message par rapport aux informations que les différents canaux de la situation communicationnelle donnent. Il gère son discours par rapport aux messages verbaux et non-verbaux des destinataires. Par opposition, à l'écrit, il n'y a pas de destinataire direct.

Deuxièmement, le langage spontané de l'oral prototypique se distingue du *délibérément travaillé* de l'écrit (*Deliberate working over*, Chafe, 1994:43). Lors d'une communication orale, les locuteurs encodent les messages de manière spontanée et rapide. En revanche, à

l'écrit, les scripteurs ont tout le temps pour produire un discours construit et pour le travailler autant que nécessaire. Ils peuvent réfléchir aux structures et aux lexiques ; ils peuvent également revenir sur leur production pour la modifier autant de fois qu'ils le souhaitent.

Troisièmement, si l'oral se définit comme possédant un rythme de production plutôt rapide, le rythme de production est plus lent à l'écrit. À l'oral, les individus doivent mobiliser rapidement les informations relatives à un message (Chafe, 1994). C'est ainsi que nous trouvons dans les productions orales une « *accumulation de parasites* » (Gadet, 1989:49-50) : les hésitations (*euh, ben, etc.*), les répétitions (*le le le le le moyen*) les phrases inachevées (*je pense que... une ville comme ça..., etc.*) (Gadet, 1989:51). Prototypiquement, la modalité écrite assouplit cette contrainte temporelle grâce à un rythme de production plus lent et plus modulable. Les locuteurs bénéficient, à l'écrit, d'un temps de planification potentiellement supérieur ce qui facilite, par exemple, la recherche lexicale (Fayol, 1997), la production d'un lexique plus riche (Johansson, 1999) ou la production de formes syntaxiques moins fréquentes (Bonin, 2003 ; Gayraud *et al.*, 2001 ; Gayraud *et al.*, 2000 ; Jisa, 2004).³³ Les parasites de l'oral et le rythme lent et autogéré de l'écrit permettent, par exemple, de réactiver en MDT ce qui a été produit pour maximiser les chances d'une meilleure cohérence et cohésion et/ou modifier le texte si la nécessité s'en fait sentir (Alamargot *et al.*, 2005). Néanmoins, ceci se fait plus aisément à l'écrit qu'à l'oral, d'autant plus que la modalité écrite à la particularité de laisser une trace visible (Tolchinsky, 2006).

En effet, la permanence de la trace écrite rend possible un retour en arrière pour modifier sa production. Lors d'une production orale, l'absence de trace implique un retour en arrière plus compliqué. Néanmoins, même si cela ne se fait pas dans les mêmes conditions, des retours sont également possibles à l'oral. Johansson (1999), qui se base sur Ong (1982), rappelle que le produit final de l'écrit perdure plus longtemps et reste visible alors que le produit final de l'oral est seulement audible. À l'écrit les individus peuvent modifier leurs textes plus aisément en intégrant éventuellement des formes qu'ils n'avaient pas mobilisées précédemment.

Cinquièmement, l'oral impose aux locuteurs une pression communicative que la modalité écrite n'exige pas. En effet, le locuteur, pour mener à bien son intervention, doit – en plus de mobiliser les informations pertinentes au message, les planifier et les encoder, capter

³³ Nous pouvons donner comme exemple des SNL contenant plusieurs noms lexicaux ou des structures telles que les subordonnées non finies.

l'attention des destinataires – prévenir et empêcher les interruptions des autres participants. D'un autre côté, la modalité écrite implique un registre académique ainsi qu'une norme rédactionnelle exigeante. Ceci implique alors la mobilisation de formes lexicales et de structures syntaxiques rares, appartenant à un registre élevé. Elle impose également une pression informative importante ; les scripteurs se doivent de fournir un maximum d'information. Nous disons alors que l'écrit apparaît comme plus planifié que l'oral.

Sixièmement, les modalités du langage n'ont pas le même degré de partage du temps et de l'espace. Prototypiquement, en production orale dialogique, le locuteur et le destinataire du message partagent le temps et l'espace, ce qui n'est pas le cas de la production écrite. La synchronisation entre l'émetteur et le destinataire d'un message ne s'effectue pas de la même manière à l'oral et à l'écrit (Auer, 2009).

Nous déduisons alors de ce sixième aspect, que, et ceci est notre septième point, que la communication écrite ne propose pas un cadre aussi collaboratif que la communication orale. Le *common ground* est plus difficile à établir. Cependant, à l'écrit, les individus pourraient exploiter le rythme de production lent de la production écrite pour se mettre à la place du destinataire ; ce qui est rendu plus difficile à l'oral, caractérisé par un rythme de production plus rapide.

Les modalités orale et écrite divergent donc sur de nombreux points. C'est dans cette mesure que la transmission de l'information peut également différer entre les deux modalités du langage. Rappelons qu'un de nos centres de préoccupations est la PAS, qui reflète la transmission de l'information à l'oral. Étant données les différences entre les deux modalités, le pattern préféré des locuteurs à l'oral (la PAS) n'est alors certainement pas le même à l'écrit. Chacune des modalités a des caractéristiques spécifiques ce qui a des conséquences sur les choix linguistiques effectués d'une production textuelle et sur le type de planification mobilisée.

Ainsi, par exemple, les unités phrastiques, à l'oral, sont de petite taille étant donné le rapide temps de traitement et le caractère éphémère de l'oral (Auer, 2009). Ces unités ne véhiculent qu'une seule information à la fois en sachant que l'information nouvelle est, en général, placée préférentiellement en position post verbale et que la position sujet est remplie par les pronoms. La modalité orale met en jeu la contrainte du *sujet léger*. Elle se manifeste par l'évitement de SNL en position sujet, fonction syntaxique réservée à l'ancienne information et donc à des formes exprimant l'accessibilité telles que le pronom (Chafe, 1994).

L'écrit irait à l'encontre de cette contrainte du *sujet léger* en facilitant la production de sujets lexicaux pouvant introduire un nouveau référent (Chafe, 1994). Cet aspect n'est pas sans rappeler les contraintes de la PAS.³⁴ Rappelons que les clauses contiennent un élément lexical, qui ne se trouve pas en position sujet. De plus, les clauses ne comportent qu'une seule nouvelle information et l'ancienne information est maintenue en position sujet. La PAS reflète les caractéristiques de la modalité orale qui impliquent une transmission de l'information spécifique. Ainsi, les contraintes de la PAS peuvent sans doute être violées en production écrite.

Les exemples ci-dessous, qui sont deux productions narratives (56) est la version écrite et (57) la version orale) d'une adolescente de 3^{ème}, illustrent ces différences.

(56) Mon frère m'a conseillé d'acheter le portable que vendait la sœur de son copain. J'ai suivi son conseil. Sa sœur me l'a vendu à 25 euros mais, au moment de payer, après qu'elle ait été conseillée par plusieurs de ses amies, elle changea les prix et m'a dit qu'elle me le vendrait 40 euros. J'ai dit « Tchao », ça lui a pas plu. Deux jours après, je l'ai vue dans la rue, elle m'a insulté, je l'ai frappée, mon frère est venu nous séparer. (3^{ème}, n°01, narratif écrit, O/E)³⁵

(57) C'était pour une histoire de portable, j(e) m(e) rappelle. C'était euh mon frère, en fait, i(l) m'avait trouvé un portable. C'était euh la p(e)tite sœur de son copain, ben elle vendait son portable et donc et ben: elle le vendait pas cher. Et puis pff c'était une pigeonne, elle l'a vendu pas cher du tout. Donc moi j'ai dit ouais, j'ai accepté et tout, et au moment où on donnait l'argent et tout elle me fait non non, on s'était pas arrangée comme ça, et tout, après j(e) lui fais ben si, on s'est arrangé comme ça, puisque tout le monde lui avait dit mais tu te fais voler et tout. Alors que moi j(e) m'en foutais tant qu(e) j'avais mon portable et tout. J(e) lui ai dit, écoute, et tout tu vas pas changer les prix à la dernière seconde, ça s(e) passe pas comme ça, moi j(e) suis pas une pigeonne et tout. Et donc ensuite ben on s'est disputé et tout. J(e) suis rentrée chez moi, j'avais pas récupéré l(e) portable, j'étais énervée et donc j(e) suis partie. J'ai dit à mon frère et tout et donc après i(l) m'a dit, c'est pas grave, et tout. Puis après dès qu'elle m'a revue dans la rue, une fois, elle m'a taillée et donc ça m'a pas plu, et donc j(e) suis partie la voir et donc après j(e) l'ai frappée. Après dès qu'on avait fini d(e) se

³⁴ Notions développées dans la Section 1.2.3 (p. 36).

³⁵ Les exemples issus des productions des individus ayant participé à l'étude ReFlex sont tous titrés de la même manière. Ainsi, – 3^{ème}, n°01, narration écrite, O/E – signifie qu'il s'agit: (a) d'un individu scolarisé en 3^{ème}; (b) de l'individu anonyme sous le numéro 01; (c) d'un texte narratif; (d) d'un texte écrit; et enfin (e) d'un individu ayant produit dans un premier temps un texte oral puis dans un second temps un texte écrit (O/E).

battre et tout, i(l) y a mon frère i(l) passait par là et donc euh ouais par chance i(l) nous a séparées puis j(e) suis rentrée chez moi. Et j(e) me suis fait éclater. (3^{ème}, n°01, narratif oral, O/E)

Ainsi, dans la production orale (57), la collégienne utilise des sujets pronominaux plus légers (*c'était mon frère, il m'avait trouvé un portable*) que les sujets lexicaux utilisés à l'écrit (56) (*Mon frère m'a conseillé, sa sœur me l'a dit*). De plus, à l'écrit (56), l'individu produit 9 items lexicaux différents alors qu'à l'oral (57) elle en utilise 7 différents, pour un texte beaucoup plus long ; l'information à l'écrit, contrairement à l'oral, est condensée. Enfin, nous pouvons noter un lexique plus familier à l'oral avec l'utilisation des termes *pigeonne, tailler, éclater*.³⁶

Produire à l'oral et à l'écrit implique des processus de haut niveau identiques : la gestion des concepts et des dimensions communicatives, l'accès lexical, la structure syntaxique ou encore l'activation et la sélection d'éléments dans la MLT et dans la MDT.

Les différences résident dans le temps consacré à ces processus cognitifs et aux phénomènes de révision beaucoup plus pertinents pour l'écrit. Les différentes contraintes de ces modalités impliquent une structuration de l'information divergente. L'apprentissage de la langue écrite impliquerait même « *une réorganisation fondamentale du système de production langagière* » (Schneuwly, 1988:167). Aussi avons-nous décidé de travailler avec deux modèles que nous mettons en relation : l'un adapté à la production orale et l'autre adapté à la production écrite.

La section qui suit se consacre alors à la présentation des modèles de production écrite et orale. Il existe plusieurs modèles de la production langagière tels que celui de Hayes (1998) basé sur le modèle de Hayes et Flower (1980), ceux de Garrett (1988) et Levelt (1989, 1999) et celui de Van Galen (1990). Nous nous basons sur le modèle de Hayes (1998)³⁷ pour la production écrite et sur celui de Levelt (1989, 1999) pour la production orale.

2.3. Modèles de production langagière

2.3.1. Modèle de production écrite

2.3.1.1. Description du modèle

³⁶ Ces items ne sont pas utilisés dans leur sens premier ; *pigeonne* signifie personne dupe ou naïve exploitée facilement par d'autres ; *tailler* a le sens de *critiquer/insulter* ; *se faire éclater* a la signification de *se faire disputer* (très fort) (Cf. liste des formes non standard en Annexe 1).

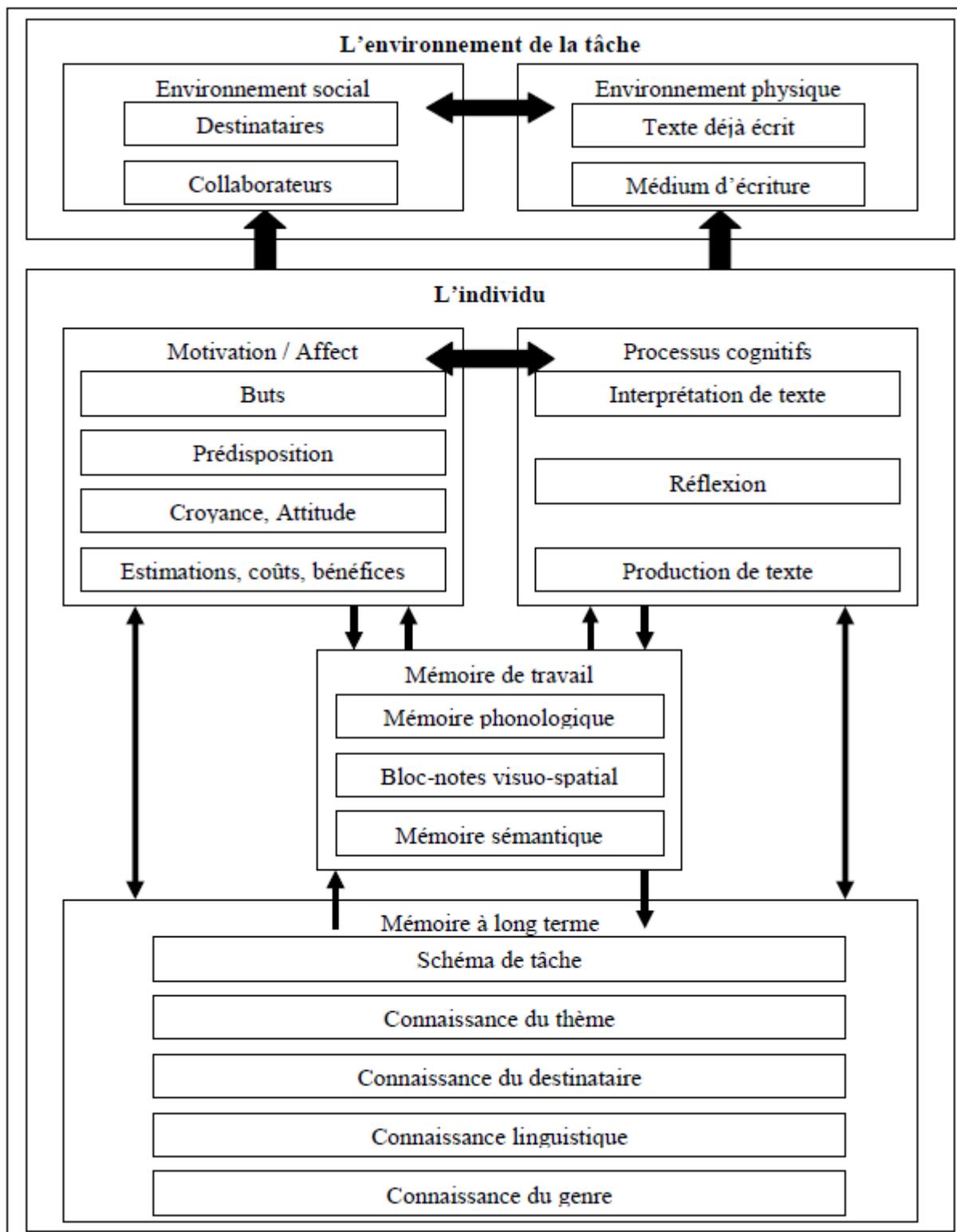
³⁷ Les deux modèles choisis présentent une schématisation pour des adultes sains.

Les recherches se concentrant sur la production écrite, qui constituerait une des activités cognitives les plus coûteuses³⁸ et complexes³⁹, débutent avec les travaux de Hayes et Flower en 1980 sur lesquels nous nous basons pour analyser la production écrite. Nous nous concentrons sur une version révisée du modèle de Hayes et Flower (1980) de la production langagière écrite (Hayes, 1995, 1998 dans Piolat et Pélissier, 1998) (Figure 5).

Figure 5 : Organisation générale du nouveau modèle du processus d'écriture (Hayes, 1998:56)

³⁸ Entre autres, Favart et Olive (2005), Kellogg (1987, 1994), Olive et Piolat (2005), Piolat, Olive et Kellogg (2005).

³⁹ Entre autres, Torrance et Galbraith (2006).



Deux composantes principales caractérisent ce modèle de la production écrite: celle de *l'environnement de la tâche* et celle de *l'individu*. *L'environnement de la tâche* est défini par *l'environnement social* et *physique*. *L'environnement social* réfère aux conventions sociales

auxquelles sont soumis et font face les scripteurs. Nos écrits sont divers et s'adaptent aux destinataires et aux contextes. *L'environnement physique* correspond au *texte déjà écrit* et au *médium d'écriture*. Il s'agit de l'intégration dans le modèle de la forme intermédiaire du produit et des différents types de moyens d'écriture tels que le traitement de texte ou le courrier électronique.

La composante de *l'individu* rassemble *motivation/affect, cognition* et *mémoire*. Le modèle de Hayes se centre sur l'individu et sur les fonctions psycho-cognitives concernant notamment les processus de planification. Ce modèle accorde une grande place à la MDT qui a un « *rôle primordial dans l'activité rédactionnelle* » (Hayes, 1998:62). Hayes donne de la MDT la même définition que Baddeley (1986). Ce système cognitif apparaît comme une ressource limitée que les locuteurs/scripteurs utilisent aussi bien pour stocker des informations que pour effectuer des traitements mobilisant des processus cognitifs. Dans ce modèle de rédaction, la planification, les prises de décisions relatives aux textes et les révisions font partie des processus de réflexion. Le scripteur élabore des plans rédactionnels, à savoir des *plans de composition* (Piolat et Roussey, 1992) adaptés à la situation de discours et guidés par les informations récupérées dans la MLT.

Les principaux *processus cognitifs* impliqués dans la rédaction sont *l'interprétation de texte, la réflexion* et *la production de texte* (Hayes, 1998). *L'interprétation de texte* permet l'élaboration de représentations internes des données linguistiques et graphiques. *L'interprétation de texte* relève également de tous les aspects liés à la révision, sur lesquels *la réflexion* se base afin de créer d'autres représentations que *la production de texte* met en lien avec le contexte de la tâche. D'autres processus existent tels que *la résolution des problèmes* (planification incluse), *la prise de décision* et *la production d'inférences* (Hayes, 1998). *La résolution des problèmes* correspond au moment où le scripteur sait qu'il a à atteindre un but spécifique mais ignore encore comment il va y arriver. Il réfléchit aux étapes par lesquelles il doit passer : de la création d'un plan à l'encodage du message. *La prise de décision* intervient lorsque l'individu doit prendre une décision concernant son texte : il évalue « *les alternatives parmi lesquelles il(s) doit(ven)t choisir* » (Hayes, 1998:89). Enfin *la production d'inférences* définit le processus « *par lequel une nouvelle information tire sa source d'une ancienne* » (Hayes, 1998:90).

La *motivation/affect* est essentielle pour une compréhension plus complète de la production écrite (Hayes, 1998). Quatre aspects sont impliqués : le *but* de la rédaction, les

prédispositions de l'individu face à un type d'écrit, la *croyance* et l'*attitude* de l'individu et enfin *les estimations, les coûts et les bénéfices* d'un écrit. Ainsi, lors d'une tâche spécifique, un scripteur doit gérer les différents buts de sa rédaction, qu'il met en place selon ses dispositions et ses croyances à l'écrit. Ce cadre de la rédaction proposé par Hayes (1998) est complété également des modèles plus spécifiques à la planification ou encore à la révision (Allal, Chanquoy et Largy, 2003).

2.3.1.2. Modèle de production écrite : conclusion

Ce modèle de Hayes et Flower (1995, 1998) met l'accent sur le contexte de production. L'environnement de la tâche semble être un facteur de variations linguistiques. Néanmoins, ce modèle, même s'il demeure encore une référence centrale (Alamargot et Chanquoy, 2002), présente quelques limites (Berninger et Swanson, 1994 ; Hayes, 1996).

Deux principales remarques peuvent être faites. Premièrement, ce modèle s'attache à la description du scripteur expert. Il ne renseigne ni ne décrit la construction progressive de l'expertise rédactionnelle (Alamargot et Chanquoy, 2002). Le modèle développemental de Swanson et Berninger (1994) complète ce modèle de Hayes et Flower. Deuxièmement, ce modèle reste flou quant aux processus linguistiques (tels que la formulation, l'encodage grammatical, *etc.*), qui sont davantage travaillés dans le modèle de la production orale de Levelt (1989, 1999). Dans ce modèle, Levelt décrit avec plus de précision les différents niveaux de la production langagière en détaillant les choix linguistiques (comme l'accès lexical ou l'encodage grammatical) tout en prenant en considération le contexte et les connaissances du locuteur.

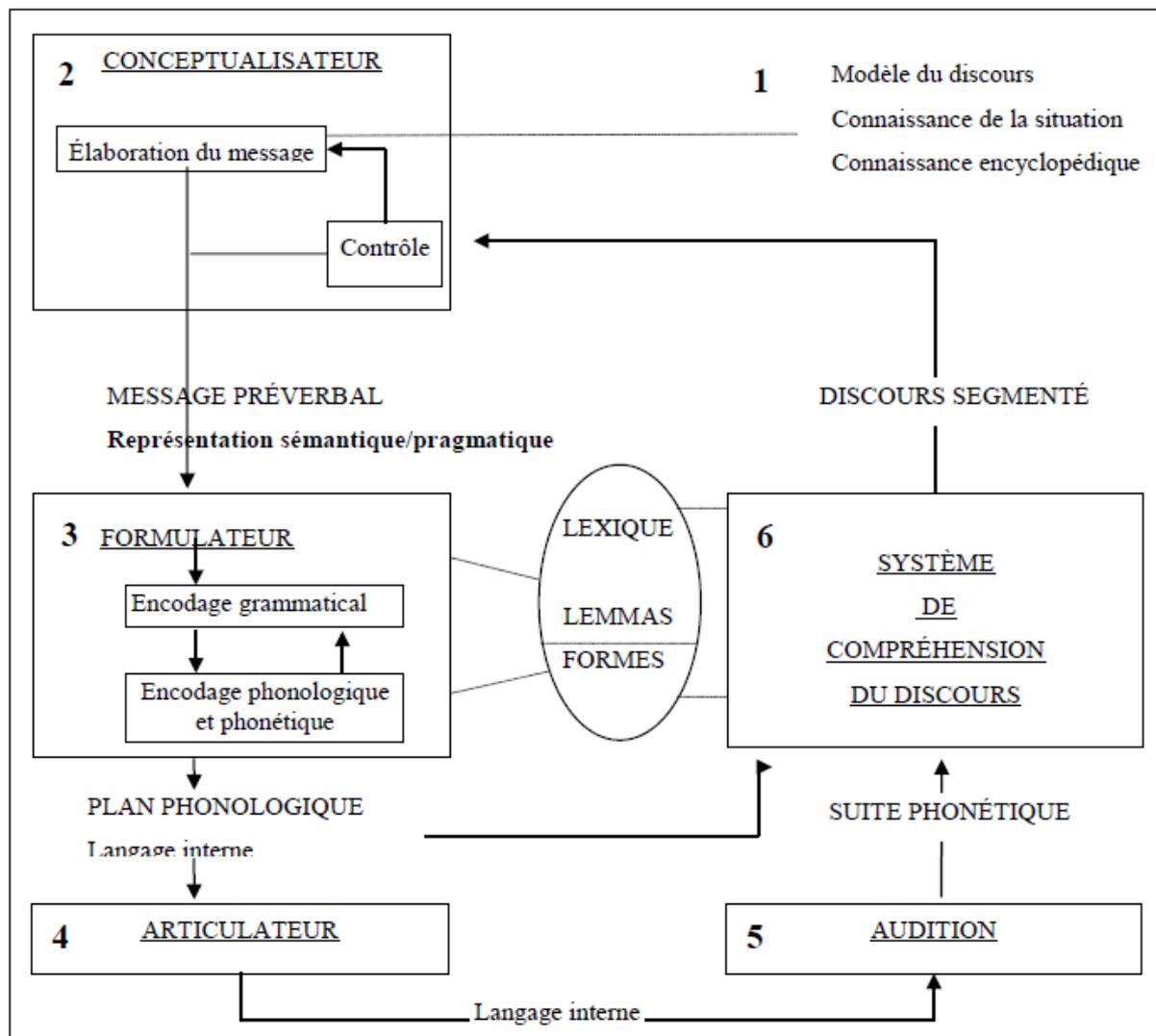
2.3.2. Modèle de production orale

2.3.2.1. Description du modèle

Le modèle de Levelt (1989, 1995, 1999), qui s'inscrit « *dans une perspective psycholinguistique* » (Fayol et Heurley, 1995:20), paraît le plus approprié pour la compréhension de la production orale. Ce modèle s'est inspiré de celui de Garrett (1975, 1980a, 1980b, 1982, 1984, 1988) qui est basé sur l'analyse systématique des erreurs en production orale (Fayol, 1997). La différence la plus importante entre Garrett et Levelt est que Levelt dépasse la phrase, observe la syntaxe inter-phrastique et considère le texte comme un

tout ; argument essentiel pour notre étude. La Figure 6 présente ce modèle (Levelt, 1989 ; Traduction de Fayol, 1997).

Figure 6 : Modèle de production orale du langage (Levelt, 1989)



Ce modèle est constitué de divers modules : un premier niveau (1) correspondant *au modèle du discours et les connaissances de la situation et encyclopédiques*, le *conceptualisateur* (2), le *formulateur* (3), l'*articulateur* (4), l'*audition* (5) et enfin le *système de compréhension du discours* (6). Le premier niveau (1) inclut la représentation des contextes communicatifs ainsi que la connaissance que les locuteurs stockent dans leur mémoire.

Le conceptualisateur (2), qui représente le *niveau cognitif* de la production textuelle, active, sélectionne et organise les idées selon l'intention communicative du locuteur en ayant recours à deux types de connaissances, le déclarative et le procédurale (Levelt, 1989). Le locuteur doit

choisir entre de nombreuses stratégies selon son intention et son destinataire (Levelt, 1995). Selon le type de discours, le locuteur mobilise des connaissances spécifiques. Ainsi, s'il produit une narration, il sollicite des connaissances relatives à la structuration conventionnelle de ce type de texte (situation, nœud, action, *etc.*) et le modèle mental⁴⁰ de son message. Selon le type de texte, la structure est plus ou moins évidente. Le texte expositif a une structure beaucoup moins clairement définie et connue des jeunes locuteurs que celle du texte narratif.

L'individu doit organiser les contenus activés. Pour cela, il fait appel à deux types de planification : *la planification par connaissances rapportées* et *la planification par connaissances transformées* (entre autres, Bereiter et Scardamalia, 1987 et 1998 ; Bereiter, Burtis et Scardamalia, 1988).⁴¹

Le conceptualisateur permet l'élaboration des discours sous forme de messages préverbaux. Ces derniers prennent la forme de propositions prélinguistiques incluant diverses représentations telles que les représentations sémantiques (prédicat, agent, patient, instrument, expérience, *etc.*) ou les représentations pragmatiques (topique, commentaire, assertion, interrogation, *etc.*). La prise en considération de la situation du discours est essentielle, le locuteur se constitue alors une représentation mentale de cette dernière (Levelt, 1989). Ce modèle mental de la situation évolue selon le type de discours (narration, description, dialogue, *etc.*) et le locuteur adapte ses choix. Ces messages prélinguistiques sont soumis à deux étapes de planification : la macroplanification et la microplanification (Levelt, 1989). La macroplanification correspond à la récupération des données pertinentes à la situation communicationnelle. Selon le contexte, ce qui a été dit, ce qui va être dit, les connaissances partagées, les locuteurs font des choix différents. Ces informations sont ensuite encodées par le biais de la microplanification impliquant divers aspects tels que l'accessibilité des référents (soit la structure informationnelle des éléments) ou la topicalisation. L'accessibilité des référents est un des aspects majeurs de la microplanification. Le locuteur fait son choix parmi les formes linguistiques à disposition en fonction du statut informationnel des référents à introduire. Pour cela, il doit porter attention à la situation discursive.⁴²

Le formulateur comprend *l'encodeur grammatical* et *l'encodeur phonologique*. *L'encodeur grammatical* élabore la forme syntaxique du message en sélectionnant les items lexicaux dans le lexique mental et en prenant en compte les caractéristiques syntaxiques de ces items dans

⁴⁰ Notion présentée dans le Chapitre 1, Section 1.1 (p. 22).

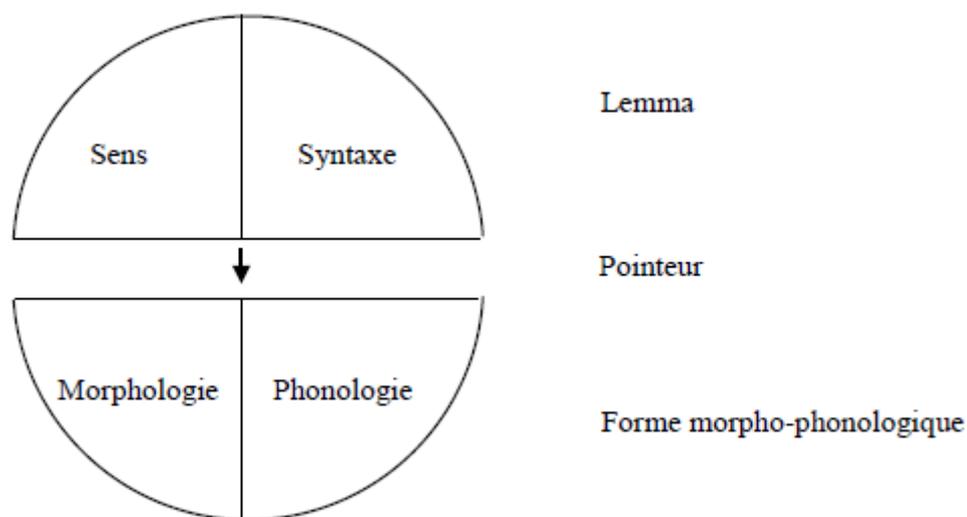
⁴¹ Notions présentées dans le Chapitre 2, Section 2.6 (p. 79).

⁴² L'accessibilité des référents selon Levelt est présentée au Chapitre 1, Section 1.2 (p. 32).

l'élaboration de la structure syntaxique du message. Un important aspect de *l'encodage grammatical* est la cohésion (*cohesive encoding*) ; l'individu construit son message en prenant en considération ce qui a déjà été mentionné dans le discours, ce qu'il est en train d'énoncer et ce qu'il va émettre. La cohésion ne se limite pas à l'interprétabilité des référents mais va au-delà. Les choix des formes linguistiques dépendent du type de référent à introduire, de son accessibilité, du contexte de production, *etc.* Ainsi, utiliser un SNL ou un pronom dépend de tous ces facteurs.

Le *formulateur* est lié au composant du lexique. Une entrée lexicale comprend quatre types d'information concernant : sa signification, ses propriétés syntaxiques, ses caractéristiques morphologiques et la spécification de sa forme (Bock et Levelt, 1994 ; Levelt, 1989, 1999). Les deux premiers types de traits correspondent aux lemmas, les deux derniers à la forme morpho-phonologique de l'entrée lexicale (Figure 7). Une fois un lemma activé, sa forme morpho-phonologique lui est attribué par le biais du pointeur lexical.

Figure 7 : Constitution d'une entrée lexicale (Levelt, 1989:188)



Ainsi, l'entrée *voiture*, se constitue : (a) d'un sens, *mode de transport* ; (b) d'informations syntaxiques, mot appartenant à la catégorie des *noms*, *nom concret*, *etc.* ; (c) de caractéristiques morphologiques comme l'ajout du morphème *s* pour exprimer le pluriel ; et (d) de spécificités phonologiques sur les phones constituant le signifié de cette entrée, à savoir des informations d'ordre structurel (unité monosyllabique du type VC, *etc.*) À ces diverses spécifications, s'ajoutent des propriétés d'ordre pragmatique et discursif qui nous intéressent

tout particulièrement. Le choix d'une forme se réalise par rapport à ce qui a été énoncé, ce qui est dit et ce qui va être dit et selon le but et le contexte communicationnel de la situation. Ainsi, le choix de l'item *voiture* dans un message, est en lien avec le message en cours. Selon que le référent de cette entrée lexicale est connu des interlocuteurs ou non, le locuteur fait des choix différents. Ainsi, le statut informationnel peut avoir un impact sur le déterminant ou encore sur la position syntaxique de l'item.

Les choix relatifs aux formes lexicales se font extrêmement rapidement et en fonction du contexte de production. Comment le locuteur fait-il un choix aussi rapide dans un lexique mental composé de milliers d'entrées ? Levelt (1989) suppose un traitement semi-automatique et parallèle. L'organisation conceptuelle du message est influencée par divers facteurs : contexte de production, l'intention communicationnelle, *etc.* En production orale, le rythme de production peut être beaucoup plus rapide qu'en production écrite. Soumis au rythme de production, l'accès lexical se fait différemment selon la modalité. Ainsi, il se fait très rapidement à l'oral, alors qu'à l'écrit, cet accès lexical peut être plus lent. Ainsi, selon le contexte de production, les locuteurs feraient des choix différents liés aux rythmes de production des modalités.

L'encodeur phonologique donne les formes phonologiques des items choisis préalablement. Ainsi le message *préverbal* prend une forme linguistique en deux temps. Dans un premier temps, le système établit une représentation fonctionnelle avec en parallèle une structure syntaxique associée à des éléments lexicaux. Dans un second temps, cette nouvelle structure syntaxique et les formes lexicales qui lui sont associées sont envoyées sous forme de chaîne ordonnée de lemmas à la composante articulatoire.

L'articulateur (4) est sollicité pour encoder les formes choisies précédemment et exécute alors les choix faits par l'encodeur phonologique.

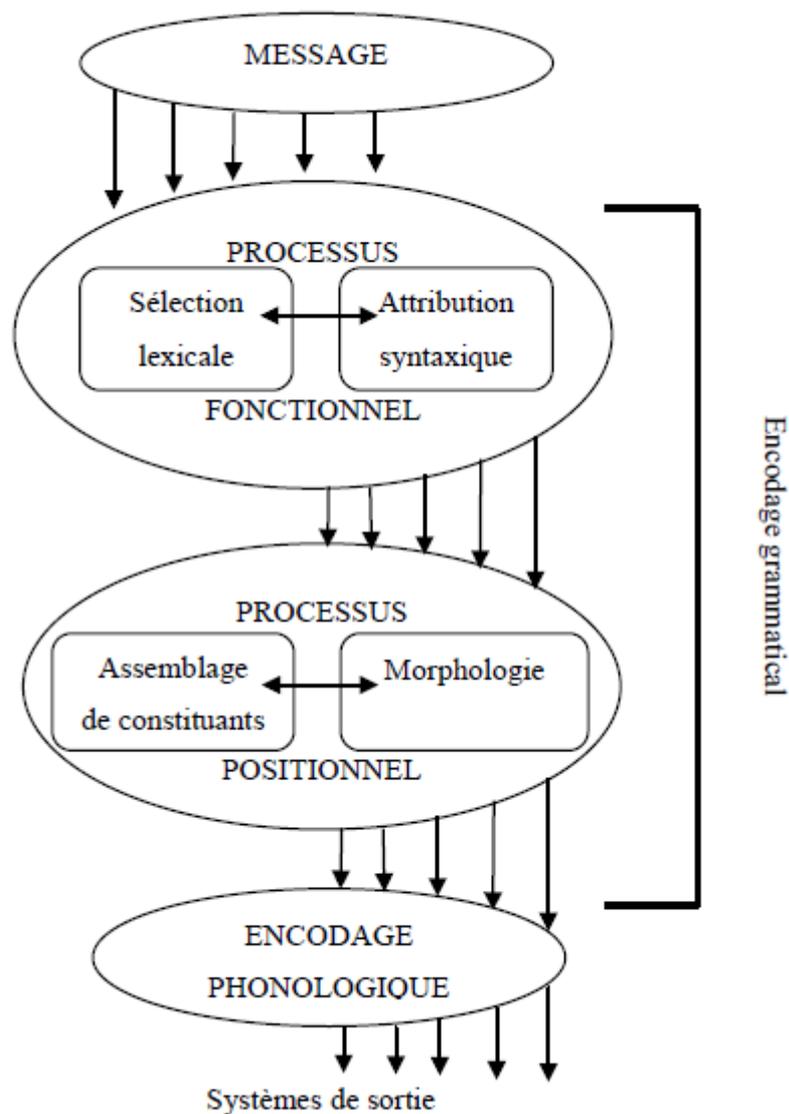
Le composant (6), *le système de compréhension du discours*, intègre des procédures de contrôle de la production langagière orale. Ce composant permet un retour du locuteur sur son message afin de corriger les erreurs éventuelles. La notion de *perceptual loop* est introduite (Levelt, 1989). Le locuteur peut effectuer un feedback sur son message avant et après l'entrée en action de l'*articulateur* ; c'est ainsi que sont distinguées l'*action externe* (quand la correction a lieu après la réalisation sonore du message) et l'*action interne* (quand la correction se déroule avant même la réalisation sonore du message). Levelt (1995) parle également de *self-monitoring* qui implique trois phases. *La première phase* concerne le

processus d'hésitation qui apparaît soit suite à une erreur de choix lexical soit suite à un manque d'information ; il s'interrompt donc. *La seconde phase* a trait à l'apparition de termes spécifiques signalant un problème à résoudre dans le message. Enfin, *la troisième phase* se définit comme la reprise du message engendrant la production de la correction de son propre message.

2.3.2.2. Encodage grammatical

La parole fluide repose sur l'interaction d'une multitude de composants dont la connaissance, pour le moment, est limitée (Levelt, 1995). Bock et Levelt (1994) proposent une étude plus approfondie de *l'encodage grammatical* du *formulateur* (Figure 8). Ce composant est davantage décrit dans la mesure où notre étude se concentre, entre autres, sur les choix syntaxiques en lien direct avec *l'encodage grammatical*.

Figure 8 : L'encodage grammatical de la production langagière (Bock et Levelt, 1994:946)



Le module de l'*encodage grammatical* est constitué de quatre niveaux principaux : le niveau du message, le niveau fonctionnel, le niveau positionnel et le niveau phonologique. *Le niveau du message* concerne l'élaboration conceptuelle du message.

Le niveau fonctionnel, pris en charge par le processeur sémantique (Leuwers, 2002), permet l'attribution d'une fonction syntaxique telle que *sujet* ou *objet direct* aux items lexicaux récupérés dans le *lexique mental*. Cette tâche implique l'identification des concepts lexicaux en adéquation avec le message que le locuteur souhaite transmettre. L'attribution des rôles syntaxiques se ferait selon le caractère plus ou moins accessible d'un élément : plus un référent est accessible, plus cet élément se verra attribuer une fonction syntaxique dite de haut niveau ; le sujet est hiérarchiquement plus haut que l'objet direct, qui est plus haut que l'objet

indirect, *etc.* (Bock et Warren, 1985 ; Bock et Levelt, 1994 ; Keenan et Comrie, 1977 ; Leuwers, 2002 ; Levelt, 1999). Les entités les plus accessibles seraient des référents animés plutôt que non animés, concrets plutôt qu'abstraites et enfin représentant une ancienne information plutôt qu'une nouvelle. L'ancienne information serait alors plus en position sujet qu'en position non sujet (Bock et Warren, 1985). Enfin, le lien entre l'accessibilité conceptuelle et l'attribution fonctionnelle se ferait également « *par l'intermédiaire des rôles évènementiels* » (Leuwers, 2002:114). Ainsi, les agents seraient plus en position sujet que les patients.

Ces informations sont, au *niveau positionnel* qui est pris en charge par le processeur syntaxique (Leuwers, 2002), soumises à un assemblage puis à un marquage flexionnel. Les éléments se voient fixer un ordre spécifique dans la clause et la morphologie appropriée leur est assignée. Si la position syntaxique est en lien étroit avec l'accessibilité des mots, il n'en est rien pour l'ordre des mots, qui ne dépend pas du caractère plus ou moins activé d'un élément (Bock et Warren, 1985). Un élément activé (soit une ancienne information) est un référent accessible qui se trouve préférentiellement en position sujet, soit en début de phrase. En français comme en anglais, ceci est lié au fait que le sujet est placé au début d'une phrase. Enfin, le dernier niveau de l'encodage grammatical est *le niveau phonologique* qui permet la concrétisation sonore du message.

Les éléments pragmatiques et sémantiques influencent les positions syntaxiques des unités langagières. Cette hypothèse est en lien direct avec notre orientation et une partie de nos résultats tente d'observer le rapport position syntaxique/aspects pragmatiques et formes linguistiques. Nous voulons vérifier que le statut informationnel d'un référent influence la position syntaxique et le choix de l'unité linguistique utilisée.

2.3.2.3. Modèle de la production orale : discussion

Un des aspects centraux du modèle de Levelt, auquel nous adhérons, est l'attention donnée au locuteur qui est au centre du langage. Le langage apparaît comme un instrument de communication que le locuteur utilise différemment selon le contexte et le but communicationnel de la situation d'énonciation. La prise en considération du contexte de production est essentielle (elle apparaît surtout au niveau du *conceptualisateur* et du *formulateur*). Le locuteur construit son message en fonction de la situation discursive, aspect

essentiel à la cohérence d'une production. Les choix linguistiques du locuteur sont, de plus, liés à la syntaxe, à la sémantique et à la pragmatique.

La pragmatique est mentionnée au niveau du modèle de discours et des connaissances encyclopédiques et de la situation : y sont regroupées les connaissances relatives à la situation d'énonciation (qui sont les participants du message, quel est la thématique, *etc.*). La pragmatique se retrouve également au niveau du *message préverbal* dans lequel interviennent également des notions comme la structure informationnelle.

Ainsi, selon les critères sémantiques, la situation pragmatique, le contexte de communication et l'intention communicative, le locuteur choisit un encodage spécifique et adapté. Nous pouvons étendre la notion de contexte de production aux différents types de textes produits par les participants de notre étude (Texte narratif/expositif, oral/écrit). Nous savons que, selon le type de discours, le locuteur mobilise des connaissances spécifiques. Lorsqu'il produit un texte, il sollicite des connaissances relatives à la structuration conventionnelle de ce type de texte et effectue le modèle mental de son message. Ainsi, selon le type de texte, les enfants de notre étude feront sans doute des choix linguistiques différents. De plus, la structure syntaxique du message se fait par le biais de la sélection d'items lexicaux dans le lexique mental tout en prenant en considération l'interprétabilité des référents. Le type de référent à introduire, son accessibilité, le contexte de production, *etc.* sont des facteurs influençant les choix linguistiques des individus (Levelt, 1989). L'utilisation d'un SNL ou d'un pronom dépend alors de la capacité linguistique du locuteur à gérer tous ces facteurs.

2.3.3. Les modèles de production : conclusion

Les modèles de Hayes pour la production écrite et celui de Levelt pour la production orale sont pertinents pour notre étude. En effet, un des aspects essentiels de ces modèles est l'importance du contexte de production qui joue un rôle majeur dans les choix linguistiques des individus (ceci est notamment mis en évidence dans le modèle de Levelt). Le contexte apparaît comme source de variation linguistique et la prise en considération du contexte lors de l'émission d'un message s'acquiert avec le temps (entre autres, Hickmann, 2002). L'enfant doit s'adapter au contexte et travailler la langue dans différentes dimensions : syntaxique, sémantique et pragmatique. Nous nous attendons alors à observer des choix linguistiques différents selon l'âge des individus et selon que les participants de notre étude produisent un texte expositif ou un texte narratif, un texte écrit ou un texte oral.

Une perspective discursive peut être donnée au modèle de Levelt (1989), notamment quant à la progression référentielle, en le complétant avec les règles de cohérence énoncées par Charolles (1978). Levelt considère que le locuteur introduit, réintroduit des référents et leur donne une forme linguistique selon leur degré d'activation et donc leur accessibilité. La construction du discours ne se fait pas aléatoirement et obéit à un certain nombre de règles. La cohérence d'un discours reposerait, entre autres, sur quatre principales méta-règles. Rappelons que pour qu'un texte soit cohérent, le texte doit comporter des éléments à récurrence stricte (*méta-règle de répétition*). Produire un texte cohérent exige également un apport en information sans cesse renouvelé (*méta-règle de progression*). Néanmoins, l'introduction de nouveaux référents ne peut se faire aléatoirement (*méta-règle de non-contradiction*). Enfin, les différents faits doivent être congruents (*méta-règle de relation*). Notre étude, à travers l'analyse de deux types spécifiques de marqueurs cohésifs, à savoir les SNL et les pronoms, va tenter de montrer comment les enfants gèrent la continuité référentielle dans des contextes de production différents.

2.4. Types de textes

2.4.1. Quel type de texte ?

Dans notre travail, le contexte de production renvoie aux deux modalités du langage, orale et écrite, ainsi qu'au type de texte. Nous avons décidé de travailler sur deux types de textes : la narration et l'exposition.

Ces deux types de textes appartiennent tous deux à la catégorie des textes *non procéduraux* à distinguer des textes *procéduraux* qui disent comment faire (Tijus, 2001). Nous pouvons donner comme exemple de textes procéduraux les notices techniques, les manuels d'utilisation ou les énoncés de résolution de problèmes. Les textes *non procéduraux*, pour lesquels une typologie est faite par rapport au contenu des textes, correspondent à tous les autres types de textes. Nous retrouvons le descriptif, le narratif et l'expositif (Brewer, 1980) et l'argumentatif (Tijus, 2001). Le texte descriptif est associé à la perception de scènes statiques ; le texte narratif est lié aux événements se déroulant dans le temps et reliés par des chaînes chronologiques ou causales ; le texte argumentatif vise à remporter l'adhésion d'interlocuteurs à un point de vue ; le texte expositif propose des points de vue en faisant intervenir des processus logiques abstraits (tels que l'induction, la classification ou encore la

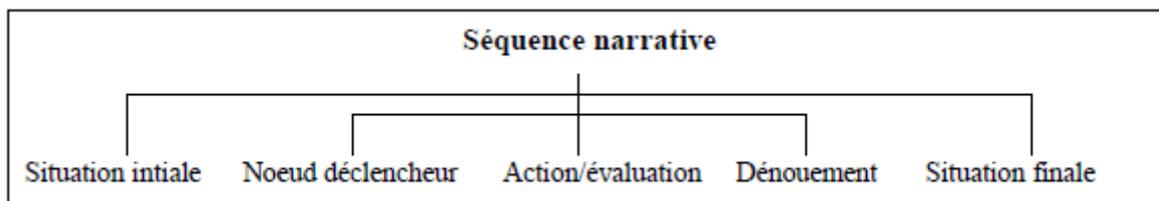
comparaison). Ce dernier type de texte représenterait le type de discours mettant le plus en jeu le langage académique, à savoir la *langue de scolarisation* (Boscolo, 1990 ; Johansson, 1999). Si le texte expositif et le texte narratif sont tous deux des textes *non procéduraux*, il n'en reste pas moins qu'ils caractérisent des discours spécifiques. S'opposent la narration et le discours logico-scientifique, reflétant *des modes de pensée bien distincts* (« *two distinct modes of thought* » de Ravid et Berman, sous presse ; Bruner, 1986), ce qui a des répercussions sur les choix linguistiques des locuteurs/scripteurs (Bruner, 1986 ; Ravid et Berman, sous presse). Il semble ainsi indispensable de les définir davantage.

2.4.2. Les types de textes narratif et expositif

2.4.2.1. La structure des textes narratif et expositif

Les types de textes narratif et expositif ont des structures différentes. Le texte narratif, prototypique est constitué des parties suivantes (Labov, 1976, 1978 ; Labov et Waletzky, 1967) : (a) le résumé, (b) les indications, (c) le développement, (d) l'évaluation, (e) le résultat ou la conclusion, (f) la chute. Un locuteur débute sa narration par un résumé puis, après avoir donné des indications sur le lieu, les personnes, l'activité, il enchaîne avec le récit proprement dit. Le locuteur termine sa narration par une conclusion et une chute, à travers laquelle il peut faire passer une morale d'ordre général. Tout le long de sa production, le locuteur emploie un procédé indiquant « *le propos de son histoire, sa raison d'être* » (Labov, 1978:471), les évaluations. Cette macrostructure narrative a été analysée par bon nombre de chercheurs, en se concentrant davantage sur le développement que sur les évaluations (Labov, 1978). Ainsi, nous trouvons d'autres schématisations de ce type de texte telles que celle d'Adam et Revaz (1996) qui parlent de macro-propositions narratives (Pn) (Figure 9). Cette organisation sous-jacente est censée guider « *la compréhension et la production des récits* » (Fayol, 1987:223 ; Mandler, 1982).

Figure 9 : Le schéma narratif canonique (Adam et Revaz, 1996), Pn = macro-proposition narrative



La figure ci-dessus présente les éléments essentiels de schéma narratif constitué d'une situation initiale, d'un nœud déclencheur, d'une action et/ou d'une évaluation, d'un dénouement et d'une situation finale dans laquelle nous pouvons retrouver par exemple une morale. L'activation de ces divers composants dépend évidemment autant « *de la connaissance du déroulement lui-même que de celle de celui à qui s'adresse la narration* » (Fayol, 2000:187).

Si le texte expositif est très difficilement schématisable, trois constituants à ce type de texte peuvent, néanmoins, être proposés (Katzenberger, 2004, 2005) : (a) l'introduction, qui aurait pour rôle d'introduire et de présenter le thème de la production ; (b) le corps de la production, dans lequel le scripteur développerait ses idées sur une thématique donnée ; et, (c) la conclusion, qui terminerait le texte. Prototypiquement, ces trois parties devraient être présentes dans les productions des locuteurs/scripteurs. Ces trois constituants impliquent différents mouvements : LA THÉMATISATION (EXPAND), L'EXPANSION (ENLARGE-ON), LA TRANSITION (MOVE-ON), L'UNIFICATION (UNITIZE) et L'ARRÊT (STOP) (Britton, 1994). LA THÉMATISATION est le mouvement principal de ce type de discours qui permet le développement d'une thématique. L'EXPANSION est un mouvement permettant de compléter LA THÉMATISATION en développant une idée précise. LA TRANSITION se situe en général entre deux EXPANSIONS ; il s'agit d'une transition entre un point précis et un autre. Il peut être introduit par des formes linguistiques telles que *le point suivant traite de ...*. L'UNIFICATION peut également suivre L'EXPANSION, il permet de réduire le contenu en mémoire dans le but de préserver l'unification du texte. Il peut apparaître sous une structure comme *en résumé, le phénomène X est causé par Y... .* L'ARRÊT clôture le discours.

Ces deux types de textes se définissent alors comme ayant des structures bien différentes. Cette différence est accentuée lorsque nous regardons plus précisément les caractéristiques de chacun de ces textes.

2.4.2.2. Spécificités et comparaison des deux types de textes

Ces deux types de textes s'opposent par leurs structures, leurs contenus et leurs buts communicationnels (Mosenthal, 1985) (Tableau 3).

[Tableau 3 : Présentation des oppositions entre le texte narratif et expositif]

Tableau 3: Présentation des oppositions entre le texte narratif et expositif

Narration	Exposition
Macrostructure essentiellement linéaire	Macrostructure non linéaire et moins identifiée
Organisation temporelle et causale	Organisation thématique
Appel à des informations spécifiques	Appel à des informations génériques
Forte implication du locuteur	Forte abstraction du texte et distance du locuteur
Registre académique non exigé	Registre académique
Forte mobilisation de la planification par connaissance rapportée	Forte mobilisation de la planification par connaissance transformée
Acquise précocement	Acquise plus tardivement

Les différences présentées dans le Tableau 3 peuvent être explicitées. Premièrement, le texte expositif se caractérise par une macrostructure non linéaire et moins clairement identifiée que celle du texte narratif (Gayraud, 2000). Cette absence d'une structure bien spécifiée a pour conséquence que la planification et l'organisation de ce texte sont assez coûteuses (Boscolo, 1990 ; Tolchinsky et *al.*, 1999). Par ailleurs, il s'agit d'un des textes les plus difficilement réalisables par les locuteurs et également d'un des plus difficilement représentables par le lecteur (Britton, 1994). Ceci se reflète dans le texte en lui-même (avec, par exemple, un manque de cohérence), ce qui entrave la construction du modèle mental du lecteur.

L'organisation linéaire du type de texte narratif répond à l'aspect antinomique de la production langagière. Tout message est verbalement linéaire alors que le modèle mental sous-jacent est, par définition, multidimensionnel. Il n'est en effet « *qu'exceptionnellement linéaire* » (Fayol, 1997:146). Le texte narratif serait un des seuls types de textes à avoir un modèle mental pouvant être linéaire. C'est la raison pour laquelle les enfants maîtriseraient plus précocement ce type de texte que tout autre (Bourdin, 1994 ; Fayol, 1984 ; Gayraud, 2000 ; Jisa, 2000a, 2000b ; Ravid, 2000 ; Ravid *et al.*, 2002 ; Wigglesworth, 1991).

Deuxièmement, le texte narratif se définit comme une suite de faits détaillés dans l'ordre chronologique (Aisenman et Assayag, 1999 ; Berman et Slobin, 1994a, 1994b ; Hickmann, 2003 ; Labov, 1978). Cette succession chronologique peut être définie comme « *une méthode*

de récapitulation de l'expérience passée consistant à faire correspondre à une suite d'événements (supposés) réels une suite identique de propositions verbales [...] Ce qui le (le récit) caractérise, c'est que les propositions y sont ordonnées temporellement » (Labov, 1978/1993:463). Cette organisation temporelle facilite la mise en texte d'un événement (Fayol, 1997). Les propositions d'une production narrative sont certes majoritairement liées temporellement mais elles peuvent également l'être causalement (Aksu-Koç et Küntay, 2001 ; Coirier *et al.*, 1996 ; Fayol, 2000 ; Jisa et Mazur, 2006).

En revanche, le texte expositif est une production dans laquelle les locuteurs créent une structure thématique, exposent des faits ou des idées théoriques et des interprétations de ces derniers (Boscolo, 1990 ; Britton, 1994 ; Katzenberger, 2004 ; Tolchinsky *et al.*, 2000 ; Tolchinsky *et al.*, 1999 ; Ravid et Berman, sous presse). Les individus peuvent procéder des faits aux interprétations ou *vice versa*. Ce type de texte, qui dérive de l'écrit académique (Mosenthal, 1985), requiert un déploiement de relations logiques, et non pas chronologiques, entre les arguments (Britton, 1994 ; Mosenthal, 1985 ; Ravid, 2005).

Troisièmement, lors d'une production narrative, les locuteurs/scripteurs se concentrent sur des gens, des actions, *etc.*, spécifiques (Berman et Nir-Sagiv, 2004 ; Berman et Slobin, 1994a, 1994b ; Berman et Verhoeven, 2002 ; Coirier *et al.*, 1996 ; Fayol, 1984, 1997, 2000 ; Peterson et McCabe, 1982). Le type de texte narratif sur lequel nous travaillons, sollicite un agencement de souvenirs personnels. Les locuteurs/scripteurs introduisent des référents spécifiques (un ou deux protagonistes) et les maintiennent tout le long de leur production.

Par opposition, le type de texte expositif fait appel à des informations d'ordre générique (Berman et Nir-Sagiv, 2004 ; Katzenberger, 2005 ; Mosenthal, 1985). Les locuteurs/scripteurs introduisent une information générique qu'ils développent, puis passent à d'autres thématiques, qu'ils maintiennent également. Ceci est lié à la pression informative de ce type de texte et son but communicationnel : développer des thématiques.

De par cette définition, le type de texte narratif exige, et ceci est notre quatrième point, que les locuteurs/scripteurs s'investissent personnellement dans leur narration (Adam, 1985) : nous parlons de discours orienté vers l'agent (Berman et Ravid, 2009 ; Berman et Nir Sagiv, 2007 ; Tolchinsky *et al.*, 1999). En effet, une des caractéristiques de ce texte est aussi « *la présence du sujet acteur qui se confond avec le narrateur dans le cas d'une narration personnelle* » (Gayraud, 2000:90).

En revanche, lors d'une production expositive, le locuteur doit avoir la capacité de se distancer, de se détacher de son texte (Berman et Nir-Sagiv, 2007 ; Johansson, 1999 ; Mosenthal, 1985 ; Ravid, 2005). Nous parlons de discours orienté vers le topic (Berman et Nir-Sagiv, 2007 ; Mosenthal, 1985 ; Tolchinsky *et al.*, 1999). Ceci implique une pression informative importante : le locuteur/scripteur doit nourrir la thématique abordée, ce qui peut se traduire par l'utilisation importante de lexique.

Cinquièmement, le type de texte expositif, beaucoup plus que le type narratif, exige la mobilisation du registre académique (type de registre utilisé, par exemple, dans les écrits scolaires et universitaires) (Johansson, 1999). Ce registre se caractérise par la mobilisation de diverses relations rhétoriques de haut niveau telles que la gestion de l'agencement des idées (Boscolo, 1990 ; Coirier *et al.* 1996 ; Mosenthal, 1985). Ce type de texte dispose alors de formes linguistiques complexes et plus coûteuses cognitivement, comme des sujets lourds (Ravid, 2005). Il requiert la mise en relation de deux niveaux : le niveau du contenu et celui de la rhétorique (Britton, 1994 ; Katzenberger, 2004 ; Mosenthal, 1985). De ce fait, sa planification et son organisation sont beaucoup plus exigeantes que celles du texte narratif (Boscolo, 1990).

Ainsi, et ceci est notre sixième point, les locuteurs/scripteurs, en contexte expositif, mobilisent sans doute davantage la planification par connaissances transformées que par connaissances rapportées, qui est davantage mobilisée lors d'une production narrative. Ces notions de planification sont davantage développées dans la Section 2.6 (p. 79).

En septième lieu, toutes ces particularités impliquent des différences de traitement et donc des différences de gestion selon le NIVEAU SCOLAIRE des locuteurs/scripteurs. Les informations, dans un texte narratif, sont organisées selon un schéma bien connu, appris relativement tôt et travaillé dès les premières années de scolarisation.⁴³ Les enfants connaissent très tôt certaines étapes de cette macrostructure narrative, telles que la situation initiale et le déclencheur, qui sont de surcroît travaillées à l'école (De Weck, 1991). Le locuteur/scripteur novice, dès 9-10 ans, est capable de produire une narration constituée de tous ces composants (Berman et Nir-Sagiv, 2007 ; Berman et Slobin, 1994a, 1994b ; Hickmann, 1995 ; Jisa, 2000 ; Peterson et McCabe, 1982).

⁴³ Ce type de texte a été longuement travaillé (entre autres, Adam, 1985, 1992 ; Adam et Revaz, 1996 ; Berman et Slobin, 1994a, 1994b ; Berman et Verhoeven, 2002 ; Coirier *et al.*, 1996 ; De Weck, 1991 ; Fayol, 1984, 1997, 2000 ; Gayraud, 2000 ; Jisa, 1998, 2000 ; Katzenberger, 1994 ; Katzenberger, 1999 ; Peterson et McCabe, 1982).

En revanche, le type de texte expositif n'est abordé que très tard dans le cursus scolaire. Ce type de texte serait maîtrisé assez tardivement par rapport au texte narratif (Berman et Katzenberger, 2004 ; Tolchinsky, Johansson et Zamora, 2002). Avec l'âge et l'évolution scolaire, le texte expositif est de mieux en mieux géré, les individus apprennent à passer des généralités à l'abstraction (Biber, Conrad et Reppen, 1998 ; Gayraud, *et al.*, 2000 ; Katzenberger, 1999, 2005). La structure du type de texte expositif est acquise plus tard par les enfants. Si l'introduction et le corps du texte sont en général produits, la conclusion apparaît plus sporadiquement dans les productions (Katzenberger, 2005). Ceci est surtout vrai pour les locuteurs/scripteurs les plus jeunes.

2.4.3. Types de textes : choix linguistiques différents ?

La reconnaissance des types textuels semble une acquisition tardive ne se mettant pas en place avant 12/13 ans (Benoît et Fayol, 1989 ; Coirier et Marchand, 1994 ; Fayol, 1991). Néanmoins, les enfants savent reconnaître des types textuels plus tôt que d'autres (Benoît et Fayol, 1989). Très tôt, ils font la différence entre les textes relevant de la narration et ceux n'en relevant pas. Lors de tâches de reconnaissance, il semble, qu'à 9 ans, les individus distinguent relativement bien les textes des non-textes. À 11 ans, ils savent isoler la description des narrations et des argumentations, ce dernier type étant seulement différencié des autres à 13 ans (Benoît et Fayol, 1989). Cependant, dès 9/10 ans, en production, les jeunes locuteurs/scripteurs, anglais, espagnol et israéliens, marquent eux-mêmes la différence entre un texte narratif (personnel) et un texte expositif (Berman et Nir-Sagiv, 2004 ; Reilly *et al.*, 2005).

Nous supposons que, pour les participants de notre étude, même les plus jeunes, le type de texte est source de variation. Nous nous attendons à ce que même les CM2 font des choix linguistiques différents selon le type de texte.

Par définition, ces deux types de textes impliquent des caractéristiques très différentes ayant des répercussions sur les choix linguistiques des locuteurs/scripteurs (Bruner, 1986 ; Ravid et Berman, sous presse). Le texte narratif se caractérise, entre autres, par l'utilisation de formes spécifiques (pronoms personnels), de noms concrets, de clauses simples et courtes (entre autres, Berman et Nir-Sagiv, 2007 ; Berman et Nir-Sagiv, 2004; Ravid, 2005, 2006).

En revanche, le texte expositif fait appel à des formes génériques (pronoms impersonnels et génériques), à des noms abstraits, à des clauses plus longues et à des structures syntaxiques

dites plus complexes (passif, subordination non finie, par exemple) (entre autres, Berman et Nir-Sagiv, 2007 ; Berman et Nir-Sagiv, 2004 ; Johansson, 1999 ; Ravid, 2005, 2006).

Deux exemples illustrent ces différences : (58) est un texte narratif et (59) un texte expositif.

(58) Un jour, j'étais partie chez ma grand-mère avec ma petite sœur qui avait deux ans (j'en avais six à l'époque). Là-bas, j'ai retrouvé mes deux cousines Ismat et Shaïna. Elles sont toutes les deux plus âgées que moi. Comme nous sommes très complices, nous avons prévu une partie de cache-cache. Au bout d'un quart d'heure, je n'ai plus voulu jouer. Mes cousines ont alors arrêté de me parler et j'étais très chagrinée. Puis, je me suis effondrée et je leur ai demandé de faire la paix. J'ai alors vu qu'elles étaient contentes. Elles se sont alors excusées pour leur comportement et tout s'est bien arrangé. Je sais que c'est peut-être banal mais, lors de cette scène, j'ai vraiment senti un sentiment d'abandon, de solitude. (5^{ème}, n°06, narratif écrit, E/O)

(59) La dispute éclate lorsque plusieurs personnes ne sont pas d'accord. Le meilleur moyen de résoudre ce problème est la paix ou l'ignorance de l'autre. Parfois, les gens concernés font appel à la violence qui a plutôt tendance à envenimer la situation. Verdict : la violence ne résout rien. Lorsque je me dispute avec quelqu'un, j'essaie d'abord de faire en sorte que mon « adversaire » se calme. Alors après, j'utilise le système de la parole : chacun parle à son tour sans que l'autre intervienne comme ça, on a moins de mal à s'exprimer au lieu de crier tout en même temps. De cette manière, on comprend mieux l'idée de cette personne. On peut donc se disputer calmement. Voilà. Voilà ce que je sais de la dispute et comment l'éviter ou « l'alléger ». (5^{ème}, n°06, expositif écrit, E/O)

Dans sa narration, l'enfant raconte une histoire personnelle en faisant appel à des informations spécifiques. Il s'agit d'un événement spécifique (*une dispute*), ayant eu lieu dans un endroit spécifique (*chez sa grand-mère*), avec des personnes spécifiques (*ses deux cousines, Ismat et Shaïna*), qui participent à une action particulière (*une partie de cache-cache*). Le discours est orienté vers l'agent, le scripteur s'investit dans son texte ; ceci se voit clairement à travers l'utilisation du déictique *je* et du pronom personnel *elles* faisant référence *aux cousines*. Les clauses sont également simples et la majorité des noms représentent des entités concrètes (*sœur, grand-mère, cache-cache*).

Ce même enfant, en production expositive, se penche sur la thématique de la dispute d'une manière objective. L'adolescente fait appel à des structures traduisant l'aspect général et impersonnel de son texte. Le discours est orienté vers le contenu. Ceci se voit notamment à

l'utilisation de formes génériques (*personnes, gens, chacun, etc.*), de noms abstraits (*dispute, moyen, paix, ignorance, violence, idée*), de structures impersonnelles (*on*), de clauses complexes (par exemple, subordonnée non finie, *le meilleur moyen de résoudre ce problème*). Le déictique *je* est également beaucoup moins utilisé, le locuteur oriente son discours vers le *topic*, et se distancie donc de sa production en utilisant, entre autres, le pronom générique *on*.

Nous pouvons voir, à travers l'observation très rapide de deux exemples, que ces types de textes, narratif et expositif, engendrent des choix linguistiques très différents. Nous nous attendons alors à ce que des différences ressortent également quant au traitement du SN et de l'information. Ainsi, par exemple, le texte expositif de l'adolescente (58) contient davantage d'items lexicaux que le texte narratif (58).

2.5. L'interface type de texte/modalité

Dans les parties précédentes de ce chapitre, ont été abordées diverses notions en lien étroit avec le contexte de production. Rappelons que, par contexte de production, nous entendons, dans cette étude, modalité de production (oral/écrit) et type de texte (narratif/expositif).

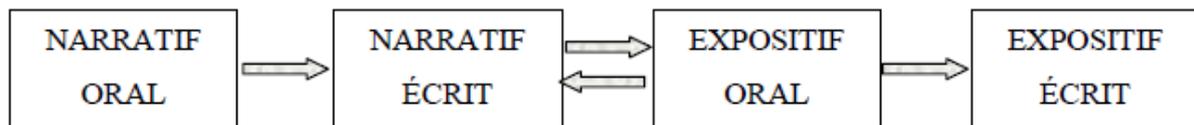
Trois notions sont étroitement liées au contexte : le registre, le genre et le mode (Berman, 1999a, 1999b ; Halliday, 1985 ; Halliday et Hasan, 1976/1989 ; Ravid, 2000 ; Ravid et Tolchinsky, 2000, 2002). Ces trois dimensions – registre, genre et modalité - sont essentielles à la bonne compréhension de la production textuelle dans le sens où chacune d'entre elles implique des variations quant aux choix linguistiques des individus (Berman, 1999a, 1999b ; Berman, 2000 ; Berman et Katzenberger, 2004 ; Cahana-Amitay et Katzenberger, 2000 ; Gayraud, 2000 ; Gayraud *et al.*, 1999, 2001 ; Jisa, 2000a, 2000b, 2000c, 2004 ; Jisa *et al.*, 2002 ; Jisa et Viguié, 2005 ; Johansson, 1999 ; Katzenberger, 1999, 2004 ; Ravid *et al.*, 2002 ; Tolchinsky *et al.*, 1999 ; Zamora et Kriz, 2000).

La modalité et le type de texte, qui sont deux facteurs intra-individuels de notre étude, sont liés aux dimensions du registre (Ravid et Tolchinsky, 2002). Selon les définitions, nous pouvons aisément penser à une échelle de registre quant aux différents types de textes produits. Ainsi, le type de texte narratif oral se caractériserait par un registre moins académique que le texte expositif écrit (Ravid et Berman, sous presse). Le texte expositif écrit est le type de texte mettant le plus en jeu les habilités linguistiques de haut niveau (Ravid et Berman, sous presse).

Cet aspect académique de la production s'exprime, par exemple, par l'utilisation massive de lexique (Blanche-Benveniste, 1990, 1995), par la présence de subordonnées non finies (Jisa, 2004) ou encore par la production de SNL lourds (constitués de plus d'un item lexical) (Ravid, 2000 ; Ravid et Berman, sous presse ; Ravid et *al.*, 2002). Ces formes sont donc davantage retrouvées dans les productions expositives écrites.

Nous pouvons alors parler d'interface entre ces deux aspects du contexte, représenté par la Figure 10 (Berman et Ravid, 2009).

Figure 10 : L'expression linguistique à travers l'interface entre la modalité et le type de texte (Berman et Ravid, 2009:106)



Cette figure présente la hiérarchie linguistique pouvant exister entre ces différents types de textes. Les productions narratives orales et expositives écrites en constituent les deux extrêmes, les deux autres types apparaissant comme des textes intermédiaires.

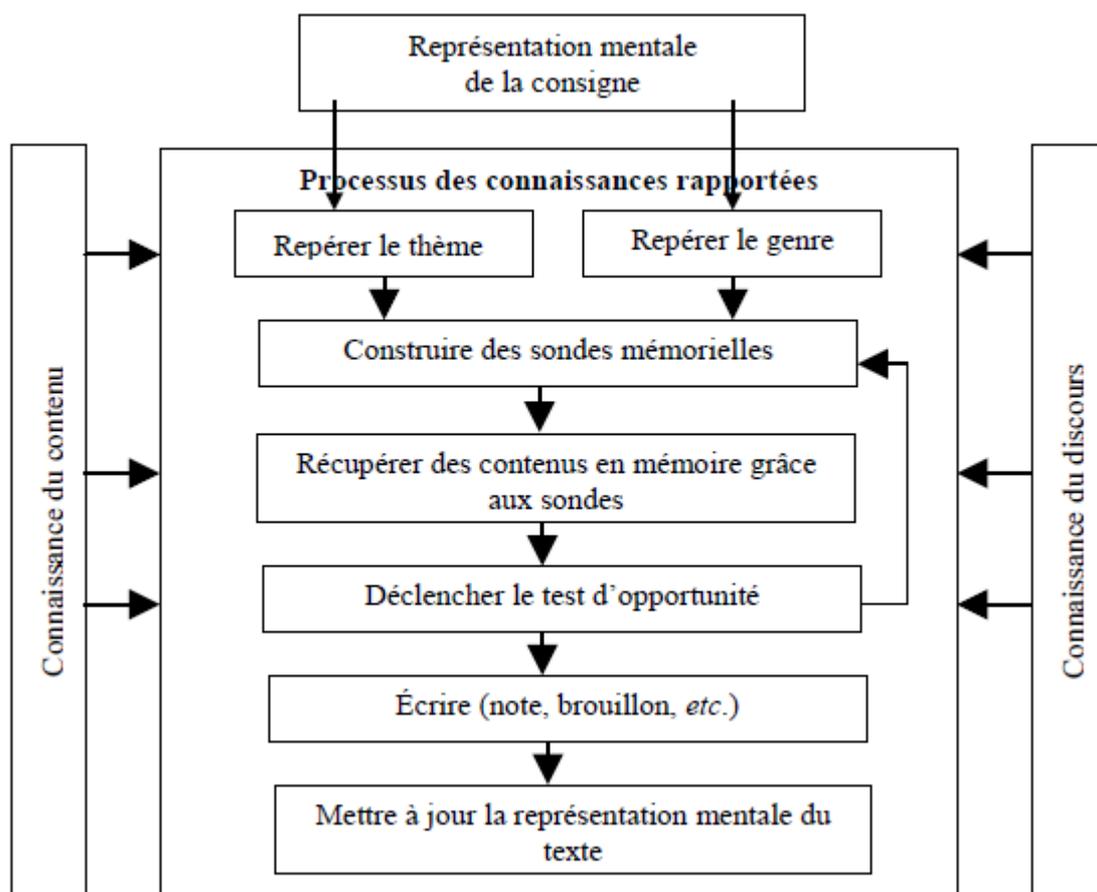
2.6. Production textuelle, planification et développement

Lors d'une production de texte, le locuteur fait appel, entre autres, à des modèles mentaux structurés. Rappelons qu'il s'agit de représentations cognitives regroupant des informations et des connaissances relatives à une situation spécifique (Hoc, 1992:72). Les individus lors de la verbalisation de leur message doivent faire face à une situation paradoxale : le message doit être linéaire alors que le modèle mental est, par définition, multidimensionnel. Pour répondre à la contrainte engendrée par l'aspect multidimensionnel de la représentation mentale et l'aspect linéaire du langage, les individus ont recours à trois stratégies de planification : la planification par connaissances rapportées (*knowledge telling strategy*), la planification par connaissances transformées (*knowledge transforming strategy*) et la planification par connaissances révisées (*knowledge crafting*).

2.6.1. Types de planification

Les deux premières stratégies de planification sont proposées par Scardamalia et Bereiter (entre autres, Bereiter et Scardamalia, 1987, 1998 ; Bereiter, Burtis et Scardamalia, 1988). *La planification par connaissances rapportées* est modélisée par la Figure 11.

Figure 11 : Stratégie des connaissances rapportées (selon Bereiter et Scardamalia, 1987 et 1998 ; Traduction de Chanquoy et Alamargot, 2002:180)

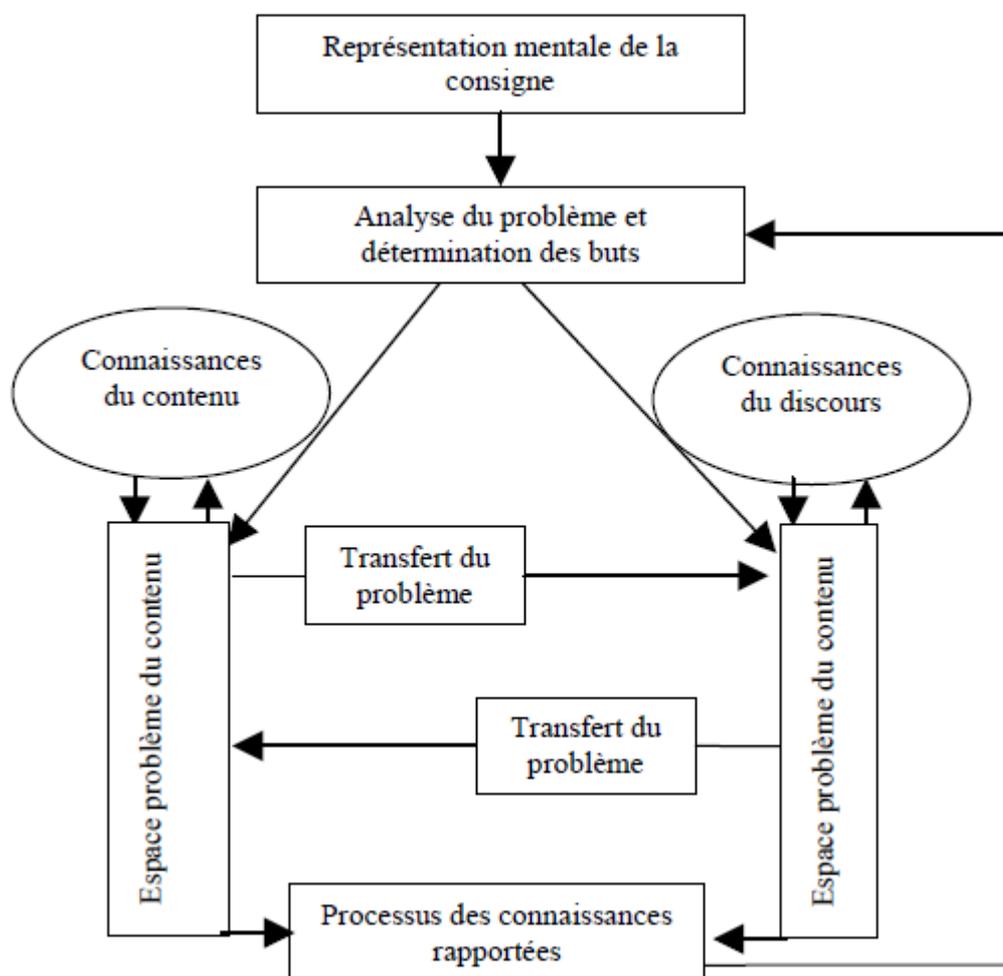


Ce type de stratégie correspond à un travail de récupération des contenus, des informations, pas à pas, dans la MLT ; puis, ces derniers sont reformulés au fur et à mesure qu'ils sont récupérés « *sans procéder à une réorganisation d'ensemble du contenu conceptuel ou de la forme linguistique du texte* » (Chanquoy et Alamargot, 2002:7). Ce type de planification se caractérise par des cycles répétitifs de récupération et de formulation (McCutchen, 1996).

Cette manière de procéder, constituer un texte pas à pas, permet aux enfants de produire un texte malgré le fait que la plupart des ressources cognitives sont consacrées à de multiples demandes de l'activité rédactionnelle (travail grapho-moteur). Cette planification permet aux scripteurs novices de prendre en considération certains aspects du contexte de production tels que le thème ou le genre du texte. Ils récupèrent leurs connaissances du contenu et du discours, prennent en considération le thème et le genre, et mettent en mot ces aspects pour concrétiser leurs représentations mentales.

La planification par connaissances transformées (Figure 12) appelle à un effort de mise en relation de deux espaces : celui des contenus (permettant l'accessibilité et l'organisation des connaissances) et celui de la rhétorique (définissant les contraintes de buts compte tenu de la situation) (Scardamalia et Bereiter, 1987, 1998).

Figure 12 : Stratégie des connaissances transformées selon Bereiter et Scardamalia, 1987 et 1998 ; Traduction de Chanquoy et Alamargot, 2002:181)



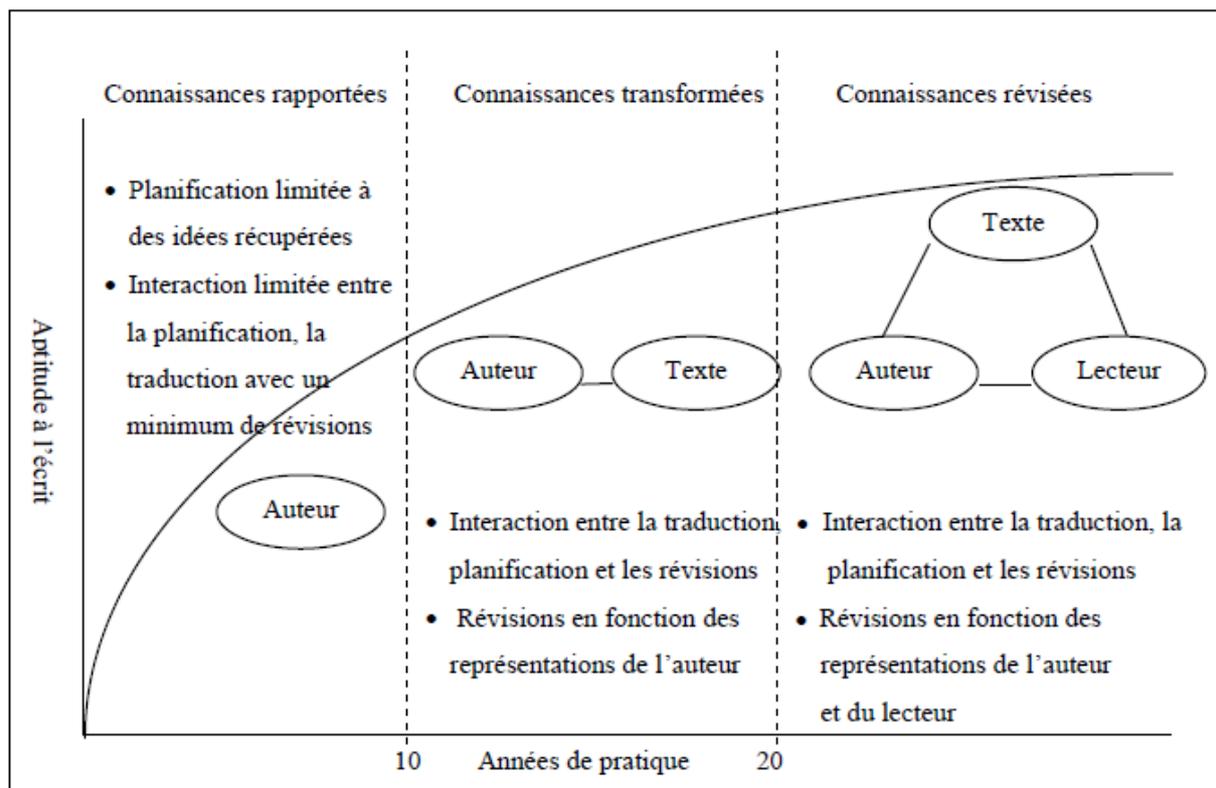
Les connaissances récupérées en mémoire sont réorganisées en prenant en compte plus de paramètres linguistiques et pragmatiques. Selon les buts communicationnels, les locuteurs/scripteurs ajustent leur message (Alamargot et Chanquoy, 2002 ; Chanquoy et Alamargot, 2002). La tâche linguistique est analysée tel un problème à résoudre. Ainsi, ce type de planification ajoute la prise en considération de ce problème (composant : *analyse du problème et détermination des buts*). L'organisation des connaissances suppose une planification incluant de nombreuses connaissances relatives aux domaines, aux contraintes rhétoriques, linguistiques et pragmatiques de la langue. Tous ces aspects sont pris en considération et sans cesse mis en relation avec la tâche linguistique à effectuer, le problème. Ainsi, *les connaissances du contenu* et *celles du discours* sont toujours liées au problème. Le locuteur/scripteur doit faire le lien entre les connaissances du contenu, portant sur les domaines, les champs d'études et les connaissances du discours. Ces dernières concernent les aspects syntaxiques (par exemple l'ordre des mots dans la phrase), les structures de textes et

les connaissances pragmatiques (telles que les règles d'utilisation de la langue selon le contexte de production). Les scripteurs entrent dans un travail de formulation et reformulation qui peut être réalisé tant que le texte ne satisfait pas la consigne, c'est-à-dire tant que la production ne répond pas au problème.

Ce type de capacité rappelle le modèle de Levelt qui intègre syntaxe, sémantique et pragmatique. La langue offre aux locuteurs/scripteurs diverses structures pour encoder un même type de message. Il choisit alors une structure linguistique spécifique en fonction des aspects pragmatiques de la situation de production et de son intention communicative (Levelt, 1989, 1999).

Une troisième stratégie est suggérée, *la planification par connaissances révisées (the Knowledge-Crafting)* (Figure 13) (Kellogg, 2008).

Figure 13 : Les macro-stages du développement des aptitudes à l'écrit (Kellogg, 2008:4)



Ce troisième niveau intègre la prise en considération des représentations mentales de l'interlocuteur ; il est atteint lorsque l'individu est capable de faire appel à *la planification par*

connaissances rapportées et à la *planification connaissances transformées* en prenant en considération les représentations mentales du destinataire du message.

2.6.2. Planification et développement

Selon le contexte de production (MODALITÉ et TYPE DE TEXTE) et/ou le NIVEAU SCOLAIRE des locuteurs, une stratégie (*par connaissances rapportées, transformées et révisées*) peut être plus sollicitée qu'une autre. *La planification par connaissances rapportées* caractérise la procédure de planification des individus novices alors que *la planification par connaissances transformées* est une caractéristique de locuteurs plus experts. *La planification par connaissances révisées* serait mobilisée assez tardivement.

Selon Chanquoy et Alamargot (2002:7), les jeunes enfants, contraints par les multiples demandes du processus rédactionnel,

« ne peuvent mettre en œuvre qu'une gestion pas à pas de leur production, en ne prenant en compte que certains paramètres rédactionnels (comme le genre et le thème du texte) et en en délaissant d'autres (comme le destinataire et le contexte de la production). Ceci les conduit à rédiger des textes se présentant comme des juxtapositions d'énoncés, sans liens clairs. Le texte est donc élaboré sans réelle réorganisation des informations récupérées et de manière assez économique. »

Ceci caractérise la planification par connaissances rapportées. Dans ce même esprit, Berninger et Swanson (1994) proposent une vision développementale de l'activité rédactionnelle.⁴⁴ Ils suggèrent que les ressources cognitives sont partagées entre différents processus. Les enfants, aux cours de leur acquisition de la langue écrite, attribuent plus de ressources à un processus spécifique plutôt qu'à un autre. La pratique et l'expérience permettent d'automatiser certains processus (comme celui de l'écriture), exigeant donc moins de ressources cognitives, qui sont alors allouées à des processus de haut niveau (comme ceux des révisions). C'est dans cette mesure que les individus, en automatisant certains processus, peuvent mobiliser davantage *la planification par connaissances transformées* (Bereiter et Scardamalia, 1987).

Trois phases de développement de l'enfant à l'écrit, décrivant « trois phases de développement de la production écrite, chacune étant caractérisée par un type de contrainte pesant sur la mémoire de travail » (Favart et Olive, 2005:278), sont proposées (Berninger et

⁴⁴ Ceci *via* le concept de MDT s'inspirant des travaux de McCutchen (1996, 1998, 2000) qui se base sur La *Théorie de la Capacité*, développée par Just et Carpenter (1992). Ce modèle développemental de l'activité rédactionnelle se base sur le modèle de Hayes et Flower (1980).

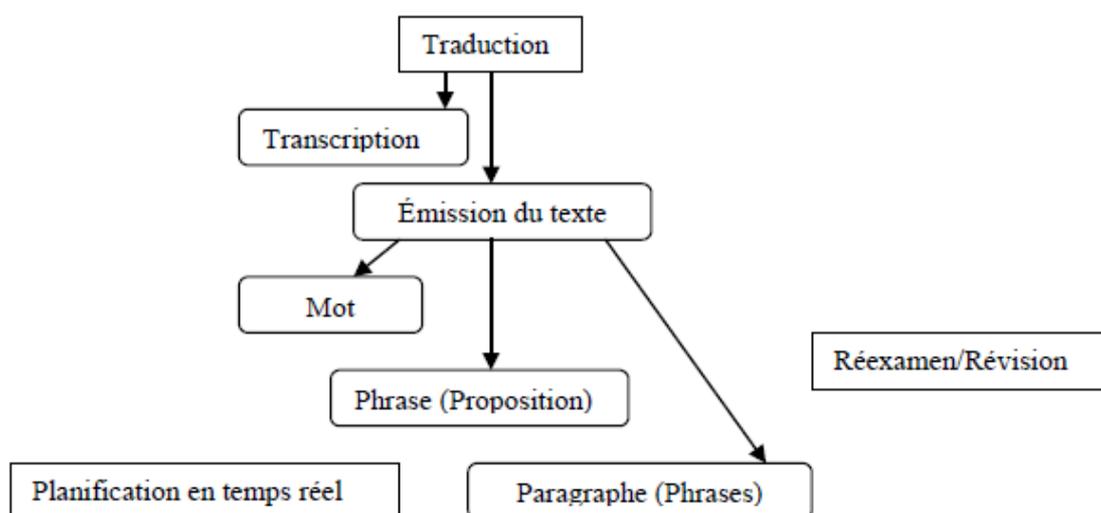
Swanson, 1994). Ces trois phases peuvent être mises en relation avec les types de planification mobilisée par les individus. Favart et Olive (2005:279) résument ces trois phases développementales de l'activité rédactionnelle en disant que :

« la mise en texte occupe le devant de la scène lors de la première phase, la mémoire de travail assure sa coordination avec la révision dans la seconde phase, et enfin avec la planification dans la troisième. [...] Enfin, parallèlement à cette interactivité croissante des trois processus, le jeune scripteur peut, lors de sa production, opérer sur des unités langagières de taille croissante : le mot puis la phrase, le paragraphe, et enfin le texte dans son ensemble. ».

Ainsi, selon la phase et selon les ressources allouées à certains processus, les enfants utilisent davantage une planification plutôt qu'une autre.

La première phase, schématisée par la Figure 14, caractériserait des enfants de CP aux CE2 (à savoir, entre 6 et 8 ans).

Figure 14 : Développement de la rédaction aux premiers niveaux scolaires (selon Berninger et Swanson, 1994 ; Berninger, Yates, Cartwright, Rutberg, Remy et Abbott, 1992 ; Traduction de Piolat, 2004:36)



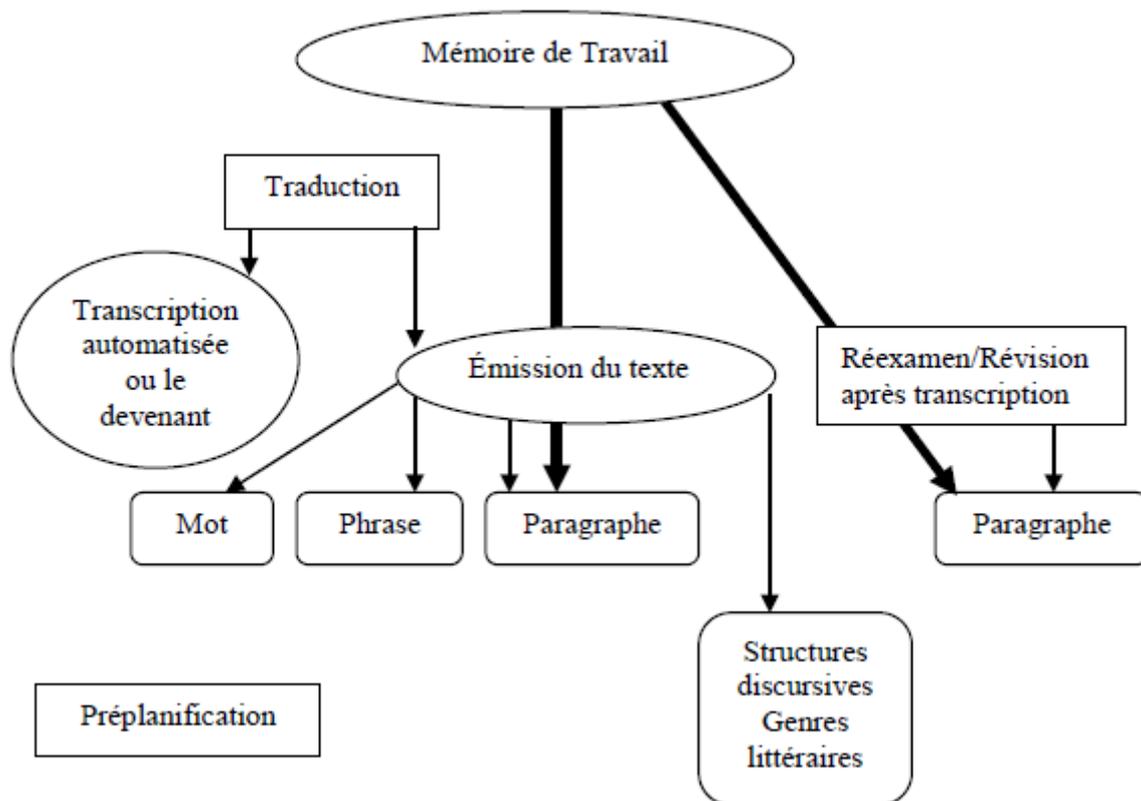
À ce premier niveau, les processus de bas niveau tels que le codage orthographique ou les processus grapho-moteurs permettant l'acte d'écriture affectent l'écrit des enfants (Berninger, *et al.*, 1992 ; Berninger et Swanson, 1994). Les ressources cognitives sont davantage allouées aux processus : (a) de *traduction* (ou de formulation) consistant à la transformation de la représentation mentale en représentations linguistiques et (b) d'*émission* (ou de génération)

assurant la transformation des représentations linguistiques en graphèmes. Les ressources cognitives étant mobilisées pour gérer la *traduction* et l'*émission du texte*, le *processus de planification* se limite à l'enchaînement des propositions et celui de *révision* à des corrections dites strictement locales (Berninger et Swanson, 1994 ; Favart et Olive, 2005 ; Piolat, 2004). Les enfants ne sont également pas conscients des enjeux de l'écrit. Aussi, des études ont montré qu'une aide supplémentaire (telle que de l'entraînement à écrire) accordée à des enfants ayant des difficultés liées à l'apprentissage de l'écriture et de l'orthographe, entraînait une amélioration de l'écriture et de la composition des phrases⁴⁵ ; même si Graham *et al.* (2000) et Fayol et Miret (2005) nuancent cet effet. Ainsi, lors de cette première phase, les sujets novices anticipent alors peu sur le contenu du message et sur la manière dont ils vont le communiquer (Fayol, 1990). De plus, les ressources cognitives sont davantage allouées aux processus de traduction et d'émission, nous pouvons alors penser que les jeunes enfants ont recours principalement à *la planification par connaissances rapportées*.

Puis ces composantes de traduction et émission du texte s'automatisent et l'enfant passe à une seconde phase où il attribuerait davantage les ressources cognitives à un nouveau composant, celui de la révision. Cette deuxième phase du développement de la rédaction des textes, qui est présentée par la Figure 15, caractérise des jeunes scripteurs scolarisés entre le CM1 et la 6^{ème} (à savoir, entre environ 9 et 11 ans).

Figure 15 : Développement de la rédaction aux niveaux scolaires intermédiaires (selon Berninger, Cartwright, Yates, Swanson et Abbott, 1994 ; Berninger et Swanson, 1994 ; Traduction de Piolat, 2004:63)

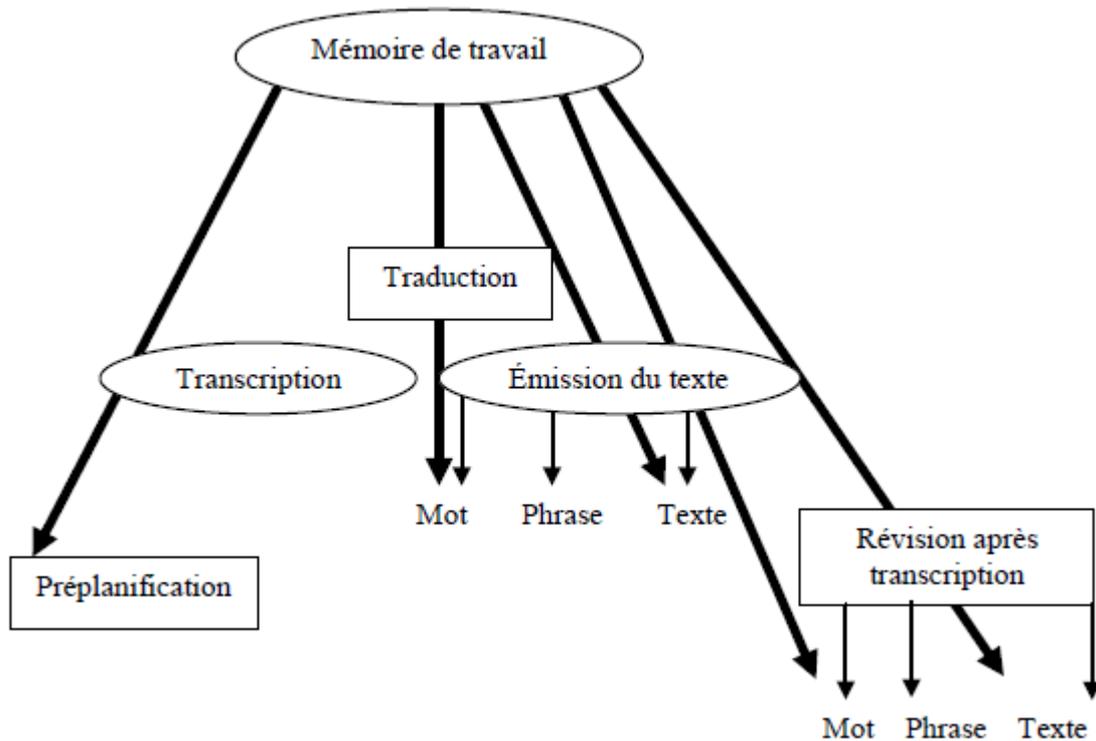
⁴⁵ Cf. entre autres : Berninger, Fuller, Whitaker (1996) ; Berninger, Vaughan, Abbott, Abbott, Rogan, Brooks, Reed et Graham (1997) ; Berninger, Vaughan, Abbott, Begay, Coleman, Curtin, Hawkins, Graham (2002) ; Berninger, Vaughan, Abbott, Brooks, Abbott, Reed, Rogan, Graham (1998) ; Graham, Harris, Chorzempa (2002) ; Graham, Harris et Kink (2000) ; Jones et Christensen (1999).



Au cours de cette seconde phase, les individus ont automatisé le processus de *traduction* et commencent à développer celui de la *planification*. Les révisions sont encore rares, même si elles commencent à dépasser le niveau orthographique du mot pour celui de la phrase ou du paragraphe. Les locuteurs/scripteurs commenceraient à produire un texte en prenant en considération le but et le contexte communicationnel. Ils deviendraient alors conscients que le texte doit être sans cesse travaillé pour représenter au mieux leur modèle mental et être adapté à la situation communicationnelle. Ce comportement laisse entendre que les individus de 9 à 11 ans commenceraient à mobiliser *la planification par connaissances transformées*. Néanmoins, à cet âge, ils mobilisent encore davantage *la planification par connaissances rapportées* (Bereiter et Scardamalia, 1987 ; Piolat, 2004).

Enfin, après avoir automatisé *le processus de traduction et de révision*, les enfants passeraient à une troisième phase de développement de la rédaction des textes (Figure 16). Cette dernière phase représente des enfants scolarisés de la 5^{ème} à la 3^{ème} (à savoir, entre environ 12 et 15 ans).

Figure 16 : Développement de la rédaction aux niveaux scolaires élevés (d'après Berninger et Swanson, 1994 ; Traduction de Piolat, 2004:64)



Durant cette troisième phase, les individus sont plus à l'aise face à la gestion de tous les processus à déployer lors d'une tâche de rédaction. Ils font également de plus en plus de révisions. Le processus de révision s'automatise. Les enfants développent la capacité de mobiliser *la planification par connaissances transformées* (Bereiter et Scardamalia, 1987 ; Chanquoy et Alamargot, 2002 ; Fayol, 1997 ; Olive et Piolat, 2005 ; Piolat, Isnard et Della-Valle, 1993 ; Piolat et Roussey, 1992). Néanmoins, le processus de planification n'est ni géré ni coordonné avec tous les autres composants. Comme le rappelle Piolat (2004:64), « *cela viendra plus tardivement vers 16 ans* ».

Avec l'âge, les sujets automatisent certains processus, tels que la transcription, et décident de donner plus de ressources cognitives à un processus (celui de la révision, par exemple) plutôt qu'à un autre, selon le type de tâche qui leur est demandée. Ils font preuve d'adaptation et de flexibilité. L'expertise en production écrite se caractérise comme une gestion des processus engagés dans l'activité langagière et par le fait d'utiliser à bon escient tous les processus de révision (impliquant lecture et relecture de sa production) pour assurer une bonne cohérence et cohésion de son texte (Alamargot *et al.*, 2005 ; Bourdin, 2002 ; Olive et Piolat, 2003).

Plus tard, les individus mobilisent *la planification par connaissances révisées* : ils deviennent capables d'assurer une certaine cohérence entre leurs idées, leurs mises en mots tout en prenant en considération le fait que leurs propres représentations mentales diffèrent de celles des destinataires de leurs messages. Durant les premières années de pratique, les représentations mentales du locuteur/scripteur dominent le message. Les novices font preuve d'*égocentrisme cognitif*,⁴⁶ ils ne prennent pas forcément en compte les représentations mentales des destinataires. Ceci se met en place tardivement et exige la mobilisation de *la planification par connaissances révisées* (Kellogg, 2008).

Nous pouvons imaginer que les enfants de notre étude, scolarisés en CM2, font plus appel à *la planification par connaissances rapportées*. Les élèves de 5^{ème} et de 3^{ème} mobiliseraient de plus en plus *la planification par connaissances transformées*, en pensant que les élèves de 3^{ème} prendraient de plus en plus en considération les représentations mentales des destinataires des messages. Les enfants les plus âgés adapteraient davantage leurs choix linguistiques à la situation de communication. Ainsi, la production de SNL, et notamment complexes et/ou en position sujet, augmenterait avec l'âge tout comme les stratégies linguistiques comme le maintien de l'information sous forme lexicale.

Pour conclure, ces types de planification (*par connaissances rapportées, transformées et révisées*) fonctionnent comme un continuum (Chanquoy et Alamargot, 2002) : les individus passent progressivement d'une stratégie à l'autre au cours de leur développement. De plus, les experts mobilisent ces trois types de planification. Mais, selon le contexte, ils ont recours davantage à une stratégie plutôt qu'à une autre. *La planification par connaissances rapportées* peut être activée par des individus experts dans des situations dénuées de tout enjeu académique ou rédactionnel (Alamargot et Chanquoy, 2002), dans des situations particulières comme la rédaction de cartes de vœux « *qui [selon McCutchen (1988)] ne suppose pas une transformation majeure des contenus* » (Chanquoy et Alamargot, 2002:9). Le type de planification sollicitée dépendrait alors également du contexte de production. Ainsi, par exemple, le type de texte narratif ferait davantage appel à *la planification par connaissances rapportées*. En effet, en narration, les individus produisent en connaissant tout le contenu du futur texte et ils ont à lier les événements temporellement. Rappelons que tout message verbal est linéaire alors que le modèle mental sous-jacent est multidimensionnel. Le texte narratif est un des seuls types de textes à faire exception et à avoir un modèle mental

⁴⁶ Ces notions (représentations mentales/égocentrisme cognitif) sont abordées dans le Chapitre 1 (p. 27/28).

linéaire. *La planification par connaissances rapportées* suffirait à répondre à cette organisation. En revanche, le type de texte expositif ferait plus appel à *la planification par connaissances transformées* qu'à *la planification par connaissances rapportées*. En effet, rappelons que ce type de texte, qui dérive de l'écrit académique (Mosenthal, 1985), sollicite un déploiement de relations logiques, et non pas chronologiques, entre les arguments. Il exige la mobilisation du registre académique se définissant par la mobilisation de diverses relations rhétoriques de haut niveau telles que la gestion de l'agencement des idées mettant en relation deux niveaux : le niveau du contenu et celui de la rhétorique. De ce fait, ce type de texte est beaucoup plus planifié et travaillé que le texte narratif.

Bien que ces stratégies de planification concernent la production écrite, elles s'adaptent également à la production orale. Nous pouvons alors nous demander si le fait d'utiliser *la planification par connaissances rapportées* est seulement dû aux contraintes rédactionnelles de l'écrit. Si c'était le cas, à l'oral, les enfants seraient capables de mobiliser les deux types de planification. Or, nous pensons que ce n'est pas le cas d'autant plus que *la planification par connaissances rapportées* reviendrait à utiliser les habilités conversationnelles de l'oral pour les tâches écrites (Chanquoy et Alamargot, 2002). Ainsi, le fait d'avoir un message moins travaillé et moins planifié, caractéristique d'un oral prototypique, se retrouve dans les écrits des novices. À l'oral, tout comme à l'écrit, les individus peuvent passer par ces différentes phases, car, outre le processus de *traduction*, à remplacer par l'articulation, les processus de production textuelle à l'oral font appel à des processus communs tels que la *planification* ou *l'émission du texte*.

Chapitre 3

Problématique et questions de recherche

3.1. Problématique générale

La production langagière implique quatre niveaux : *le niveau cognitif, lexico-grammatical, lexico-phonologique* et *le niveau articulatoire ou grapho-moteur* (Coirier et al., 1996). Cette

étude souhaite contribuer à la compréhension du *niveau cognitif* ainsi que celui *lexico-grammatical* de la production textuelle. Les langues offrent une multitude de combinaisons possibles pour encoder un même message ; ce qui est intéressant est de savoir quels sont les critères poussant les locuteurs à choisir une forme plutôt qu'une autre. Le modèle de Levelt (1989) ou encore la conception de la littéracie linguistique de Ravid et Tolchinsky (2000, 2002) suggèrent l'importance de la prise en considération du contexte de production et de la pragmatique. L'individu, selon le contexte et son intention communicative, mobilise des connaissances spécifiques. Ainsi, par exemple en production narrative, il sollicite des connaissances relatives à ce type de texte. Puis, tout en prenant en considération la situation discursive, il introduit et maintient des référents. Pour organiser les contenus activés, le locuteur fait appel à deux types de planification : la *planification par connaissances rapportées* et la *planification par connaissances transformées* (entre autres, Bereiter et Scardamalia, 1987, 1998 ; Bereiter, Burtis et Scardamalia, 1988).

Problématique Ainsi, comment les enfants deviennent-ils des locuteurs/scripteurs performants ? Comment ce développement se traduit-il linguistiquement ? Le contexte de production a-t'il un impact sur les choix linguistiques de nos participants, tout comme les adultes (Levelt, 1989) ? Est-ce que les enfants et adolescents de notre étude, selon le type de texte à produire (narratif oral, narratif écrit, expositif écrit, expositif oral), adaptent leurs choix linguistiques et mobilisent une planification plutôt qu'une autre, comme cela l'est suggéré par le modèle de Levelt (1989), par Bereiter et Scardamalia (1987) ou encore par Ravid et Tolchinsky (2000, 2002) ?

Aux questions soulevées par notre problématique, nous avons les attentes suivantes. Nous supposons que les enfants, avec l'âge et l'expérience de l'école, passent du statut de locuteur natif (*native speaker*, Berman, 2004:9, 2008:2) à celui de locuteur expert (*proficient speaker*, Berman, 2004:9, 2008:2). L'individu acquiert à gérer une situation langagière complexe qui l'oblige à jongler efficacement entre ses connaissances linguistiques, cognitives, discursives et culturelles (Berman, 2004, 2008).

Nous suggérons que les enfants font des choix linguistiques de plus en plus adaptés aux contextes communicationnels ; ceci est davantage développé dans les deux parties qui suivent. La première partie présente les questions de recherche en se concentrant sur les variables indépendantes de l'étude, à savoir le NIVEAU SCOLAIRE, le TYPE DE TEXTE, la MODALITÉ et l'ORDRE DE PRODUCTION. La seconde partie aborde les questions de recherche différemment en se focalisant

sur les variables dépendantes de l'étude, à savoir : la proportion des SNL et des pronoms (Section 3.2.2.1, p. 99), le degré de lexicalité des SNL (nombre de nœud lexical) (Section 3.2.2.2, p. 100), la proportion de SN selon leur statut informationnel (Section 3.2.2.3, p. 102), l'utilisation non conventionnelle des pronoms (Section 3.2.2.4, p. 103), la forme utilisée pour maintenir une information (Section 3.2.2.5, p. 105), et la PAS (Section 3.2.2.6, p. 106).

3.2. Questions de recherche

3.2.1. Questions de recherche liées aux variables indépendantes

3.2.1.1. Le facteur niveau scolaire

Pour avoir une vision étendue de l'acquisition tardive du langage de locuteurs francophones natifs monolingues concernant un domaine sémantico-syntaxique du français, nous avons choisi d'observer trois populations différentes : des enfants de CM2 (âge moyen = 10;9), de 5^{ème} (âge moyen = 12;7) et de 3^{ème} (âge moyen = 15;2). La première VI dans ce travail est donc le NIVEAU SCOLAIRE.

Nous nous intéressons au développement tardif, à savoir à la capacité de récupérer et d'utiliser des connaissances linguistiques chez des individus de plus de cinq ans. L'adolescence apparaît comme une période intéressante pour comprendre l'acquisition tardive du langage (Berman, 1997, 2006 ; Ravid, 2005). Pendant cette période, de grands changements s'opèrent. Ceci se voit à travers (a) le lexique utilisé ; (b) la manipulation du registre, de la morphologie dite complexe, *etc.* ; (c) la syntaxe, avec une diversification et une complexification des structures syntaxiques ; (d) la gestion syntaxique à deux niveaux, à l'intérieur d'une clause et entre plusieurs clauses.⁴⁷

Ces changements sont liés aux facultés des locuteurs/scripteurs qui se développent avec le NIVEAU SCOLAIRE et avec l'âge. Berninger et Swanson (1994) suggèrent trois phases de développement de l'activité rédactionnelle des enfants : (a) une première phase se caractérisant par une importance donnée aux processus liés au code orthographique ; (b) une

⁴⁷ Cf. les travaux, entre autres, de : Aisenman et Berman (2000), Anglin (1993), Argerich et Tolchinsky (2000), Berman (1995, 1999a, 1999b, 2001, 2008), Berman et Katzenberger (2004), Berman et Slobin (1994), Berman et Verhoeven (2002), Jisa (1998, 2000a, 2000b, 2004), Johansson (1999), Gayraud (2000), Gayraud *et al.* (1999), Nippold (2002, 2004), Ravid (2000, 2002, 2005), Ravid et Shyldkrot (2005), Ravid, Shyldkrot et Berman (2005), Ravid et Tolchinsky (2002).

seconde phase durant laquelle les individus commencent à faire appel les processus de planification ; (c) une troisième phase durant laquelle les enfants mobilisent de mieux en mieux les processus de planification en révisant davantage, par exemple.

Qu'en est-il pour la population de notre étude ? Nos individus devraient avoir dépassé le premier stade de développement de la rédaction des textes étant donné que le groupe le plus jeune de notre population est constitué d'élèves scolarisés en CM2 (âge moyen = 10;9). Cette population en est alors théoriquement à la seconde phase de développement de la rédaction des textes. Il paraît probable que notre groupe de 5^{ème} (âge moyen = 12;7) et de 3^{ème} (âge moyen = 15;2) en soit à la troisième phase de développement de la rédaction de texte.

Nous voulons également voir si les enfants deviennent de plus en plus *matures linguistiquement*, ce qui présuppose la disponibilité d'un répertoire linguistique important intégrant un large éventail de registres et la capacité d'adapter les choix linguistiques aux contextes et buts communicationnels (Berman, 2006 ; Berman et Ravid, 2009 ; Ravid et Tolchinsky, 2002).

Pour cela, les individus apprendraient à mobiliser *la planification par connaissances transformées*. En effet, les plus jeunes de nos individus feraient appel à *la planification par connaissances rapportées* plutôt que *par connaissances transformées* (Bereiter et Scardamalia, 1987). Puis, avec l'expérience scolaire, les enfants développent la capacité d'opérer une *planification par connaissances transformées* (Chanquoy et Alamargot, 2002 ; Fayol, 1997 ; Olive et Piolat, 2005 ; Piolat *et al.*, 1993 ; Piolat et Roussey, 1992). Les individus passent progressivement d'une stratégie à l'autre au cours de leur développement. Les locuteurs novices ont également un éventail de formes lexicales et syntaxiques plus limité que les plus locuteurs/scripteurs plus âgés et sont également peu conscients de la diversité des formes dont ils disposent (Fayol, 1990).

Ces différences laissent supposer que les comportements linguistiques de nos sujets seront différents d'un groupe à l'autre.

3.2.1.2. Le facteur TYPE DE TEXTE

Le TYPE DE TEXTE est la seconde VI de cette étude. La pertinence d'une telle variable se trouve confirmée dans le fait que les choix linguistiques sont liés à plusieurs facteurs, dont le type de texte.⁴⁸

Nous avons choisi de faire produire à chacun des individus de notre étude deux types de textes : un texte narratif et un texte expositif. Des différences de choix linguistiques entre les productions expositives et narratives sont attendues. En effet, le contexte de production semble avoir un impact sur les choix linguistiques des locuteurs (Levelt, 1989 ; Ravid et Tolchinsky, 2000, 2002).

Le type de texte expositif fait appel à un registre académique et implique la production de formes linguistiques plus coûteuses que celles exigées par le texte narratif. Ce type textuel impose aux locuteurs/scripteurs une pression informative importante, ce qui devrait se concrétiser par une utilisation plus importante de SNL et un taux plus élevé d'informations. Ainsi, le type de texte expositif, de par son exigence cognitive, fait plus appel à *la planification par connaissances transformées* qui correspond à un effort d'adéquation entre deux espaces : celui des contenus et celui de la rhétorique.

En revanche, le texte narratif est le récit d'expériences personnelles avec des informations spécifiques que les locuteurs introduisent et dont certains sont maintenues tout le long de leur production. Ceci devrait se traduire par une utilisation plus massive des pronoms. Le type de texte narratif, contrairement au texte expositif, fait plus appel à *la planification par connaissances rapportées*, à savoir à un travail de récupération des informations, pas à pas, dans la *mémoire à long terme*. Le locuteur/scripteur connaît le début, le milieu et la fin de son histoire. Ces informations sont formulées au fur et à mesure qu'elles sont récupérées sans qu'il n'y ait nécessité d'une quelconque réorganisation de ces dernières.

3.2.1.3. Le facteur MODALITÉ

Pour cette étude, deux types spécifiques de langues sont étudiés : (a) la langue écrite de scolarisation ; et, (b) la langue orale (semi-)spontanée en situation non prototypiquement dialogique. Ces deux types de langues sont naturels dans le contexte spécifique de l'école.

⁴⁸ Divers travaux montrent que les choix linguistiques varient selon le type de texte (entre autres, Berman, 2005a, 2005b ; Berman et Nir-Sagiv, 2005, Berman et Strömqvist, 1999 ; Biber *et al.*, 1998 ; De Weck, 1991 ; De Weck et Schneuwly, 1994 ; Jisa, 1998, 2000a ; Ravid *et al.*, 2002 ; Reilly, Zamora et McGivern, 2005 ; Tolchinsky et Rosado, 2005 ; Ravid et Berman, 1999 ; Ragnarsdottir et Strömqvist, 2005 ; Ravid et Katzenberger, 1999 ; Van Hell, Verhoeven, Tak et Van Oosterhout, 2005)

L'étude des deux modalités est essentielle pour une bonne évaluation des compétences langagières de l'enfant ; la MODALITÉ DE PRODUCTION (orale/écrite) est la troisième VI manipulée dans cette étude. Le fait que l'écrit et l'oral impliquent des différences significatives dans des travaux antérieurs⁴⁹ et le fait que ces deux modalités diffèrent aussi bien par leurs caractéristiques linguistiques (la syntaxe, par exemple) que par leurs demandes cognitives (modalités de planification, par exemple),⁵⁰ justifient le choix de cette VI.

Nous postulons qu'une production écrite est plus planifiée qu'une production orale.⁵¹ De plus, la modalité écrite appelle un registre plus académique que la modalité orale. La modalité écrite se caractérise principalement par deux aspects : une contrainte temporelle assouplie et un temps de planification supérieur (Bonin, 2003 ; Fayol, 1997). Ceci facilite la production de formes syntaxiques dont le traitement s'avère plus lourd, par exemple les sujets lourds (Gayraud *et al.*, 2001 ; Jisa, 2004). En revanche, la modalité orale se caractérise par son temps de planification moindre et sa pression communicative supérieure (Blanche-Benveniste, 1990, 1995, 2000 ; Fayol, 1997 ; Gayraud, 2000). Aussi, nous attendons-nous à relever des différences quant aux choix linguistiques des participants.

Ainsi, l'accès lexical permettant la formation des SNL s'opère rapidement à l'oral du fait de son rythme de production alors qu'à l'écrit cet accès lexical peut être plus lent. En production écrite, nous devrions relever une proportion plus importante de SNL, et moins de répétitions d'un même item lexical.

3.2.1.4. Le facteur ORDRE DE PRODUCTION

Le protocole de notre étude prévoit également une autre VI liée à la MODALITÉ : l'ORDRE DE PRODUCTION. Notre population est répartie en quatre groupes de passation.⁵² Des travaux montrent que l'ORDRE DE PRODUCTION peut influencer les choix linguistiques (Fayol, Jisa et Mazur-Palandre, 2007, 2008 ; Jisa, 2004 ; Gayraud *et al.*, 1999, 2001 ; Mazur Palandre, 2008a, 2008b). L'ORDRE DE PRODUCTION a, par exemple, un effet significatif sur l'utilisation de la subordination non finie qui serait davantage utilisée dans les productions d'individus qui

⁴⁹ Nous pouvons citer, entre autres, Berman (2000), Berman, Ragnarsdottir et Strömquist (2002), Ravid, Van Hell, Rosado et Zamora (2002), Berman et Verhoeven (2002), Cahana-Amitay et Berman (2000), Gayraud *et al.*, (2000), Hurell (1956), Jisa (1998, 2000a, 2000b), Johansson (1999), Ravid et Berman (1999), Ravid et Cahana-Amitay (2005), Ravid et Katzenberger (1999), Ravid et Shlesinger (1995), Ravid, Van Hell, Rosado, et Zamora (2002), Tolchinsky et Rosado (2005), Van Hell *et al.* (2005).

⁵⁰ Nous pouvons citer, entre autres, Blanche-Benveniste (1990, 1995, 2000), Fayol (1997), Gayraud (2000), Halliday (1989), Johansson (1999).

⁵¹ Se reporter à la Section 2.2 (p. 54).

⁵² Se reporter au Tableau 3 présentant les ordres de passation.

réalisent un texte écrit puis un texte oral (Gayraud *et al.*, 1999, 2001). En effet, une fois que des connaissances (des structures syntaxiques, par exemple) sont activées et déployées à l'écrit, elles peuvent alors se répercuter dans les productions orales. Nous pouvons également penser que l'inverse soit possible.

Nous attendons que les participants fassent des choix linguistiques différents selon l'ORDRE DE PRODUCTION. Ainsi, les individus ayant élaboré une production écrite puis une production orale, usent davantage de SNL que de pronoms dans leurs textes écrits et oraux. Ces formes, activées lors de la production écrite, restent activées et sont rapidement recouvrables et sélectionnées lors de la production orale. L'inverse peut également être vrai. Des formes plus caractéristiques des textes oraux peuvent se retrouver dans les textes écrits, pour les individus ayant produit dans un premier temps à l'oral et dans un second temps à l'écrit.

3.2.2. Questions de recherche liées aux variables dépendantes

Dans les productions, nous nous intéressons principalement à la nature et à la fonction des SN tout en mettant ces informations en lien avec les contraintes discursives du discours et le statut informationnel de ces unités. Deux nouvelles questions surgissent : la position syntaxique influence-t-elle les choix des formes linguistiques ? Le statut informationnel des référents dans le discours influence-t-il la position syntaxique ? Selon divers travaux (Bock et Warren, 1985 ; Bock et Levelt, 1994 ; Keenan et Comrie, 1977 ; Leuwens, 2002 ; Levelt, 1999), l'attribution des rôles syntaxiques se ferait selon le caractère plus ou moins activé d'un élément : plus un référent serait activé, plus cet élément serait encodé sous forme pronominale et plus ce référent se verrait attribuer une fonction syntaxique dite de haut niveau (position sujet). C'est pourquoi certaines analyses de ce travail de recherche prennent en considération les positions syntaxiques.

Dans un premier temps, la lexicalité des textes est observée en analysant notamment *la proportion de SNL et de pronoms*.

Dans un second, en partant du postulat que la production de SNL est un indice de maturité linguistique, nous nous concentrons sur *le degré de lexicalité des SNL* (le nombre d'items lexicaux les constituant).

Dans un troisième temps, notre vision de la production langagière est complétée en mettant en relation les formes linguistiques avec le statut référentiel, qui influence les choix linguistiques

des individus (entre autres, Ariel, 1990, 1996 ; Du Bois, 1987, 2003 ; Khorounjaia et Tolchinsky, 2004 ; Levelt, 1989 ; Sanford et Garrod, 1981a, 1981b).

Dans un quatrième temps, une étude concernant les utilisations non conventionnelles des pronoms est effectuée ; il a été référencé les utilisations suivantes : (a) les pronoms réintroduisant un référent (*réintroductions pronominales*) ; (b) les pronoms utilisés sans SNL introduit précédemment (*référent manquant*) ; (c) les pronoms pouvant se référer à plus d'un référent préalablement intégré (*référents multiples*) ; et (d) les pronoms utilisés sans respect des accords en genre et/ou en nombre avec le référent auquel il se réfère (*problème d'accord*). Ces cas ont été analysés en pensant que ces utilisations étaient un indice (notamment pour les trois premiers cas) à la fois de non-cohérence et de respect des principes de transmission de l'information. Si les enfants ont une proportion importante de ces utilisations non conventionnelles, ceci serait dû au fait que les locuteurs/scripteurs sont submergés de tâches à réaliser. Face à cette situation, les utilisations non conventionnelles apparaissent. Ce résultat refléterait alors que la convention *SNL pour une nouvelle information ou pour désambigüiser* ne serait pas acquise. Néanmoins, si cette proportion est assez faible, nous pouvons nous dire que les individus respectent le fait d'utiliser un *SNL pour introduire une information ou pour éviter les ambigüités*.

En partant de l'idée que les productions des enfants ont une faible proportion de ces types d'utilisation, nous pensons alors que le traitement de l'ancienne information mérite réflexion, puisque le principe *SNL pour une nouvelle information ou pour désambigüiser* serait acquis. L'ancienne information peut être encodée de deux manières : par un pronom ou par un SNL. Le maintien sous forme pronominale est le choix le plus naturel et le plus économique, ce qui répond aux contraintes de la transmission de l'information à l'oral. Mais, nous pouvons nous demander si ce principe persiste avec l'avancée dans le cursus scolaire, à l'écrit et dans les différents types de textes. Ainsi, dans un cinquième temps, nous nous concentrons sur la forme des SN maintenant une information.

3.2.2.1. La lexicalité des textes : La proportion des SNL et de pronoms

Dans le Chapitre 8, la lexicalité des textes est observée. Nous nous attendons à ce que *la proportion des SNL et des pronoms* varie selon les facteurs NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE.

Nous nous attendons à un effet du NIVEAU SCOLAIRE. Les jeunes scripteurs allouent leurs ressources cognitives majoritairement aux processus de traduction et de transcription au détriment des tâches de haut niveau telles que l'accès lexical (Bereiter et Scardamalia, 1987 ; Berninger et Swanson, 1994). L'accès lexical des plus jeunes peut en être entravé, ce qui se traduirait par une proportion moins importante de SNL dans leur production.

Nous prédisons également une influence du contexte de production (TYPE DE TEXTE et MODALITÉ) sur l'emploi des SNL et des pronoms. Le type de texte expositif exige un registre assez académique et impose une pression informative que le type de texte narratif n'impose pas. Ainsi, le texte expositif se caractériserait par une proportion plus importante de SNL. En revanche, dans le texte narratif, les locuteurs/scripteurs se concentrent sur un protagoniste principal, qui, une fois introduit, est souvent maintenu. Le maintien se faisant préférentiellement sous forme pronominale,⁵³ ce type de texte se caractérise sans doute par une proportion plus importante de pronoms.⁵⁴

Concernant les deux modalités de production, de par leurs caractéristiques,⁵⁵ nous supposons que *la proportion des SNL* est plus importante dans les textes écrits que dans les oraux, en revanche, *la proportion des pronoms* est sans doute plus élevée dans les textes oraux que dans les textes écrits. En effet, la modalité écrite se caractérise, entre autres, par un rythme de production lent qui engendre un temps de planification plus modulable et plus long que la modalité orale. Les enfants peuvent éventuellement se servir de ce temps supplémentaire pour mobiliser et produire des SNL.

Nous supposons également un effet de l'ORDRE. Ainsi, les individus ayant produit un texte écrit puis un texte oral réaliseraient sans doute davantage de SNL (puisqu'ils sont déjà activés et déployés à l'écrit) que de pronoms.

Enfin, *les proportions de SNL et de pronoms* selon les positions syntaxiques devraient également varier selon les facteurs de cette étude. En effet, les SNL se trouveraient davantage en position non sujet qu'en position sujet (Blanche-Benveniste, 1995, 1990 ; François, 1974 ; Jeanjean, 1980 ; Lambrecht, 1987). Nous attendons des résultats différents selon la position des SN.

⁵³ En effet, nous avons vu que un des principes de la transmission de l'information, et notamment de l'oral, est l'encodage d'une ancienne information par le biais d'un pronom.

⁵⁴ Les spécificités de chacun des types de textes sont développées à la Section 2.4 (p. 70).

⁵⁵ Ces caractéristiques sont présentées à la Section 2.2 (p. 54).

Nos prochaines prédictions sont faites dans l'optique que les hypothèses, émises précédemment, se confirment (principalement l'hypothèse selon laquelle la proportion des SNL augmente avec l'âge). Le SNL apparaît comme un indice de maturité linguistique, d'autant plus que cette unité, rappelons-le, semble jouer un rôle essentiel dans l'acquisition du langage et la construction de la structure du discours (Dockrell et Messer, 2004 ; Francis, 1986 ; Ravid 2005, 2006 ; Ravid et Berman, sous presse). De plus, la manipulation de cette unité est un des aspects reflétant le passage de l'individu de *native speaker* à *proficient speaker* (Berman, 2008 ; Ravid et Tolchinsky, 2000, 2002). C'est en partant de ce postulat que nous nous concentrons sur le degré de lexicalité des SNL.

3.2.2.2. Le degré de lexicalité des SNL

Dans le Chapitre 9, *le degré de lexicalité des SNL* est observé. Par degré de lexicalité, nous entendons le nombre d'items lexicaux constituant un SNL, indice de complexité des SNL.⁵⁶ Nous avons distingué *les SNL constitués d'un nœud lexical* et ceux *de plus d'un nœud*. Nous nous attendons à ce que *la proportion des SNL constitués d'un nœud lexical* et celle *des SNL constitués de plus d'un nœud* varient selon les facteurs NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE.

Nous pouvons imaginer que les enfants de notre étude produisent de plus en plus de SNL complexes. Ils développent des capacités leur permettant de produire des formes cognitivement plus coûteuses et ce grâce à leur expérience avec les différents contextes de production (Ravid, 2000).

Nous pensons également que les individus ont tout de même des préférences pour un type de SNL particulier selon le contexte de production (TYPE DE TEXTE et MODALITÉ). Le texte expositif exigeant un registre plus académique que le texte narratif pourrait appeler la production de SNL plus complexes que le type narratif : les enfants apprendraient à s'adapter au contexte de production, tout comme les adultes (Levelt, 1989).

Les participants de notre étude pourraient produire des SNL composés de plus d'un nœud lexical davantage en production écrite qu'en production orale. En effet, la production écrite, de par son rythme de production lent, permet un temps de planification plus important, ce qui faciliterait l'accès lexical, et donc également la production de plus d'un nœud lexical à l'intérieur d'un même SNL.

⁵⁶ Le nombre de nœud comme indice de complexité du SNL est argumenté à la Section 1.3.4 (p. 46/47).

L'ORDRE ferait également varier *la proportion des SNL constitués d'un nœud lexical, celle des SNL constitués de deux et celle des SNL constitués de plus de deux nœuds*. L'ordre écrit/oral faciliterait davantage la production de SNL composés de plus d'un nœud.

Enfin, nous nous attendons à ce que *les proportions des SNL composés d'un nœud et de plus d'un nœud*, varient également selon leurs positions syntaxiques. En effet, les locuteurs/scripteurs combleraient la position syntaxique sujet, le plus souvent, en ayant recours à un sujet léger (Blanche-Benveniste, 1995, 1990 ; Chafe, 1994 ; Du Bois, 1987, 2003 ; François, 1974 ; Jeanjean, 1980 ; Lambrecht, 1987). Ainsi, les individus de notre étude suivent sans doute cette tendance et s'ils placent un SNL en position sujet, nous supposons qu'il sera généralement le moins lourd possible.

Partons de l'idée que les résultats de la lexicalité des textes et de la constitution des textes fassent ressortir un lien fort entre formes linguistiques, positions syntaxiques, contexte de production et âge. Nous pouvons alors compléter notre connaissance de la production langagière en mettant en relation tous ces aspects avec le statut référentiel, qui semble lui-aussi avoir un impact sur les choix linguistiques des individus (entre autres, Ariel, 1990, 1996 ; Du Bois, 1987, 2003 ; Khorounjaia et Tolchinsky, 2004 ; Levelt, 1989 ; Sanford et Garrod, 1981a, 1981b).

3.2.2.3. Le flux de l'information : statut informationnel des syntagmes nominaux

C'est ainsi que le Chapitre 10 se focalise sur le statut informationnel des SN. Nous nous attendons à ce que *les proportions des SN introduisant une nouvelle information et maintenant une information* varient selon les variables : NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE.

Le NIVEAU SCOLAIRE aurait un effet significatif. L'introduction de nouveaux référents est considérée comme exigeant plus de ressources cognitives que l'ancienne information (Chafe, 1994 ; Khorounjaia *et al.*, 2004), d'autant plus qu'introduire une nouvelle information implique l'utilisation de lexique. C'est dans cette mesure que la *proportion des SN introduisant une nouvelle information* devrait augmenter avec l'avancée dans le cursus scolaire.

Néanmoins, si l'introduction d'informations dans un texte constitue une des règles assurant la cohésion et la cohérence des textes (Charolles, 1978 ; Reinhart, 1980 ; Schneuwly, 1988), une

autre règle est le maintien, dans le texte, des référents déjà introduits (Charolles, 1978 ; Reinhart, 1980 ; Schneuwly, 1988). En effet, un texte cohérent est le passage d'une nouvelle information à une autre, chacune d'entre elles est, de plus, développée. Ainsi, nous pouvons nous demander si le maintien référentiel, qui semble être une acquisition tardive (Hickmann, 2000, 2003 ; Hickmann et Hendricks, 1999 ; Kail et Hickmann, 1992), n'augmenterait pas avec le NIVEAU SCOLAIRE. Nous pouvons alors également supposer que *la proportion des SN maintenant une ancienne information* est plus importante chez les locuteurs plus experts. Nous attendons les résultats pour statuer sur cette question.

Nous prévoyons également que *la proportion des SN introduisant et maintenant une information* varie selon le contexte de production (TYPE DE TEXTE et MODALITÉ). Nous supposons que *la proportion des SN introduisant une nouvelle information* a une proportion plus importante dans les textes expositifs que dans les textes narratifs. En revanche, *la proportion des SN maintenant une information* est sans doute plus élevée dans les textes narratifs. En effet, rappelons que le type de texte expositif impose une pression informative plus importante que le texte narratif dont une des caractéristiques est le maintien de référents une fois qu'ils sont introduits.

De plus, la variable MODALITÉ ferait également varier la proportion des SN introduisant une nouvelle information et celle des SN maintenant une information. La proportion des SN introduisant une nouvelle information est sans doute plus importante dans les textes écrits que dans les textes oraux. En revanche, la proportion des SN maintenant une information est sans doute plus élevée dans les productions orales que dans les productions écrites. En effet, l'introduction de nouveaux éléments implique la mobilisation de formes lexicales, cognitivement plus lourdes que les formes pronominales. De plus, introduire un nouveau référent est bien plus coûteux que d'en maintenir un déjà connu (Chafe, 1994). Les individus doivent activer le nouveau référent et l'introduire sous forme lexicale pour assurer la bonne compréhension du destinataire. Nous avons déjà vu que la modalité écrite facilite sans doute, de par son rythme de production lent, la production de telles formes.

Dans le même esprit, nous supposons également que les individus ayant produit dans l'ordre de passation écrit/oral introduisent davantage de nouvelles informations que les individus ayant produit dans l'ordre inverse. Des formes mobilisées à l'écrit restent activées lors de leur production orale et les locuteurs les utilisent de nouveau.

Enfin, nous faisons l'hypothèse que l'observation de *la proportion des SN introduisant et maintenant une information*, selon les positions syntaxiques, peut compléter les résultats que nous trouvons lors des analyses ne prenant pas en considération la position syntaxique. En effet, nous avons déjà mentionné le fait que l'attribution des rôles syntaxiques dépendrait de l'accessibilité d'un référent (Bock et Warren, 1985 ; Bock et Levelt, 1994 ; Keenan et Comrie, 1977 ; Leuwers, 2002 ; Levelt, 1999). Nous pouvons nous attendre à une distribution particulière de l'ancienne et de la nouvelle information : l'ancienne information se retrouverait davantage en position sujet alors que la nouvelle information en position non sujet.

Une étude mettant en relation statut informationnel et formes linguistiques pourrait compléter ces résultats. Nous savons que la nouvelle information est forcément lexicale. Néanmoins, nous pourrions trouver dans les productions des enfants des réintroductions sous forme pronominale, des introductions de nouveaux référents sous formes pronominales sans référence à un SNL préalablement introduit ou encore des ambiguïtés référentielles. C'est pourquoi nous analysons les utilisations non conventionnelles de pronoms par les enfants.

3.2.2.4. L'utilisation non conventionnelle des pronoms

Le Chapitre 11 présente une analyse de l'utilisation non conventionnelle des pronoms. En effet, lors du codage des données, nous avons été confrontée à quatre catégories d'utilisation non conventionnelle des pronoms. La première catégorie, la *réintroduction pronominale* est l'utilisation d'un pronom pour réintroduire un référent déjà présent dans le texte mais trop loin pour que l'utilisation d'un pronom n'entrave pas la compréhension. La réintroduction d'un tel référent se réalise sous forme lexicale. La deuxième catégorie, *référent manquant*, concerne l'utilisation de pronoms anaphoriques qui apparaissent sans mention d'un référent au préalable. La troisième catégorie, *référents multiples*, correspond à l'utilisation d'un pronom pouvant se rapporter à plus d'un référent mentionné au préalable. Enfin, la quatrième catégorie, *problème d'accord*, est celle des pronoms anaphoriques qui sont accordés de manière non conventionnelle ; l'accord en genre et en nombre du pronom avec son référent s'écarte de la norme de la *langue de scolarisation*.

Ces types d'utilisation devraient diminuer avec le NIVEAU SCOLAIRE et être plus fréquents dans les textes expositifs que dans les textes narratifs. Nous ne faisons pas d'hypothèses concernant

la MODALITÉ et l'ORDRE étant donné certaines limites de codage. En effet, en production orale, nous ne pouvons pas détecter certaines erreurs notamment celles de nombre. De plus, l'oral n'engendre pas forcément un respect de l'accord en genre et en nombre du pronom avec le SNL avec lequel il entretient une relation de coréférence (Berrendonner, 1994 ; Reichler-Béguelin, 1989, 1993).

L'utilisation non conventionnelle des pronoms (et notamment, *réintroduction pronominale*, *réfèrent manquant* et *référénts multiples*) peut montrer deux choses. Premièrement, si les résultats de nos analyses montrent une proportion importante de ces utilisations, cela peut être l'indice que les enfants, submergés de tâches cognitives, ne prennent pas encore tout à fait en compte les connaissances de leur destinataire. En effet, la communication implique un effort coopératif constant entre les participants d'un discours ; il doit y avoir un partage des connaissances (entre autres, Levelt, 1989 ; Haviland et Clark, 1974). Les enfants ont alors trop de choses à faire en même temps, ce qui se traduit par l'utilisation non conventionnelle de pronoms. Deuxièmement, si les participants de notre étude ont une faible proportion d'utilisation non conventionnelle des pronoms, cela peut laisser entendre qu'ils respectent les principes du traitement de l'information, c'est-à-dire que s'il y a ambiguïté, ou encore introduction d'une nouvelle information, ils utilisent le lexique plutôt que les pronoms.

Partons de l'idée que les trois populations de notre étude obéissent aux principes de transmission de l'information, et notamment à celui d'introduire une nouvelle information sous forme lexicale. Si ce cas s'avère confirmé par nos résultats, cela reflète qu'introduire une nouvelle information sous forme lexicale est acquis relativement tôt. Il est alors sans doute plus intéressant de se concentrer sur les formes linguistiques relatives aux anciens référents pour lesquelles la langue offre deux possibilités : les pronoms et les SNL. En effet, même si le maintien de l'information semble réalisé préférentiellement sous forme pronominale, ce qui est un des principes de transmission de l'information à l'oral, la langue permet également le maintien sous forme lexicale. Si l'introduction d'une nouvelle information sous forme pronominale est considérée comme une entrave à la compréhension, le maintien sous forme lexicale est tout autrement envisagé. En effet, maintenir un référent sous forme lexicale, en faisant preuve de variabilité quant aux lexèmes choisis, est une des caractéristiques des productions écrites d'experts.

3.2.2.5. La forme utilisée pour maintenir l'ancienne information

Le Chapitre 12 se concentre sur *la proportion des pronoms et des SNL maintenant une ancienne information*. La forme linguistique choisie pour encoder un référent dépend de son statut informationnel dans le discours (Ariel, 1990, 1996 ; Khorounjaia et Tolchinsky, 2004 ; Levelt, 1989 ; Sanford et Garrod, 1981a, 1981b). Un nouveau référent, qui est de faible accessibilité, serait encodé sous la forme de SNL, alors qu'un élément actif, donc ancien, serait encodé sous forme pronominale. Les participants de notre étude devraient suivre cette tendance. *La proportion de l'ancienne information sous forme pronominale serait donc plus importante que celle de l'ancienne information sous forme lexicale.*

Nous supposons un effet du NIVEAU SCOLAIRE sur *la proportion des SN maintenant une information sous forme pronominale et sous forme lexicale*. Avec l'expérience, les individus de notre étude maintiennent certainement de plus en plus une information sous forme lexicale, ce qui va à l'encontre des principes de la transmission de l'information de la langue orale. Les locuteurs/scripteurs apprennent à utiliser le SNL, comme outil rhétorique, pour maintenir une information.

Nous attendons également un effet du TYPE DE TEXTE. Les participants de notre étude privilégieraient le maintien sous forme lexicale pour les textes expositifs et écrits. En effet, nous pouvons penser que la substitution lexicale se réalise davantage en production expositive qu'en production narrative, ce qui serait dû au registre académique de ce texte ainsi qu'à sa pression informationnelle.

Concernant la MODALITÉ DE PRODUCTION, l'écrit, par son rythme de production lent, facilite l'accès lexical et donc la mobilisation et la production de SNL. De plus, cette modalité de production est sans doute celle permettant le plus aux individus d'aller à l'encontre des principes naturels de la transmission de l'information, à savoir maintenir une forme de préférence par un pronom.

Nous suggérons que les individus ayant produit dans l'ordre écrit/oral maintiennent plus l'information sous forme lexicale que les individus ayant produit dans l'ordre inverse ; et ce pour les mêmes raisons évoquées précédemment : les formes lexicales mobilisées à l'écrit restent activées lors de la production orale et les locuteurs les utilisent de nouveau.

Si notre hypothèse – selon laquelle *la proportion du maintien sous forme lexicale* augmente avec les années de scolarisation – se confirme, alors le maintien sous forme lexicale apparaît comme un signe de maturité. Les enfants de notre étude doivent suivre une trajectoire, qui aboutira au comportement linguistique caractéristique de celui des scripteurs experts (De Weck, 1991 selon Moirand, 1975), à savoir maintenir une information sous une forme lexicale différente de la première apparition. De ce fait, il peut être intéressant de voir le degré de lexicalité des SNL maintenant une information. Nous observons alors : (a) *la proportion des SNL maintenant une information en n'impliquant pas de changement linguistique par rapport à la mention précédente* (**Mon chat s'est enfui, mon chat me manque**) ; et (b) *la proportion des SNL maintenant une information impliquant des changements linguistiques par rapport à la mention précédente* (**Mon chat s'est enfui, ce petit félin me manque**).

Nous supposons un effet du NIVEAU SCOLAIRE. Les enfants avec l'expérience vont sans doute de plus en plus modifier la forme des SNL qu'ils maintiennent.

De plus, nous nous attendons à ce que les individus maintiennent plus fréquemment une information en variant la forme linguistique à l'écrit et en texte expositif, qu'à l'oral et qu'en texte narratif. Nous pouvons également penser que les individus appartenant à l'ordre de passation écrit/oral maintiennent plus une information sous forme lexicale en variant la forme.

3.2.2.6. Le concept de la PAS de Du Bois (1987, 2003)

Nous souhaitons examiner si les choix linguistiques des enfants de notre étude obéissent aux mêmes contraintes que celles gouvernant les choix des adultes dans les différentes langues décrites par Du Bois (1987, 2003) et ses collègues. Les textes des locuteurs/scripteurs sont analysés dans le but de déterminer dans quelle mesure le NIVEAU SCOLAIRE, le TYPE DE TEXTE et la MODALITÉ peuvent faire varier la PAS. Nous voulons vérifier si la position syntaxique de la nouvelle information est gouvernée par des contraintes discursives (Du Bois, 1987 ; Khorounjaia et Tolchinsky, 2004).

Rappelons que la PAS sous-entend quatre contraintes : (A) une clause est constituée au maximum d'un élément lexical ; (B) cet élément lexical n'apparaît pas en position sujet d'un verbe transitif ; (C) une clause est constituée d'au maximum un seul nouvel élément ; et (D) l'ancienne information se trouve en position sujet de verbe transitif.

La transmission de l'information se fait différemment selon la modalité de production. La production orale soumet les individus à de telles contraintes (principalement, rythme de

production rapide et contraintes conversationnelles), que leurs choix linguistiques en sont influencés. Ces choix reflètent les contraintes de la PAS. Les travaux en acquisition prenant comme objet d'étude la PAS révèlent que les individus observés (des jeunes enfants aux adultes) ont comme pattern préférentiel la PAS. Elle est *naturellement* utilisée à l'oral. Nos analyses apportent sans doute des informations complémentaires sur la PAS, dans la mesure où elles prennent en considération la deuxième modalité du langage, la modalité écrite ainsi que deux types de discours (narratif/expositif). Les études présentées, excepté celle de Khorounjaia et Tolchinsky (2004), ne traitent pas d'une telle situation.

Les choix linguistiques découlent de la transmission de l'information à l'oral, qui diffère selon les modalités du langage. Les modalités écrite et orale ont des caractéristiques spécifiques et donc la transmission de l'information ne se réalise pas dans les mêmes circonstances. Les locuteurs/scripteurs vont-ils avoir recours aussi naturellement à la PAS à l'écrit ? En partant du postulat que la syntaxe émerge du discours, la modalité orale pourrait être le contexte qui laisse à la PAS le plus de liberté pour se réaliser. L'écrit forcerait l'enfant francophone à aller à l'encontre de la transmission naturelle de l'information.

D'une façon générale, avec l'expérience, les enfants apprendraient à aller à l'encontre des contraintes de la PAS, caractéristique de la modalité orale. Le fait de ne plus se soumettre aux contraintes de la PAS serait en lien direct avec les exigences de l'écrit scolaire.

Selon le type de texte, les choix de l'enfant reflèteraient davantage le pattern de la PAS. En texte expositif, les locuteurs/scripteurs vont certainement à l'encontre de ces contraintes discursives naturelles de la langue parlée. En revanche, en texte narratif, ce pattern pourrait davantage ressortir.

Ainsi, nous pouvons nous attendre à ce que :

- le nombre moyen d'éléments lexicaux par clause (contrainte A) augmente avec le niveau scolaire, et il serait plus important dans les textes expositifs et écrits que dans les textes narratifs et oraux, et dans les textes des individus ayant produit dans l'ordre écrit/oral.
- la proportion des sujets lexicaux (contrainte B) augmente avec le niveau scolaire, et elle serait plus importante dans les textes expositifs et écrits que dans les textes narratifs et oraux, et dans les textes des individus ayant produit dans l'ordre écrit/oral.

- le nombre moyen de nouvelle information par clause (contrainte C) augmente avec le niveau scolaire, et il serait plus important dans les textes expositifs et écrits que dans les textes narratifs et oraux, et dans les textes des individus ayant produit dans l'ordre écrit/oral.
- la proportion de l'ancienne information en position sujet (contrainte D) diminue avec le niveau scolaire, et elle serait plus importante dans les textes narratifs et oraux que dans les textes expositifs et écrits, et dans les textes des individus ayant produit dans l'ordre oral/écrit.

3.3. Récapitulation des questions de recherche

Le Tableau 4 résume l'ensemble des attentes en fonction des variables indépendantes de cette étude, à savoir le NIVEAU SCOLAIRE, le TYPE DE TEXTE, la MODALITÉ et l'ORDRE DE PRODUCTION.

[Tableau 4 : Récapitulation des attentes]

Tableau 4 : Récapitulation des attentes⁵⁷

			NIVEAU SCOLAIRE	TYPE DE TEXTE	MODALITÉ	ORDRE ⁵⁸
SNL (Chapitre 8)			Augmente	Plus en expositif	Plus à l'écrit	Plus écrit/oral
Pronoms (Chapitre 8)			Diminue	Plus en narratif	Plus à l'oral	Plus oral/écrit
SNL (Chapitre 9)	1	nœud	Augmente	Plus en narratif	Plus à l'oral	Plus oral/écrit
SNL (Chapitre 9)	+1	nœud	Augmente	Plus en expositif	Plus à l'écrit	Plus écrit/oral
Ancienne (Chapitre 10)	information	Attentes non posées		Plus en narratif	Plus à l'oral	Plus oral/écrit
Nouvelle (Chapitre 10)	information	Attentes non posées		Plus en expositif	Plus à l'écrit	Plus écrit/oral

⁵⁷ Nous parlons de proportions.

⁵⁸ Par *écrit/oral*, nous entendons les individus ayant produit dans un premier temps un texte écrit puis dans un second temps un texte oral. Par *oral/écrit*, nous entendons les individus ayant produit dans un premier temps un texte oral puis dans un second temps un texte écrit. Ainsi, par *plus écrit/oral*, nous entendons que les individus appartenant à cet ordre de production ont une proportion plus importante de SNL, par exemple.

Pronoms non conventionnels (Chapitre 11)	Diminue	Plus en expositif	/	/
Ancienne information sous forme lexicale (Chapitre 12)	Augmente	Plus en expositif	Plus à l'écrit	Plus écrit/oral
Ancienne information sous forme pronominale (Chapitre 12)	Diminue	Plus en narratif	Plus à l'oral	Plus oral/écrit
Substitution lexicale par le biais de la répétition (Chapitre 12)	Diminue	Plus en narratif	Plus à l'oral	Plus oral/écrit
Substitution lexicale : changements linguistiques (Chapitre 12)	Augmente	Plus en expositif	Plus à l'écrit	Plus écrit/oral

Le Tableau 5 résume les attentes de cette étude relatives aux contraintes de la PAS par rapport aux variables indépendantes.

[Tableau 5 : Attentes relatives à la Preferred Argument Structure]

Tableau 5 : Attentes relatives à la Preferred Argument Structure

	Niveau scolaire	Type de texte	Modalité	Ordre
Respect de la Preferred Argument Structure	Diminue	Plus en narratif	Plus à l'oral	Plus oral/écrit
Contraintes				
A. Un seul élément lexical par clause (Chapitre 8)	Augmente	Plus en expositif	Plus à l'écrit	Plus écrit/oral
B. Position sujet non lexicale (Chapitre 8)	Augmente	Plus en expositif	Plus à l'écrit	Plus écrit/oral
C. Une seule nouvelle information par clause (Chapitre 10)	Augmente	Plus en expositif	Plus à l'écrit	Plus écrit/oral
D. Ancienne information en position sujet	Diminue	Plus en narratif	Plus à l'oral	Plus oral/écrit

(Chapitre 10)

Tous les résultats concernant la nature, la fonction, le statut informationnel des formes pronominales et lexicales ainsi que ceux concernant la PAS, fourniront un état de la gestion de la cohérence et de la cohésion pour ces groupes d'individus. Ceci se verra notamment par le biais de la gestion des statuts informationnels et des marques de cohésion à adapter selon le degré d'activation d'un référent. Durant l'adolescence de grands changements linguistiques s'opèrent et ceci dans le but de devenir un *proficient speaker* (Berman, 2004, 2008). Nous voulons, entre autres, voir comment les enfants réagissent face aux diverses contraintes de la langue et si leurs choix linguistiques sont adaptés en fonction des contextes et des buts communicationnels.

Partie II

Méthodologie

Chapitre 4

Collecte des données et Protocole

Notre travail de recherche est dans la continuité d'une étude psycholinguistique internationale⁵⁹ (le projet Spencer) qui avait pour but l'analyse du développement de l'expression orale et écrite de quatre populations socio-économiquement aisées : des enfants de 9/10 ans, des pré-adolescents de 12/13 ans, des adolescents de 15/16 ans et des adultes BAC+5. L'équipe expérimentale du projet, auquel appartient ce travail de thèse, s'est inspirée du protocole expérimental de ce projet Spencer.

Après avoir mis au point le protocole, le travail de terrain, qui s'est étalé sur un an, a débuté. Durant cette première année du projet, nous avons sélectionné les lieux dans lesquels nous avons par la suite constitué notre population (trois groupes de niveau scolaire : CM2, 5^{ème} et 3^{ème}). Nous nous sommes ensuite concentrées sur la collecte des données consistant à obtenir quatre textes par individu : un narratif oral, un narratif écrit, un expositif oral et un expositif écrit.

Toutes les décisions relatives à la collecte des données, au protocole ou encore aux choix des logiciels ont été prises par l'équipe ReFlex, dont je fais partie. Cette équipe a été différente selon les besoins et l'avancée des travaux. Ainsi, l'équipe permanente se constitue du porteur (Harriet Jisa), du coordinateur (Florence Chenu), des collaborateurs principaux (Michel Fayol, Bernard Lété et Séverine Maggio) et de moi-même.⁶⁰ Lors de la collecte des données, nous avons demandé de l'aide à des expérimentateurs. Puis des étudiants ont effectué divers

⁵⁹ Rappelons que ce projet regroupait sept pays, l'Espagne, la France, l'Islande, l'Israël, les Pays-Bas, la Suède et les États-Unis d'Amérique. Cette étude était financée par la fondation Spencer (USA) et intitulée *Developing literacy in different contexts and different languages*.

⁶⁰ Harriet Jisa et Florence Chenu : Laboratoire DDL, UMR 5596, Université de Lyon ; Michel Fayol et Séverine Maggio : Laboratoire LAPSCO, UMR 6024, Université Blaise Pascal ; Bernard Lété : EMC, EA3082, Université de Lyon et INRP.

travaux tels que Maggio (2007) sur l'orthographe, Cautain (2007) sur les évaluations ou encore Begnet (2007) sur les pauses.

Ce chapitre se concentre, dans un premier temps, sur le déroulement de la collecte en elle-même avec la présentation du protocole et les consignes données aux expérimentateurs. Dans un second temps, nous présentons les critères de sélection des lieux et des individus. Puis, le matériel utilisé pour effectuer la collecte des données est présenté. Enfin, les restrictions méthodologiques de ce travail closent cette partie *Méthodologie*.

4.1. Déroulement de la collecte des données

La sélection des individus et la collecte des données se sont étalées sur dix mois de septembre 2005 à juin 2006. Tous les élèves ont participé sur la base du volontariat avec une autorisation parentale (Annexe 2).

La collecte des données, au collège, s'est déroulée sur six semaines, en mars et avril 2006. Notre équipe était présente dans l'établissement quatre demi-journées par semaine. Durant ces demi-journées, nous organisons des séances de deux heures pendant lesquelles huit individus produisaient deux textes chacun. Les collégiens ont participé à l'étude en dehors des cours. Dans les écoles primaires, la collecte des données s'est déployée sur deux semaines, en mai et juin 2006. Notre équipe était présente une demi-journée par semaine dans chaque école. Deux semaines de collecte ont été ajoutées, à raison de deux demi-journées par semaine, pour récupérer les productions des élèves qui avaient été absents. Durant ces demi-journées, nous organisons des séances d'environ une heure et quart pendant lesquelles huit individus produisaient deux textes chacun. Les écoliers, contrairement aux collégiens, ont participé à l'étude durant les horaires de d'école ; organisation proposée par les institutrices et instituteurs.

4.1.1. Protocole expérimental

Notre projet consiste à susciter pour chaque individu participant à notre étude la production de quatre textes : un texte narratif oral, un texte narratif écrit, un texte expositif oral, et un texte expositif écrit, sachant que les productions écrites ont une version brouillon et une propre.⁶¹ Tous ces textes avaient une thématique commune : les problèmes entre les gens. Diverses

⁶¹ Pour notre travail de thèse, nous avons utilisé les versions orales ainsi que les versions propres des écrits.

études ont prouvé la pertinence d'une telle thématique ; les individus se sentent tous concernés par ce sujet et produisent davantage que pour un autre thème (Gayraud, 2000 ; Labov, 1978/1993 ; Viguié-Simon, 2001). Les locuteurs ont tous visionné une vidéo muette, produite spécifiquement pour l'étude antérieure dont nous nous sommes inspirée, le projet Spencer. Ce stimulus a été créé spécialement pour le projet Spencer par une équipe journalistique de la Télévision Éducative d'Israël, qui a eu le souci de produire une vidéo « *moralement inoffensive et visible* » par les différents groupes de participants (Viguié-Simon, 2001:46). Cette vidéo, qui présente une série de clips portant sur des problèmes relatifs au milieu scolaire tels que la tricherie, les disputes, l'exclusion, *etc.*, a été produite sans parole afin d'assurer son adaptabilité internationale mais également pour ne pas influencer les choix linguistiques des individus.

Nous avons également essayé de motiver au maximum nos individus: (a) nous leur avons clairement expliqué ce que nous attendions d'eux et le but de leur participation ; (b) nous avons mis en avant les enfants choisis en les valorisant et en les définissant comme des participants à une étude scientifique ; (c) nous leur avons signalé que sans eux, nous ne pourrions mener cette étude. Il en est ressorti que tous les sujets sélectionnés, même ceux présentant un fort degré d'absentéisme dans le système scolaire, sont venus à leur rendez-vous, même le mercredi après-midi (demi-journée durant laquelle les collégiens de notre population n'avaient jamais cours). Tous nos individus ont produit et ont pris la tâche au sérieux.

La collecte, pour les sujets, s'est organisée sur deux sessions. Une semaine séparait les deux sessions et, entre les deux productions d'une même session, nous avons organisé des tâches de distraction afin d'éviter le plus possible les transferts (mot à mot) d'une modalité à l'autre. Chaque sujet a donc eu deux rendez-vous. Lors du premier rendez-vous (Session 1), nous avons fait une présentation rapide de notre projet puis nous avons projeté la vidéo aux individus qui ont poursuivi en produisant deux textes. Entre les deux textes qu'ils avaient à produire, nous leur avons demandé de répondre à un questionnaire (Questionnaire 2, Annexe 3) rempli par le soin de nos investigateurs afin d'éviter une surcharge de l'écrit pour nos sujets (surtout pour les plus jeunes). Lors du second rendez-vous (Session 2), les individus sont passés directement à la production textuelle sans regarder une nouvelle fois la vidéo. Nous leur avons demandé de produire à nouveau deux textes. Comme lors de la première

session, ils ont répondu à un questionnaire (Questionnaire 3, Annexe 4), toujours avec l'aide de nos investigateurs.

Trois groupes de participants constituent notre population : un groupe de CM2, un de 5^{ème} et un de 3^{ème}. Les individus dans un même groupe de niveau scolaire sont répartis de manière équilibrée en quatre ordres de passage qui sont illustrés dans le Tableau 6. L'ordre (oral-écrit, écrit-oral) a été contrôlé dans la mesure où des études pilotes ont révélé un effet significatif de l'ordre de production pour certains critères d'analyse (Fayol, Jisa et Mazur-Palandre, 2007, 2008 ; Jisa, 2004 ; Jisa et Mazur, 2006a, 2006b ; Gayraud *et al.*, 1999, 2001 ; Mazur-Palandre, 2007a, 2007b, 2008a, 2008b).

[Tableau 6 : Ordre de passage]

Tableau 1 : Ordre de passage

	Semaine I	Semaine II
A	Narratif oral – Questionnaire – Narratif écrit	Expositif oral – Questionnaire – Expositif écrit
B	Narratif écrit – Questionnaire – Narratif oral	Expositif écrit – Questionnaire – Expositif oral
C	Expositif oral – Questionnaire – Expositif écrit	Narratif oral – Questionnaire – Narratif écrit
D	Expositif écrit – Questionnaire – Expositif oral	Narratif écrit – Questionnaire – Narratif oral

Ainsi, si un individu appartient à l'ordre de passation A, lors de la première semaine, il regarde la vidéo puis produit un texte narratif oral. Ensuite, l'expérimentateur lui fait remplir un questionnaire et le sujet produit un texte narratif écrit. Une semaine après, il revient et produit un texte expositif oral, remplit un autre questionnaire et fait un texte expositif écrit.

Ces ordres permettent de voir s'il existe un éventuel impact de l'écrit sur l'oral, de l'oral sur l'écrit, de la narration sur l'exposition et enfin de l'exposition sur la narration.

4.1.2. L'équipe expérimentale

4.1.2.1. Consignes communes

Des consignes spécifiques et adaptées aux contextes de production imposés (narratif oral, narratif écrit, expositif oral, expositif écrit) ont été données. Il a été demandé aux

expérimentateurs de ne jamais oublier qu'ils étaient face à des enfants et adolescents. Les consignes étaient à dire lentement, en observant les individus afin de vérifier s'ils saisissaient la consigne. Les expérimentateurs devaient alors rassurer les sujets et nous leur avons donné une liste d'expressions qu'ils pouvaient verbaliser telle que *vas-y, continue, c'est bien, etc.* Lors de la formation des investigateurs, l'équipe a insisté aussi sur le fait que nous recherchions à obtenir des monologues. S'ils étaient dans la position inconfortable de faire face à une longue pause, ils n'avaient pas à la combler. Une discussion a été engagée pour leur montrer la nécessité des pauses lors l'encodage du message.

Chaque expérimentateur, selon le type de texte auquel il était affecté, avait en sa possession un document comportant les questions que pouvaient éventuellement poser les individus aux investigateurs et les réponses que ces derniers devaient formuler (Annexe 5). De plus, les expérimentateurs responsables des textes expositifs, étaient munis d'un document présentant une relance. Cette relance était destinée à débloquent une situation de silence total due par exemple à un problème de compréhension de la consigne malgré sa répétition (Annexe 6). Enfin, les expérimentateurs étaient aussi pourvus d'une fiche *observation de l'individu* (Annexe 7) sur laquelle il leur était demandé de noter des observations sur le déroulement de la séance telles que *panne de courant, sonnerie du collège, etc.*, ainsi que des observations éventuelles sur l'individu telles que *enrhumé, timide, etc.* Il a été demandé aux expérimentateurs de remplir cette fiche une fois que l'élève avait quitté la salle.

4.1.2.2. Expérimentateurs et scénarios

Pour mener à bien cette étude et afin de ne pas trop indisposer les écoles primaires et le collège, il a été décidé de collecter les données en équipe. Plus nous étions nombreux pour assurer ce travail de terrain, moins nous passions de temps dans les établissements scolaires à les importuner. L'équipe responsable de la collecte comportait sept personnes, chacune ayant un rôle bien défini et des consignes spécifiques. Chacune des personnes devait respecter le protocole en prenant le moins de liberté possible.

Une première personne (A) s'occupait de l'organisation générale de la collecte. Elle gérait les groupes d'enfants, les amenait dans les bonnes salles au bon moment et réglait les problèmes éventuels tels que l'absentéisme des individus. Elle était également chargée de la présentation générale du projet aux individus avant la projection de la vidéo (Annexe 8).⁶²

⁶² Il s'agissait de l'auteur de cette thèse.

Deux autres expérimentateurs (B et C) collectaient le texte narratif oral. Leur rôle était de faire produire aux enfants une histoire personnelle et de leur faire remplir un questionnaire. Pour la narration orale, un investigateur s'occupait des garçons et une investigatrice des filles. Cette décision a été prise en pensant que les enfants et adolescents seraient plus à l'aise face à une personne de même sexe pour raconter une histoire personnelle (Annexe 8).

Une quatrième et cinquième personne (D et E), étaient affectées à la collecte du texte expositif oral. Leur rôle était de faire produire aux enfants un exposé sur les problèmes entre les gens et de leur faire remplir un questionnaire. Contrairement au texte narratif, il nous a semblé que le texte expositif n'exigeait pas une similitude de sexe entre les locuteurs et les expérimentateurs. Le type de texte expositif étant un texte plus scolaire, le choix de l'investigatrice pour les sujets garçons et filles ne semblait pas gênant, le monde de l'éducation étant encore aujourd'hui très féminin (Annexe 8).⁶³

Un sixième expérimentateur (F), était chargé du texte narratif écrit pour les filles et les garçons. Son rôle était de faire écrire aux enfants une histoire personnelle (Annexe 8). Contrairement au texte narratif oral, il nous a semblé que le texte narratif écrit n'exigeait pas une similitude de sexe entre les individus et l'expérimentateur. En effet, il est moins gênant de raconter une histoire personnelle à l'écrit car les scripteurs ne s'adressent pas directement à l'expérimentateur.

Une septième personne (G) était assignée à la collecte du texte expositif écrit pour les filles et les garçons. Son rôle était de faire écrire aux enfants un exposé sur les problèmes entre les gens (Annexe 8). Comme pour le texte expositif oral, nous n'avons pas fait de différence de genre entre les investigateurs et les scripteurs.

Une fois la collecte terminée, nous avons remercié les participants et les équipes pédagogiques en organisant un après-midi détente durant lequel nous leur avons proposé une présentation adaptée de la psycholinguistique ainsi que de notre travail. Des livres pour les BCD⁶⁴ ont été offerts aux écoles primaires. Une liste de livres avait été préalablement demandée au corps enseignant. Nous avons remercié chacun des collégiens en leur offrant deux places de cinémas valables dans le centre culturel de leur ville.

⁶³73,4% des personnes travaillant dans l'éducation, dans la santé et dans les actions sociales sont des femmes (INSEE, chiffres 2005 « *Femmes et Hommes – Regards sur la parité* », édition 2005).

⁶⁴ La BCD, Bibliothèque Centre Documentaire, est « *un lieu de vie, de travail et de détente, de rencontre et d'échange, de production et d'exposition* ». (<http://www.accreteil.fr>)

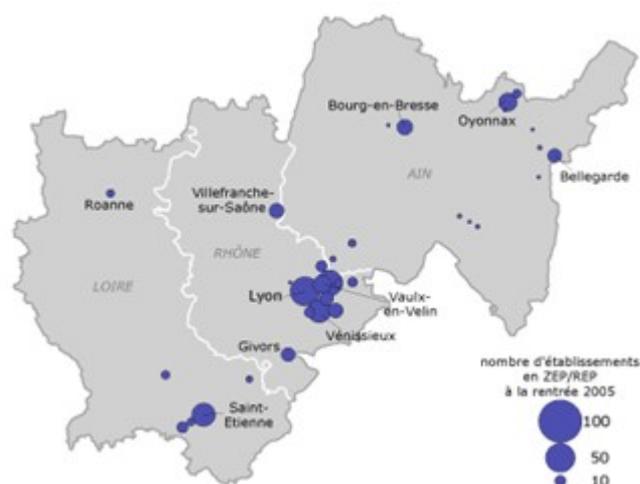
4.2. Sélection des lieux

En 1981, face à l'inégalité scolaire et à la persistance de l'échec scolaire parmi les élèves les plus défavorisés, le gouvernement de l'époque se joint aux équipes éducatives afin de mettre en place les ZEP (*Zones d'Éducation Prioritaire*). Ces zones sont créées dans certaines régions françaises afin d'introduire dans le système scolaire « l'idée de discrimination positive qui s'exprime de deux façons : « à des situations inégalitaires, solutions inégalitaires » ou « donner plus à ceux qui ont moins » (Chauveau, 2000:53). Le 1^{er} Juillet 1981, une circulaire du bulletin officiel de l'Éducation Nationale annonce la création du système ZEP qui était au centre des discussions des pédagogues depuis une douzaine d'années. Une ZEP constitue une zone rassemblant divers établissements, de l'école élémentaire, en passant par l'école primaire puis par le collège, au lycée professionnel (dans la région Rhône-Alpes, les lycées généraux n'entrent pas dans ce système) que les recteurs des Académies désignent. Les ZEP sont créées dans le but de répondre à un besoin d'améliorer les résultats scolaires dans des lieux où les indicateurs socioéconomiques et socioculturels étaient plutôt faibles. Il s'agit d'une concrétisation d'une politique de discrimination positive. Plusieurs critères furent utilisés afin de sélectionner les établissements d'une ville ou d'une région spécifique qui forment la *Zone d'Éducation Prioritaire*. Ils sont de plusieurs ordres : premièrement, les recteurs se basent sur une analyse géographique des résultats scolaires, sur des critères externes et internes de détermination des zones prioritaires (Circulaire n° 81-536 du 28 décembre 1981 (B.O.spécial n°1 du 21-01-1982)). En 1997, une politique de relance engendre la création des REP, *Réseaux d'Éducation Prioritaire*. Concrètement, les établissements classifiés ZEP deviennent alors des ZEP/REP et de nouveaux établissements, sous l'étiquette de REP, se joignent à une organisation ZEP/REP. Ces nouvelles écoles ou collèges REP profitent, par exemple, de projets communs sans pour autant bénéficier des avantages des ZEP tels que des financements et des professeurs supplémentaires ou des effectifs de classe abaissés.

Pour notre étude, nous nous sommes limitée à l'Académie de Lyon, qui, en 2004/2005, rassemble quarante REP dont : (a) cinq REP ne contiennent pas d'établissement en ZEP, (b) cinq REP avec certains établissements en ZEP, (c) trente REP composés d'établissements tous catégorisés ZEP. Généralement, les ZEP/REP se situent géographiquement plus fréquemment dans les zones et agglomérations urbaines que dans les zones rurales (Kociemba, 2004) ce qui

se confirme dans l'Académie de Lyon. Trois départements constituent l'académie de Lyon : l'Ain, la Loire et le Rhône. Six REP sont situés dans le département de l'Ain, huit dans celui de la Loire et vingt-six dans celui du Rhône (Cf. Carte 1).

Carte 1 : Carte des Réseaux d'Éducation Prioritaire (REP) de l'Académie de Lyon⁶⁵



En 2004, dans l'Académie lyonnaise, 383 établissements sont recensés dont 310 sont consacrés au premier degré (avec 150 écoles maternelles et 160 écoles primaires) et 73 sont voués au second degré (dont 50 collèges). Ces établissements regroupent 80 783 élèves dont 47 267 dans le premier degré (17 778 élèves en écoles maternelles et 27 489 en écoles primaires). Cette masse d'individus implique 7 286 postes dont 2 706 dans le premier degré et 4 580 dans le second degré.

Nous nous sommes focalisée sur des établissements d'une commune du Rhône de l'est-lyonnais pour collecter les données. Deux lycées, un lycée général et un lycée d'enseignement professionnel (LEP), huit écoles primaires et deux collèges se situent dans cette agglomération. Le LEP, trois des écoles primaires et un des collèges sont qualifiés d'établissements ZEP/REP. Nous avons décidé de nous concentrer sur ces établissements ZEP/REP excepté le LEP.

Une autorisation de l'Inspection Académique était nécessaire. Nous avons assisté à des réunions ZEP/REP afin de rencontrer les directeurs des écoles primaires, la principale du collège et les responsables académiques. La participation à ces réunions a permis une

⁶⁵ Source : <http://www.educationprioritaire.education.fr/>

meilleure connaissance du terrain et des principaux acteurs adultes tels que les instituteurs/institutrices, les responsables des centres sociaux, *etc.* Après avoir exposé notre projet et répondu à leurs questions, un dossier, sous-couvert de la Principale et de l'Inspectrice de l'Éducation Nationale a été envoyé au rectorat dans le but d'obtenir une autorisation administrative. L'autorisation officielle ainsi obtenue, nous avons pu commencer notre travail de terrain. La première étape était de réunir les équipes pédagogiques des écoles primaires ainsi que celles du collège afin de leur faire une présentation du projet. Une fois que les instituteurs/institutrices et les professeurs ont accepté de participer, nous avons pu commencer notre travail avec les élèves.

4.3. Sélection des participants

Le travail de terrain s'est poursuivi avec la sélection des participants (Septembre 2005) ce qui a débuté avec un premier questionnaire (Annexe 9) auquel tous les élèves de CM2, les collégiens de 5^{ème} et ceux de 3^{ème} ont répondu. Parmi les sujets potentiels, nous avons écarté, même si nous leur faisons passer le questionnaire pour éviter tout sentiment d'exclusion, différents profils. Premièrement, les élèves ayant redoublé au minimum deux fois ont été exclus. Deuxièmement, les élèves scolarisés dans les parcours SEGPA⁶⁶, Insertion⁶⁷ et relais⁶⁸ n'ont pas été retenus. Troisièmement, les élèves qui avaient plus de deux ans d'avance ont également été écartés de l'étude.

Ce questionnaire a été rempli par 390 individus au total et avait quatre objectifs principaux. Par le biais de ce questionnaire, nous souhaitons écarter certains types de populations. Nous

⁶⁶ Définition officielle de SEGPA (Section d'Enseignement Général et Professionnel Adapté) : « Ces sections accueillent des élèves présentant des difficultés scolaires graves et persistantes auxquelles n'ont pu remédier les actions de prévention, de soutien, d'aide et d'allongement des cycles dont ils ont pu bénéficier. Ces élèves ne maîtrisent pas toutes les compétences attendues à la fin du cycle des apprentissages fondamentaux. [...] Ils présentent sur le plan de l'efficacité intellectuelle des difficultés et des perturbations qui peuvent être surmontées ou atténuées que sur plusieurs années et qui [...] se traduisent par des incapacités et des désavantages tels qu'ils peuvent être décrits dans la nomenclature des déficiences, incapacités et désavantages (arrêté du 9-1-1989) ». (Circulaire n°96-167, du 20 juin 1996)

⁶⁷ Définition officielle des classes d'insertion : « Le dispositif s'adresse à des élèves qui ont des lacunes majeures dans les apprentissages fondamentaux et des difficultés à se situer dans les processus collectifs d'apprentissage ou dans l'institution scolaire. Il offre un cadre dans lequel les élèves puissent s'inscrire ou se réinscrire dans un projet de formation [...] pour favoriser l'insertion des élèves dans une formation qualifiante de niveau V, essentiellement en CAP, sous statut scolaire ou sous contrat d'apprentissage, sans exclure l'accès à un BEP. » (http://www.ac-creteil.fr/mission-college/cycle_orientation/troisieme_insertion.htm#1, circulaire n° 97-134 du 30 mai 1997)

⁶⁸ Définition officielle de classe relais : « elle est constituée d'élèves de collège, faisant parfois l'objet d'une mesure judiciaire d'assistance éducative, qui sont entrés dans un processus évident de rejet de l'institution scolaire et qui ont même souvent perdu le sens des règles de base qui organisent leur présence et leur activité au collège. » (Circulaire no 98-120 du 12 juin 1998)

voulions connaître la CSP⁶⁹ des parents des élèves afin d'écartier de l'étude les individus appartenant à une classe socioprofessionnelle favorisée. Nous voulions également connaître le pays de naissance des enfants et des parents afin d'exclure les enfants n'ayant pas toujours été scolarisés en France. Leur environnement linguistique nous intéressait aussi ; nous voulions exclure les bilingues afin de ne pas avoir à gérer des variables liées à une seconde langue. Enfin, nous nous sommes renseignée sur les éventuelles pathologies langagières dont certains individus auraient pu être atteints. Nous avons demandé aux enfants s'ils voyaient ou avaient vu un orthophoniste et nous leur en demandions la raison. Tous les individus ayant ou ayant eu une éventuelle pathologie du langage (du bégaiement à des pathologies plus sévères) ont été exclus.

À partir de ce premier questionnaire, nous avons pu sélectionner nos trois groupes de niveaux de scolarisation : 44 écoliers de CM2, 43 collégiens de 5^{ème} et 45 de 3^{ème} soit une population de 132 individus au total.

Ce questionnaire laisse transparaître deux critères principaux de sélection : la CSP et l'environnement linguistique de l'enfant auxquels est ajouté le niveau en expression française. Ces trois principaux critères sont développés par la suite.

4.3.1. CSP

L'un des buts principaux du questionnaire 1 était de nous renseigner sur les métiers exercés par les parents des enfants. Pour cela, nous avons demandé le métier des parents, que nous avons vérifié sur les bases de données des établissements regroupant les informations officielles concernant les situations familiales et socioprofessionnelles des individus. Ensuite, par le biais de documents de l'INSEE (<http://www.insee.fr>, Annexe 10), nous avons pu faire une correspondance pertinente et stable entre les métiers des parents et une classe sociale. Quatre classes sociales sont définies par l'INSEE : la classe favorisée A, la classe favorisée B, la classe moyenne et la classe défavorisée. Étant donné notre orientation scientifique, nous nous sommes alors intéressée seulement aux individus appartenant à la classe sociale défavorisée.

Nous tenons à signaler que la CSP ne constitue pas une VI dans notre étude dans la mesure où nous voulions connaître certains aspects du profil linguistique de cette population.

⁶⁹ Catégorie Socio-Professionnelle ; nous nous sommes basée sur les grilles de l'INSEE, Annexe X et Tableau X, que l'on trouve sur internet : <http://www.insee.fr>

Néanmoins, lorsque nous disposons de résultats concernant des populations favorisées pouvant être comparé avec les nôtres, nous les mentionnons. Nous avons notamment mentionné des travaux concernant la population francophone du projet Spencer (dont Jisa et Mazur, 2005 ; Mazur, 2005 ; Viguié-Simon, 2001) ou encore la population francophone des travaux de Gayraud (2000) (étude-pilote du projet Spencer). Ceci nous permet de montrer que notre population défavorisée suit, en règle générale, les mêmes tendances que ces populations favorisées (classe moyenne supérieure).⁷⁰

4.3.2. Environnement linguistique

À partir, principalement, du questionnaire 1, nous avons pu avoir une représentation de la situation linguistique du milieu familial et social des individus. Nous avons obtenu des informations sur les langues côtoyées ou parlées dans le cadre familial. Dans notre étude, nous avons voulu nous restreindre à des francophones natifs monolingues pouvant être, éventuellement, en relation avec une ou plusieurs langues étrangères dans leur milieu familial et social.

4.3.3. Niveau en expression langagière

Notre population a été scindée en deux groupes : nous avons sélectionné des individus ayant des difficultés d'expression langagière, ainsi que des individus ne présentant aucune difficulté d'expression ; nous avons tenu à assurer une parité de niveau d'expression.⁷¹ Ce critère du niveau d'expression langagière se base principalement sur les notes des individus dans les matières Français et Histoire-Géographie. Nous avons également pris en compte les avis et appréciations des professeurs enseignant ces deux matières et, en plus, des professeurs principaux, pour le collège. Pour obtenir ces diverses informations, nous avons organisé des réunions avec le corps enseignant. La population de notre étude se compose de trois groupes de rang scolaire comprenant chacun une quarantaine d'individus, avec autant d'individus ne présentant pas de difficulté d'expression que d'individus ayant des difficultés d'expression.

La sélection des individus de CM2 a été complétée par un test de repérage de dyslexie, l'Alouette (Annexes 11 et 12), afin d'écartier tout individu dont le niveau faible de lecture

⁷⁰ Néanmoins, comme nous le mentionnons dans la Partie Résultats, un décalage d'âge existe entre notre population de 3^{ème} et leur population 15/16 ans.

⁷¹ Ce critère de sélection n'est pas considéré comme une VI mais plus comme un contrôle.

laisserait penser à une éventuelle pathologie du langage. Les résultats de ces tests nous ont révélé qu'aucun élève de CM2 choisi ne présentait de dyslexie. Pour la passation de ce test, nous avons été aidée par deux étudiantes en orthophonie habituées à faire passer ce genre de test. Il leur a été demandé de suivre le protocole de base (Annexe 12). Afin ne pas influencer leur notation, chacune d'entre elles a vu aussi bien des garçons que des filles et des individus sans difficulté scolaire que des individus avec difficultés scolaires.

Nous n'avons pas fait passer ce test aux collégiens en pensant que s'ils étaient affectés par une pathologie langagière, ils auraient été détectés ; nous nous sommes totalement fiée aux systèmes mis en place par l'éducation nationale. En effet, l'établissement dans lequel sont scolarisés les collégiens de notre étude, présente une équipe médicale (se composant d'une psychologue scolaire, d'une infirmière, d'un médecin et d'une assistante sociale) qui possède les compétences pour repérer des enfants manifestant des pathologies langagières.

4.3.4. Sélection des participants : conclusion

Nous avons alors constitué trois groupes : des écoliers de CM2, des collégiens de 5^{ème} et de 3^{ème}. Nous avons tenu à une parité de genre et de niveau d'expression langagière. Le Tableau 7 résume ces informations.

[Tableau 7 : Constitution de la population]

Tableau 7 : Constitution de la population

	CM2	5^{ème}	3^{ème}
Âge			
Moyenne	10;9	12;7	15;2
Intervalle	[10;00-11;10]	[12;03-14;04]	[14;06-16;07]
Nombre total d'individus			
	44	43	45
Nombre total de filles			
	22	22	22
Nombre total de garçons			
	22	21	23
Nombre total d'individus sans difficulté d'expression			
	24	21	23

Nombre total d'individus avec difficultés d'expression		
20	22	22
Nombre total de textes		
176	172	180

Nous avons choisi d'associer le niveau de scolarisation et l'âge. En effet, nous nous penchons sur le degré de maîtrise de la langue dans la production textuelle, ce qui se traduit par des niveaux scolaires et pas seulement par l'âge (Schneuwly, 1988). Cette décision est également liée à un choix protocolaire à savoir le fait que nous avons sélectionné des individus ayant pu redoublé au maximum une fois dans leur cursus scolaire.

Les individus de CM2 ont été choisis pour représenter les plus jeunes de nos sujets et ce pour deux principales raisons. Premièrement, dans la classe de CM2, nous trouvons des individus ayant dépassé, en grande partie, les problèmes dus au développement grapho-moteur. Dans la mesure où nous leur avons demandé deux productions écrites sur les quatre textes, il est plus que nécessaire que les enfants n'aient pas d'importantes lacunes relatives au processus grapho-moteur de l'écriture. Deuxièmement, les enfants plus jeunes ne sont pas encore assez avancés dans le cursus scolaire pour être habitués à l'écrit de l'école. Les narrations produites par les enfants sans support et en situation monologique apparaissent plutôt tardivement et « *concernent d'abord les expériences vécues par le sujet avant de s'appliquer aux autres* » (Fayol, 1985:78). 9/10 ans semble alors être l'âge minimum dans la mesure où « *le schéma narratif apparaît utilisé dès sept ans* » en monologue et que « *l'établissement conscient d'une hiérarchie entre composantes de celui-ci ne se manifeste guère avant neuf-dix ans* » (Fayol, 1985:95). Ainsi, nous sommes certaine que les individus les moins experts de notre population possèdent les capacités de produire au moins ce type de texte. De plus, en CM2, les écoliers se familiarisent avec le type expositif *via* la présentation d'exposé sur des thématiques spécifiques. Les deux types de textes exigés sont ainsi connus un minimum par les plus jeunes de nos locuteurs/scripteurs.

4.4. Matériel de la collecte des données

4.4.1. Productions orales

Pour la collecte des données orales, quatre enregistreurs édirol R-1 ont été utilisés (Photo 1). Cet enregistreur offre des capacités d'enregistrement/lecture sans compression et d'une clarté

cristalline. L'appareil stocke les données sous fichiers stéréo sur une carte mémoire Compact Flash. L'édirol est, certes, équipé de micros internes mais offre tout de même la possibilité de se servir de micros externes *via* les entrées externes, solution que nous avons décidée de suivre comme le montre la Photo 2 qui présente une mise en situation de la collecte orale.

Photo 1 : Enregistreur édirol



Photo 2 : Collecte des données orales



Lors des sessions orales, l'individu (représenté en premier plan) était assis en face de l'expérimentateur collectant sa production (représenté en arrière plan) (Photo 2). L'expérimentateur écoutait la production de l'enfant en évitant toute intervention afin de maximiser les chances d'obtenir un monologue. Le sujet pouvait prendre le micro dans la main ou le laisser sur la table ; l'important était qu'il soit le plus à son aise possible.

4.4.2. Productions écrites

Pour la collecte des données écrites, nous avons utilisé quatre tablettes graphiques *intuos 3* de *WACOM* ainsi que quatre ordinateurs *DELL Latitude D610*. La Photo 3 présente le type de

tablette graphique utilisé pour la collecte de l'écrit. La tablette *WACOM* convient aussi bien aux gauchers qu'aux droitiers. La tablette est accompagnée d'un stylet *intuos 3 Grip Pen* qui possède une précision d'environ 0,25 mm dont la hauteur de reconnaissance maximale est de 6 mm. Chacune des tablettes était raccordée à un des ordinateurs (Photo 4).

Photo 3 : La tablette intuos 3 de WACOM



Photo 4 : Collecte des données écrites



La collecte des productions écrites, contrairement à celle des données orales, se déroulait en binôme. Les enfants (représenté au premier plan) était assis face à l'expérimentateur (représenté en arrière fond). Pour rendre le contexte de production le plus naturel possible, nous avons pris diverses décisions protocolaires : (a) les élèves ont réalisé deux versions, un brouillon et une version définitive, comme cela leur est habituellement demandé lors d'activité rédactionnelle à l'école ; (b) lors du choix de la tablette graphique, nous en avons sélectionné une des moins épaisses ; (c) le stylet sans fil accompagnant la tablette est composé d'une véritable mine à encre, tout comme les stylos standard, ce qui permet une adaptation

rapide des élèves, (d) la tablette permet d'écrire sur des feuilles ordinaires A4 (avec interligne et marge) auxquelles sont habitués les individus.

4.5. Restrictions méthodologiques

Deux principales restrictions méthodologiques peuvent être reprochées à notre étude : l'absence d'un groupe contrôle adulte et l'absence d'une comparaison avec une population plus favorisée.

Notre recherche a une perspective développementale qui peut être abordée de deux façons : soit nous choisissons de comparer notre population défavorisée à une population favorisée (en les considérant alors comme un groupe de comparaison) soit nous préférons nous pencher essentiellement sur l'étude de la population défavorisée. Nous avons favorisé la seconde approche qui nous paraît plus intéressante. Pour cette recherche, nous avons décidé de nous concentrer davantage sur un type de population afin d'en établir un profil linguistique. Cette étude complète le projet Spencer qui s'était lui focalisé sur une population plus favorisée (classe moyenne supérieure). Bien que cette thèse se focalise sur une population défavorisée, nous pensons à réaliser, dans un futur proche, une éventuelle étude comparant la population favorisée du projet Spencer et la population défavorisée du projet ReFlex.

L'équipe expérimentale a également préféré travailler sur une importante cohorte d'enfants et d'adolescents, étant donné que des travaux précédents montrent que l'adolescence est une période clef du développement tardif. Le but est d'établir certains traits caractéristiques du profil linguistique de cette population. Nous avons tenté d'obtenir une population adulte défavorisée mais nous nous sommes heurtées à des problèmes de faisabilité. Nous sommes allées prospecter dans des LEP ainsi que dans des centres sociaux. Il s'est avéré qu'il était délicat de rentrer en contact avec cette population. Divers problèmes difficilement gérables sont apparus en plus de toutes les obligations de notre projet global. Ainsi, par exemple, les jeunes adultes n'étaient guère enthousiastes à l'idée d'avoir affaire à des universitaires ; la peur d'être analysés et évalués en était sans doute la cause. Si ce travail de thèse se concentre sur les enfants et adolescents, il n'en reste pas moins que la constitution d'un groupe contrôle adulte est à envisager. Néanmoins, la collecte de productions langagières d'adultes appartenant à une CSP défavorisée doit être le thème d'un projet spécifique incluant une

collaboration avec des sociologues et des professionnels habitués à encadrer de tels adultes afin de gérer au mieux la situation.

Ce Chapitre nous a permis de présenter le protocole expérimental de ce travail de recherche afin de mieux comprendre dans quelles circonstances ont été produits les textes que nous analysons. Une fois les productions collectées, nous avons dû les informatiser afin de la travailler. Le Chapitre 5 qui suit fait le point sur les logiciels utilisés pour le traitement des données.

Chapitre 5

Exploitation des Données

Ce chapitre est consacré à l'exploitation des données ; les logiciels utilisés pour collecter, coder et analyser les données y sont présentés. Ces logiciels sont décrits globalement en nous concentrant principalement sur les fonctionnalités qui nous ont été utiles. Ainsi, la première section de ce chapitre (p. 132) concerne la transcription des données écrites avec la présentation du logiciel *Eye and Pen*©. La seconde section (p. 133) se concentre sur le logiciel *Transcriber* permettant l'exploitation des données orales. Puis, la troisième section (p. 135) aborde le logiciel *CLAN* utilisé à la fois pour les données écrites et orales.

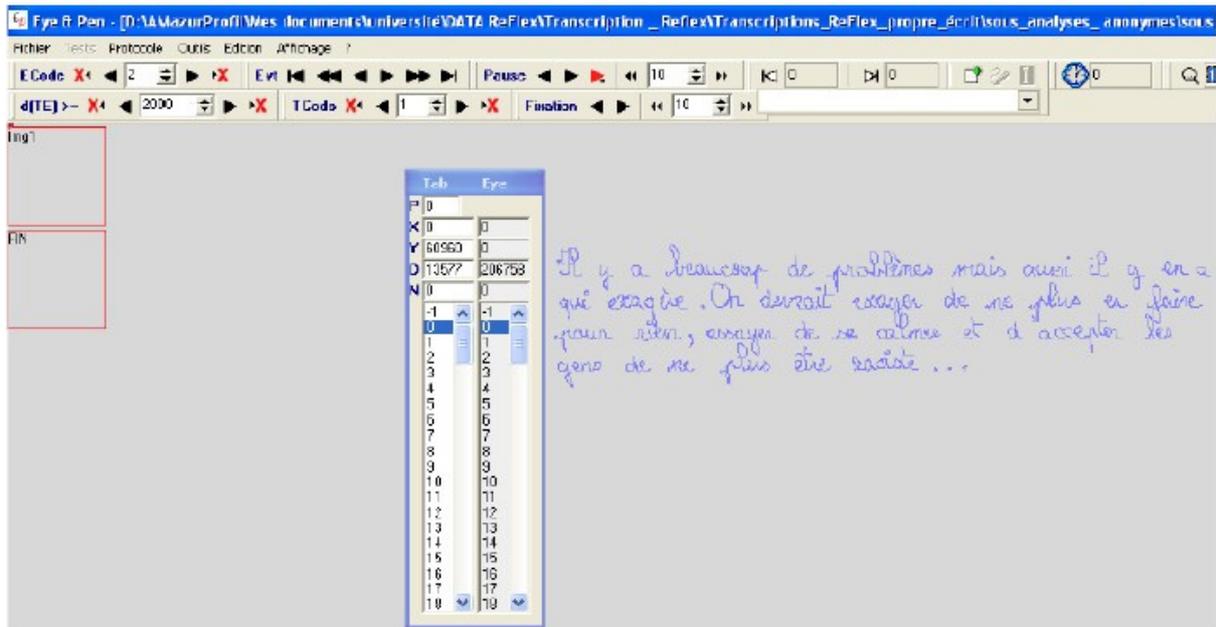
5.1. Le logiciel *Eye and Pen*©

Les productions écrites ont été traitées de deux manières différentes étant donné que les deux versions, *brouillon* et *propre*, n'ont pas été analysées de la même façon. Pour notre travail de thèse, nous nous sommes concentrée sur les versions *propres*. Ces versions propres ont été récupérées, tout comme les brouillons, par le biais de la tablette graphique grâce au logiciel *Eye and Pen*©, développé par Chesnet et Alamargot (2005), qui est une évolution du logiciel G-Studio (Chesnet, Guillabert et Esperet, 1994). *Eye and Pen*© a été conçu afin que les chercheurs en sciences cognitives et en linguistique puissent étudier les processus en temps réel lors de la production écrite. Ce logiciel peut enregistrer à la fois l'activité d'écriture et à l'activité oculaire. Pour notre étude, nous n'utilisons que la partie *Pen* du logiciel. Le texte est digitalisé par le biais de la tablette, puis affiché sur l'écran de l'ordinateur auquel la tablette

est reliée. Ainsi, il est possible de reconstituer à l'écran et à volonté la trace laissée par le stylet.

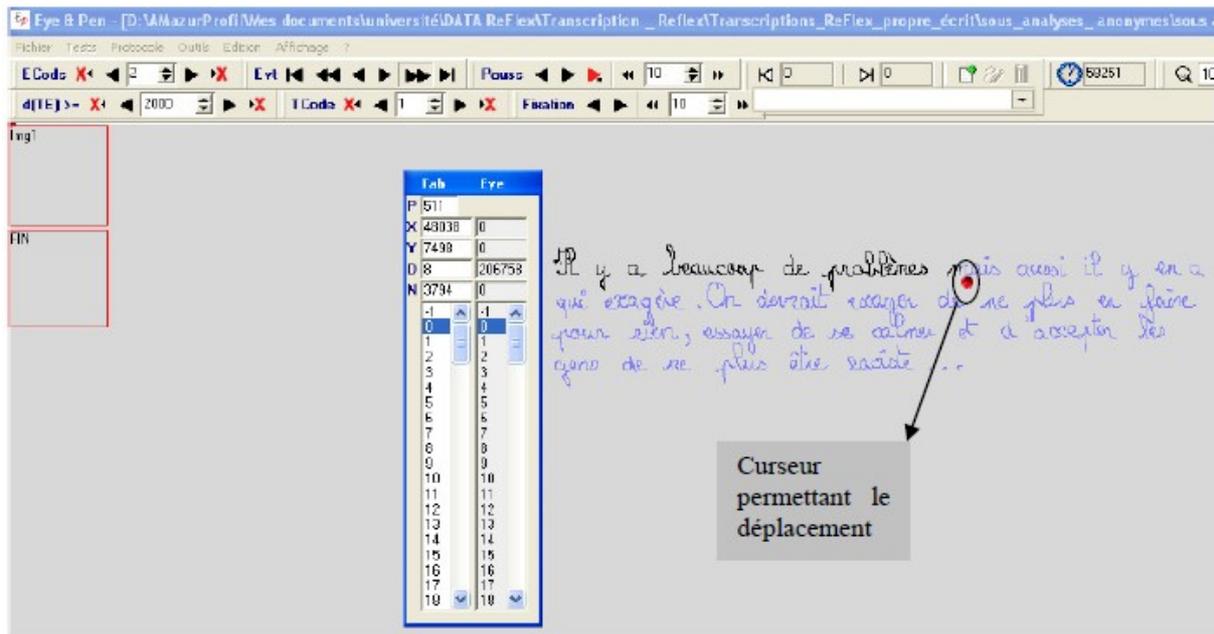
Lors de l'ouverture d'un fichier Eye and Pen©, la production apparaît (Illustration 1).

Illustration 1 : Fichier Eye and Pen©



Le logiciel permet de *rejouer* le déroulement de la rédaction en temps réel (Illustration 2). Lorsque nous souhaitons effectuer une telle manipulation, un curseur, représentant la mine du stylo, apparaît et rejoue les mouvements du stylo.

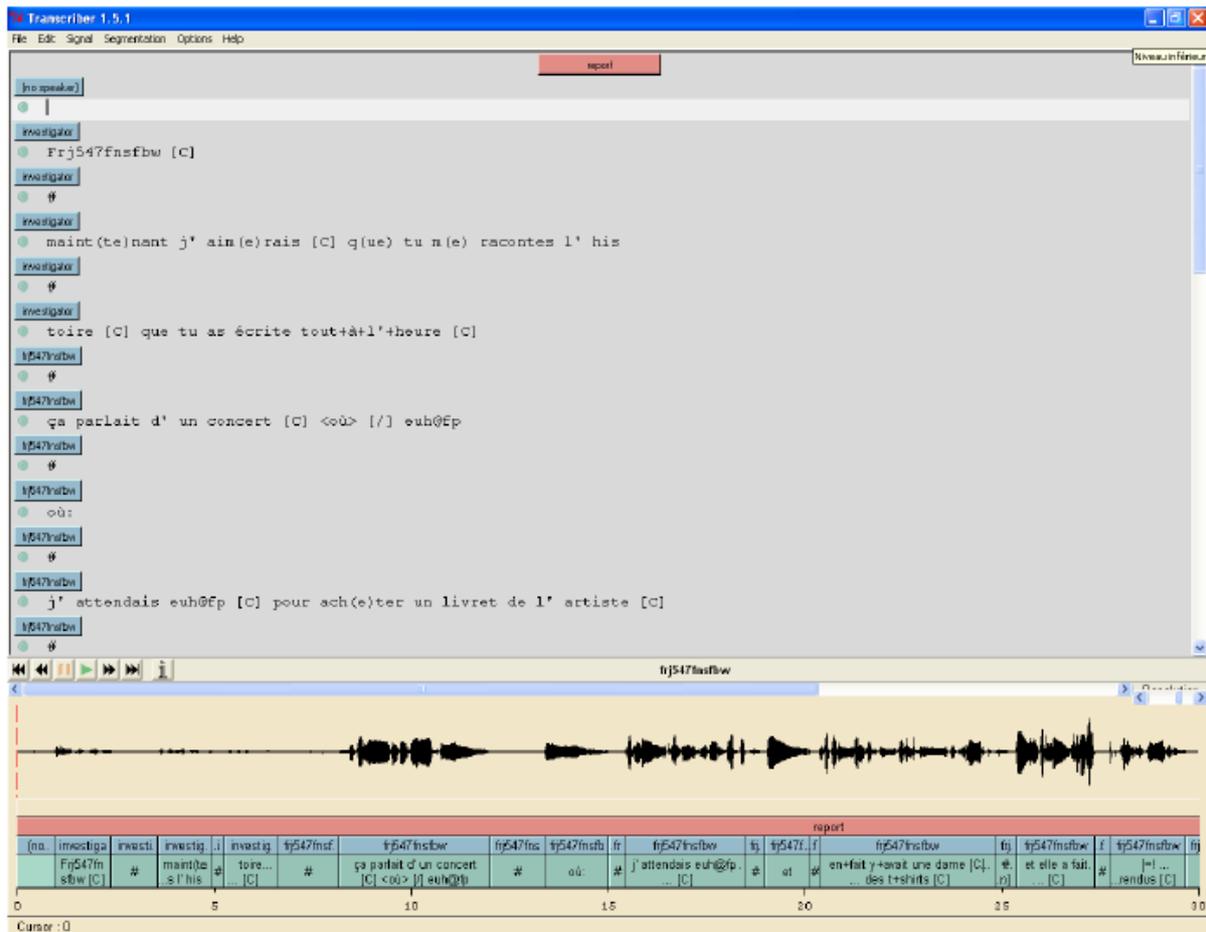
Illustration 2 : Représentation du déroulement de la production



5.2. Le logiciel *Transcriber*

Les fichiers audio originaux ont été traités en quatre temps. Premièrement, ils ont été normalisés et passés en mono avec le logiciel *Praat* (Boersma et Weenink, 2009). Deuxièmement, ils ont été traités par le programme *StkCVnoFo.tcl* pour faciliter le traitement des pauses. Ce logiciel a été développé par F. Pellegrino (Laboratoire Dynamique Du Langage) et permet le repérage automatique des zones de pauses vides. Troisièmement, les productions orales ont été transcrites grâce au logiciel *Transcriber*, qui est le seul logiciel, à notre connaissance, à pouvoir s'associer au programme *StkCVnoFo.tcl*, ce qui nous a permis de traiter les processus en temps réel tels que les pauses (Illustration 3). Les conventions de transcriptions des textes oraux se trouvent en Annexe 13. Quatrièmement, une fois les transcriptions réalisées dans *Transcriber*, elles ont été exportées dans le logiciel *CLAN* grâce à un filtre adapté (réalisé par Florence Chenu, Laboratoire Dynamique Du Langage).

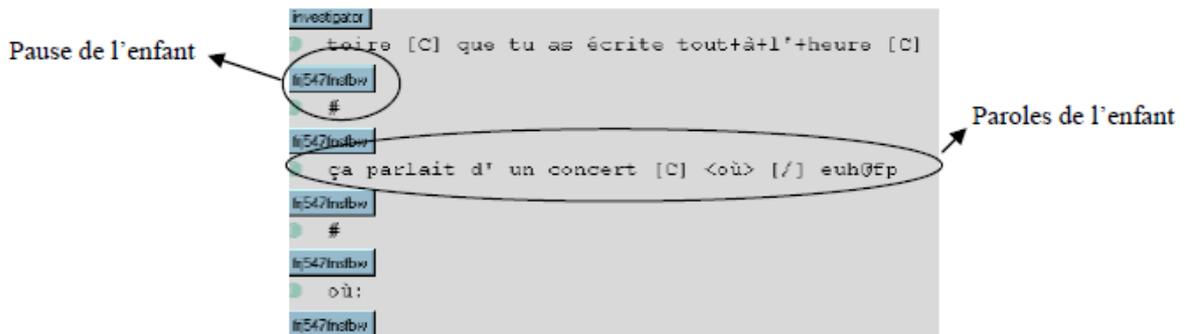
Illustration 3 : Fenêtre d'un fichier *Transcriber*



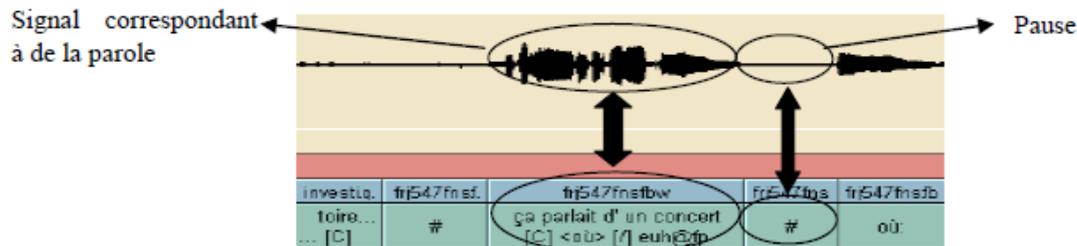
Le logiciel *Transcriber* permet de structurer et de transcrire des enregistrements. La première étape est la segmentation de la bande son, la seconde correspond à la transcription orthographique de la parole de chaque locuteur du fichier son. Il s'agit de la transcription proprement dite. L'étape de segmentation vise à structurer les enregistrements, c'est-à-dire à segmenter et à décrire le signal acoustique à différents niveaux jugés pertinents pour le traitement ultérieur (Illustration 11). Il s'agit ici principalement de noter l'identité des locuteurs, de séparer les tours de paroles, de délimiter les séquences de bruits et de paroles. L'enregistrement est découpé en tours de paroles (Illustration 4) délimités par des pauses.

Illustration 4 : Exemple de segmentations et d'identification de tours de parole

Partie 1 : haut de l'écran



Partie 2 : bas de l'écran



La fenêtre principale de *Transcriber* se décompose en deux parties principales : (a) dans la partie 1, situé en haut de l'écran (Illustration 4), se trouvent les prises de parole des locuteurs où peuvent être placées les transcriptions orthographiques correspondantes. Le symbole # représente une pause de plus de 250 millisecondes ; (b) la partie 2, située en bas de l'écran (Illustration 4), présente une vision acoustique du message et propose également une mise en relation entre la vision acoustique et la transcription. Ainsi dans cette Illustration 4, nous distinguons les sections suivantes : (a) les paroles de l'investigateur (« *toire [C] que tu as écrite tout+à+l'+heure [C]* ») ; (b) la pause de plus de 250 millisecondes de *frj547fnsfbw* (#) ; (c) le tour de parole de *frj547fnsfbw* (« *ça parlait d'un concert* ») (d) puis une autre pause de 250 millisecondes de *frj547fnsfbw* (#) et enfin (e) de nouveau une prise parole de *frj547fnsfbw* (« *où :* »).

5.3. Le logiciel CLAN

Les textes écrits et oraux ont été ensuite transférés dans *CLAN*, outil de traitement de corpus, selon les conventions *CHILDES* (*Child Language Data Exchange System*, <http://childes.psy.cmu.edu/>). Ce logiciel permet trois actions principales : (a) la transformation des données manuscrites en données dactylographiées, (b) le codage, et (c) le dénombrement d'occurrences. *CLAN* permet la transcription de productions orales puis grâce à certaines

adaptations permet le traitement des données écrites. Des codes sont insérés afin de coder des phénomènes typiques tels que les erreurs d'orthographe, les majuscules, les révisions à l'écrit *etc.* Les fichiers CLAN se composent de deux parties principales : un en-tête (Illustration 5) et le corps de la production (Illustration 6).

Illustration 5 : L'en-tête et ses fonctions principales

@Begin		
@Languages: fr	← Langue dans laquelle est produit le texte	
@Participants: SBJ Chérifa Target_Child, EXP Florence Investigator	} Résumé des informations sur les participants et les expérimentateurs	
@ID: fr Jisa_ReFlex SBJ 11;00.19 female dsw low Target_Child		
@Birth of SBJ: 26-APR-1995		
@Age of SBJ: 11;00.19	} Informations sur les participants (sexe, âge, classe, <i>etc.</i>)	
@Sex of SBJ: female		
@Group of SBJ: gs		
@Grade-Level: CM2		
@Expression Level: without_difficulty		
@Text type: expo_written		
@Version: final		
@Order: B		
@Session: 2		
@Time Duration: 00:03:26.76		
@Date: 22-MAY-2006	} Informations sur le lieu de la collecte et les expérimentateurs	
@Location: école primaire XXX		
@Investigator: Florence Chenu		
@Transcriber: Audrey Mazur		
@Reliability of transcription:		
@Coder:		
@Reliability of coder:		
@Number of Clauses: 08	} Informations syntaxiques : nombre de clauses et d'Unités T du texte	
@Number of T-Units: 03		

L'en-tête permet de donner à chaque fichier des informations importantes aussi bien sur le sujet, le type de texte, les transcripteurs, *etc.* La seconde partie du fichier CLAN présente donc le corps du texte, la production de l'individu (Illustration 6).

Illustration 6 : Corps du texte

@BG: TU1		
*SBJ: il+y+a@cok beaucoup de problèmes .	→ @cok signale que l'enfant a mis une majuscule conformément à l'usage standard	
@EG: TU1		
@BG: TU2		
*SBJ: mais aussi il y en a .		
*SBJ: qui exagère@err [: exagèrent][^.] .	→ Signal d'erreur orthographique	
@EG: TU2		
		→ [] signale la ponctuation de l'enfant

Dans ce fichier, nous mettons divers codes soit exigés par *CLAN* soit dont nous avons besoin tels que **SBJ* indiquant que la clause qui suit correspond aux direns/écrits du sujet. Les productions des sujets ont été segmentées en clauses (Conventions de transcription des textes écrits en Annexe 13 et 14).⁷² Ces productions sont ensuite codées sous ce même logiciel qui permet la mise au point d'un fichier de codage automatisant le codage et réduisant ainsi les erreurs (Illustration 7).

Illustration 7 : Une partie du fichier de codage

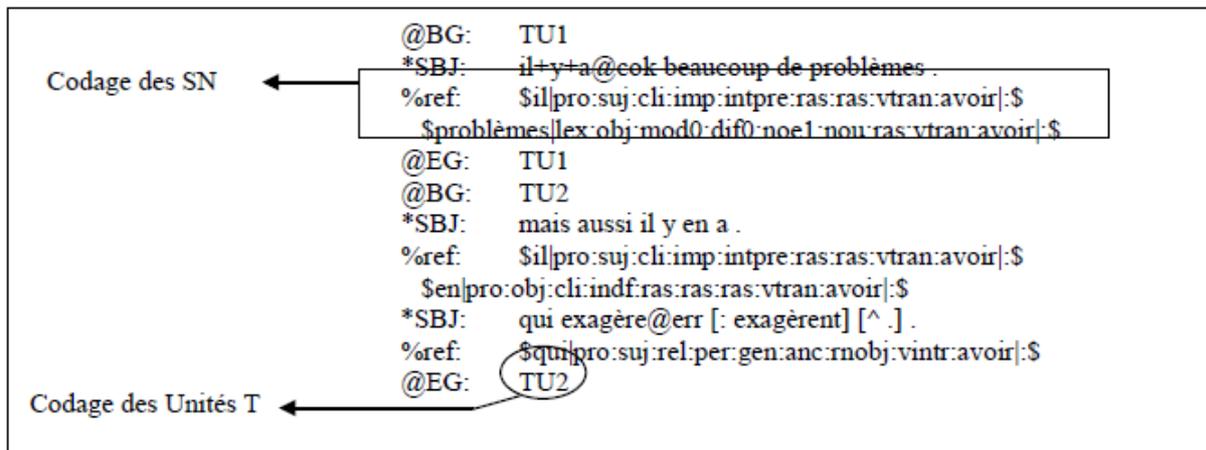
28-Apr-08 11:36	D:\AMazurProfil\Bureau\ref\%ref_ReFlex-1.cut
\ +b50 +l1 +d +s1	
%ref:	
\$	
?	Explicitation de la forme, exemple : le garçon
lexique:	Information sur la nature de la forme codée : un SNL
objet:	Indication de la position syntaxique de la forme
1 nœud:	Première information sur la constitution du SNL
0 modifieur:	Deuxième information sur la constitution du SNL
0 modifieur différent:	Troisième information sur la constitution du SNL
nouvelle:	Information sur la valeur référentielle de la forme
ras:	Information sur la position syntaxique de la forme maintenue
? :\$	Explicitation de la forme verbale de la clause, exemple : être

Nous ajoutons au fichier de transcription *CLAN*, une ligne nommée *%ref*, qui contient les informations entrées dans le fichier de codage pour coder les SN (Illustration 8). Les productions ont également été segmentées en UT.⁷³

Illustration 8 : Partie d'un fichier transcrit et codé

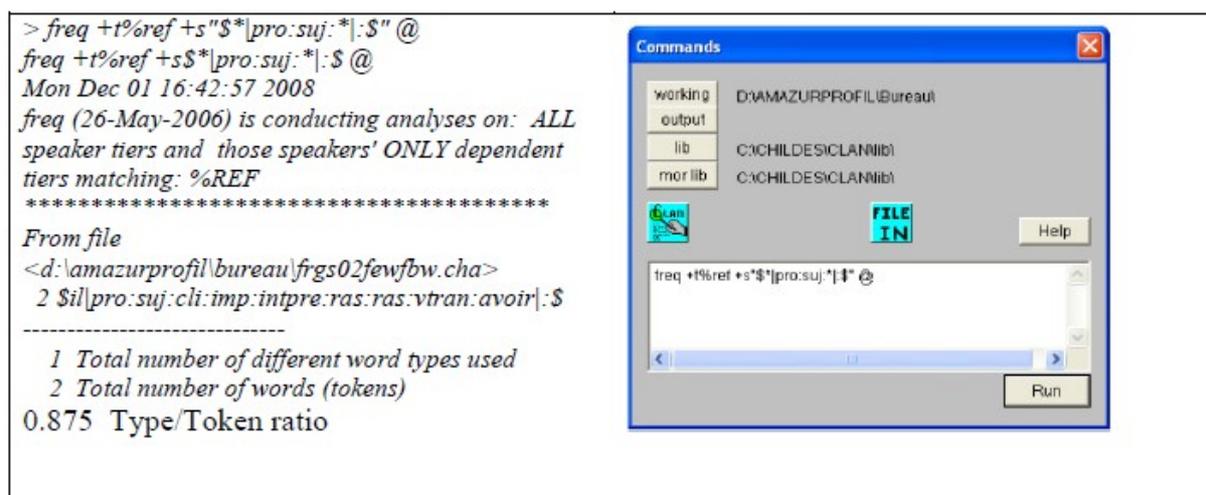
⁷² Le choix de segmenter en clause est abordé dans le Chapitre 6.

⁷³ Le choix de segmenter en UT est expliqué dans le Chapitre 6.



Ensuite, *CLAN* permet de questionner ces lignes de codages ce qui permet de comptabiliser des catégories (Illustration 9).

Illustration 9 : Exemple d'un questionnement



Ainsi, dans cet exemple, nous avons par le biais de la fenêtre de commande *CLAN* demandé au logiciel de nous donner le nombre d'occurrences de pronoms en position sujet codé sur la ligne *%ref*. Pour se faire, nous avons recours à la commande `freq freq + t%ref + s$*|pro:suj:*|:$"`

Ce Chapitre 5 était alors consacré à la présentation des logiciels – *Eye and Pen*©, *Transcriber* et *CLAN* – utilisés pour collecter, coder et analyser les données. Ces logiciels ont été décrits partiellement puisque nous nous sommes strictement focalisée sur les fonctions des logiciels dont nous nous sommes servie. Le Chapitre 6 qui suit apporte des précisions quant aux choix et positions théoriques concernant le codage des pronoms et des SNL.

Chapitre 6

Codage

Après la collecte et la transcription, nous passons au codage des données. Dans un premier temps, afin d'obtenir des indicateurs généraux (nombre de mots, clauses et d'UT), nous avons divisé les textes en clauses (ou propositions) et en UT (Unité Terminale). Puis, dans un second temps, nous avons attribué à chaque clause une ligne de codage spécifique nommée *%ref*. Pour chacun des SNL et des pronoms, nous avons donné des informations sur sa fonction syntaxique dans la clause, sa nature et son statut informationnel.

6.1. Segmentation en clause et Unité Terminale

Dans les deux modalités de production, orale et écrite, cette étude *off line* du langage implique une difficulté de méthodologie qui est la suivante : quelle unité choisir pour segmenter la production langagière ? Une unité est immédiatement bannie, celle de la phrase dans la mesure où « *elle n'est que l'approximation graphique, intuitive et informelle d'une unité de langue* », et donc « *un instrument à peu près inefficace lorsqu'il s'agit de segmenter un discours oral* » mais également lorsqu'il s'agit d'un discours écrit (Berrendonner, 1991:25 ; Berrendonner et Reichler-Béguelin, 1989).

Nous suivons la position de Coirier *et al.* (1996) qui suggèrent de choisir comme unité de base pour la description et le traitement du contenu des textes la proposition (ou la clause) qui « *apparaît comme l'unité la plus pertinente* » (Coirier *et al.*, 1996:220). Nous considérons, comme Coirier *et al.* (1996), Kintsch (1974), Le Ny (1979) ou Gineste et Le Ny (2002) entre autres, que la proposition est « *une configuration sémantique de base sous laquelle fonctionne le langage* » (Gineste et Le Ny, 2002:118). Elle permet aux individus de raisonner, de s'exprimer et de comprendre les messages de leurs congénères. La proposition ou la clause est une unité de signification constituée de prédicats et d'arguments. Les prédicats assignent des propriétés aux arguments et ces derniers sont des entités référentielles qui peuvent constituer des objets, des êtres, des idées, *etc.* Nous pouvons donner les exemples de Coirier *et al.* (1996:14) réunis en (60) et (61) afin d'explicitier les notions de propositions, prédicats et arguments.

(60) *Aboyer (chien) [le chien aboie]*

La proposition contient un argument et un prédicat qui définit une propriété à l'argument (chien) auquel il est joint.

(61) *Ronger (chien, os) [le chien ronge un os]*

La proposition contient un prédicat « ronger » qui définit et met en relation deux arguments (chien) et (os).

Les textes des individus ont également été découpés en unité terminale (UT). Aparici, Tolchinsky et Rosado (2000) suivent la définition de Hunt (1970) pour définir l'UT, qui apparaît comme un bon critère d'évaluation pour l'évaluation du développement syntaxique dans la production textuelle. Ainsi, l'UT est définie comme une unité composée d'une clause principale et de toutes ses clauses dépendantes telles que ses subordonnées, *etc.* (Hunt, 1970). Il s'agit d'une «*minimal terminable unit*» (Hunt, 1970:4) entre la clause et la phrase. En nous basant sur les considérations de Hunt (1970) et d'Aparici *et al.* (2000), nous avons considéré comme UT les unités intégrant des clauses qui entretiennent des relations de dépendance à savoir : (a) les principales et leurs subordonnées ; (b) les clauses liées aux suivantes par une relation de coordination avec élision du sujet ou du sujet et du verbe. L'UT apparaît comme un bon indicateur pour évaluer la complexité syntaxique des textes d'adolescents (Berman, 1998 ; Ravid, 2005 ; Viguié-Simon, 2001). L'exemple (62) propose un extrait d'un texte transcrit puis codé en UT.

(62) *Texte codé (CM2 n°02, expositif écrit, E/O)*

<i>Début de l'UT 1</i>	←	@BG: TU1
		*SBJ: <i>il y a beaucoup de problèmes.</i>
<i>Fin de l'UT1</i>	←	@EG: TU1
<i>Début de l'UT 2</i>	←	@BG: TU2
		*SBJ: <i>mais aussi il y en a.</i>
		*SBJ: <i>qui exagèrent.</i>
<i>Fin de l'UT 2</i>	←	@EG: TU2
<i>Début de l'UT 3</i>	←	@BG: TU3
		*SBJ: <i>on devrait essayer.</i>
		*SBJ: <i>de ne plus s'en faire pour rien.</i>
		*SBJ: <i>essayer de se calmer.</i>
		*SBJ: <i>et d'accepter les gens.</i>
		*SBJ: <i>de ne plus être raciste.</i>
<i>Fin de l'UT 3</i>	←	@EG: TU3

Ainsi cet exemple de production codée comporte huit clauses et trois UT : l'UT 1 contient une clause *il y a beaucoup de problèmes* ; l'UT 2 contient deux clauses 1. *mais aussi il y en a* et 2. *qui exagèrent* ; l'UT 3 contient cinq clauses 1. *on devrait essayer* 2. *de ne plus s'en faire pour rien* 3. *essayer de se calmer* 4. *et d'accepter les gens* et 5. *de ne plus être raciste*.

6.2. Codage des syntagmes nominaux lexicaux et des pronoms

6.2.1. La fonction syntaxique de l'unité

Dans un premier temps, après avoir mentionné le type de SN à coder (SNL ou pronom), nous avons indiqué la fonction syntaxique de la forme. Nous avons distingué les sujets grammaticaux (de verbes transitifs, intransitifs et de copules), les objets ou les fonctions autres. Nous avons pris la décision de ne considérer comme verbes transitifs que les verbes qui se construisent avec un objet direct. Nous nous sommes basée sur l'utilisation que l'enfant faisait du verbe. Un verbe transitif (63) et (64) peut être défini « *si et seulement si il peut se construire avec les deux termes exactement de la même façon qu'un verbe d'action prototypique se construit avec les deux constituant nominaux représentant l'agent et le patient* » (Creissels, 2006:283).

(63) [*Le chien*] rapporte [*le gibier*] à son maître.

(64) [*Ce film*] enthousiasme [*les spectateurs*].

La catégorie des verbes intransitifs regroupe tous les verbes ayant une construction sans complément (65) ou les verbes ayant une construction avec des compléments non objet direct (66). Les *verbes copules* sont tous les verbes appelés traditionnellement verbes d'état (être, sembler, etc.).

(65) *Il court.*

(66) *Il court vers sa mère.*

Dans notre étude, la catégorie des objets rassemble les objets strictement directs. Toutes les autres fonctions syntaxiques telles qu'objet indirect, objet second ou encore complément circonstanciel sont regroupées dans la catégorie autre.

Dans un second temps, nous avons donné des informations sur la nature des SNL et des pronoms.

6.2.2. La nature de l'unité

6.2.2.1. Les syntagmes nominaux lexicaux

Ainsi, pour les SNL, nous donnons (en valeur réelle) le nombre de modifieurs, le nombre de modifieurs différents et le nombre de nœuds lexicaux dans un même SNL (67). C'est ce dernier aspect que nous avons étudié dans ce travail de thèse.

(67) Une **bande de filles** est venue me « parler ». (3^{ème} n°07, narratif écrit)

Cet exemple contient un SNL, *une bande de filles*, constitué de deux nœuds lexicaux : *bande* et *filles*.

Nous souhaitons également discuter de trois cas pouvant poser problème lors du codage des unités : les *mots composés*, les *conversions* et certaines *expressions exclues de la catégorie de SNL*.

Pour notre étude, il est utile d'identifier les *noms composés* afin de ne pas les confondre avec des SN formés selon les règles productives de la construction de SN.⁷⁴ La composition apparaît alors comme étant le procédé « *qui consiste à former un mot en assemblant deux ou plusieurs mots* » (Lehmann et Martin-Berthet, 2000:167). Les éléments entrant dans la formation de noms composés sont des unités lexicales autonomes par ailleurs (Lehmann et Martin-Berthet, 2000 ; Riegel *et al.*, 1994/2002). Nous pouvons trouver diverses constructions (Creissels, 2006 ; Lehmann et Martin-Berthet, 2000 ; Riegel *et al.*, 1994/2002) :

- Nom+nom : chou-fleur, timbre-poste, *etc.* ;
- Nom+préposition+nom : chemin de fer, salle de bain, salle à manger, *etc.* ;
- Nom+adjectif : coffre-fort, professeur principal, cours moyen, *etc.* ;
- Adjectif +nom : petit copain, grandes vacances, gros mot, *etc.* ;
- Verbe+nom : brise-glace, *etc.* ;
- Verbe+verbe : cache-cache, savoir-faire, *etc.*

⁷⁴ En Annexe 15, est présentée la liste des noms composés.

Trois critères sont proposés afin de reconnaître un syntagme figé d'un syntagme libre (notamment Lehmann et Martin-Berthet, 2000) : le critère référentiel, le critère sémantique et le critère syntaxique.

À partir du critère référentiel, nous pouvons dire que le figement d'une unité lexicale repose sur l'existence d'un référent unique. Selon Goosse (1993/2004:176), un mot « *quoique formé d'éléments graphiquement indépendants, est composé dès le moment où il évoque dans l'esprit non les images distinctes répondant à chacun des mots composants mais une image unique. Ainsi les composés **hôtel de ville**, **pomme de terre**, **arc de triomphe** éveillent chacun dans l'esprit une image unique et non les images distinctes d'**hôtel** et de **ville**, de **pomme** et de **terre**, d'**arc** et de **triomphe**. ».*

Selon le critère sémantique, un nom composé n'accumule pas les sens de chaque item composant le composé mais a son propre sens. C'est ainsi, que dans *pomme de terre*, les sens des éléments *pomme* et *terre* ne sont pas accumulés pour obtenir celui du composé. Cependant, ce critère peut poser problème pour des SN tels que *chaise longue* ou *mauvaise herbe* dans lesquels les composants sont nécessaires au sens du composé mais des particularités spécifiques s'ajoutent au composé.

Le critère syntaxique concerne le test de la substitution paradigmaticque et celui des modifications syntaxiques telles que l'ajout d'un modifieur dans le SN. Dans le syntagme figé *chaise longue*, nous ne pouvons remplacer l'un des deux composants sans impliquer un changement de sens **chaise courte* (substitution paradigmaticque) ou nous ne pouvons pas dire **une chaise très longue* (modifications syntaxiques).

Les conversions sont des adjectifs (68), des pronoms (69), des infinitifs (70), des prépositions (71), les adverbes (72) et les conjonctions (73) qui se comportent comme des noms. En effet, ils perdent les caractéristiques relatives à leur classe syntaxique et prennent certaines caractéristiques du nom (Lehmann et Martin-Berthet, 2000).

(68) **Les grands** m'ont embêté.

(69) **Un rien** le vexe.

(70) **Se battre** n'est pas bien.

(71) (au foot) Ils se sont pris **un contre**.

(72) Une vente de **dessous** féminins.

(73) Avec **des si**, on refait le monde.

Toutes les formes qui prennent les caractéristiques des noms suite à une conversion, sont considérées comme des formes lexicales.

En codant la nature des SNL, nous avons rencontré des cas qui demandaient réflexion, des *expressions exclues de la catégorie de SNL*. En effet, trois types d'expressions idiomatiques n'ont pas été considérés comme des SNL étant donnée la perte de leurs caractéristiques nominales : le type [Verbe+Nom], le type [préposition+nom ou nom+préposition] et les noms utilisés comme adjectifs.

Le type [Verbe+Nom] (*dire du mal, avoir confiance, etc.*) dans lequel le nom perd totalement ou partiellement ses particularités nominales est fréquent. Parmi ces locutions verbales, Lehmann et Martin-Berthet (2000) proposent de distinguer deux types de locutions verbales. Le premier type concerne les syntagmes [verbe+nom] sans déterminant (*avoir lieu, avoir confiance, avoir envie, etc.*). L'absence de déterminant reflète la perte d'une des caractéristiques principales du nom. De plus nous ne pouvons ajouter de déterminant sans que le sens ne varie (74) versus (75).

(74) **La scène a le lieu dans la forêt*

(75) *La scène a lieu dans la forêt.*

Le deuxième type est celui des syntagmes [verbe+déterminant+nom] (*perdre la face, prendre la mouche, etc.*). Les principaux critères de figement sont l'invariabilité du déterminant, l'impossibilité de la mise au passif et de la pronominalisation du nom puis le sens non compositionnel (76).

(76) *Pierre a pris la mouche. / * Pierre a pris les mouches. / * La mouche a été prise par Pierre. / *Pierre l'a prise.*

Le type [préposition+nom ou nom+préposition] n'est pas à considérer comme un SNL dans la mesure où les noms, ici encore, perdent leurs spécificités syntaxiques nominales (Lehmann et Martin-Berthet, 2000). Les syntagmes [préposition+nom] comme *au lieu de, à la rigueur, etc.* sont à considérer comme des adverbes. Les noms ne fonctionnent plus entièrement en tant que tels puisqu'ils ne sont plus pronominalisables, qu'ils sont invariables et que nous ne pouvons pas les compléter avec un adjectif (77).

(77) *Il s'amuse au lieu de faire ses devoirs. / *Il s'en amuse de faire ses devoirs. / *Il s'amuse des lieux de faire ses devoirs. / *Il s'amuse au grand lieu de faire ses devoirs.*

Les syntagmes [*nom+préposition*] tels que *grâce à, faute de, etc.* sont à considérer comme des prépositions. Comme précédemment, les noms de ces types de syntagmes ne sont plus à considérer comme des noms car ils en ont perdu les caractéristiques nominales (78).

(78) *Grâce à ton aide, j'ai réussi. / *Les grâces à ton aide, j'ai réussi. / *Grande grâce à ton aide, j'ai réussi.*

Les noms utilisés comme adjectifs sont également des structures à ne pas considérer comme des SNL. Nous avons rencontré des morphèmes qui, à l'origine, appartiennent à la classe des noms et qui sont transférés dans la classe des adjectifs ; comme précédemment, le morphème perd ses qualités de nom (79).

(79) *Son pull est bleu roi. / *Son pull est bleu le roi. / *Son pull est bleu grand roi.*

6.2.2.2. Les pronoms

Lors du codage des données, même si nous avons pris en considération forme et fonction des pronoms, nous n'avons pas forcément suivi les classements traditionnels des pronoms.

Premièrement, nous notions si le pronom était un clitique, un pronom libre ou relatif.

Deuxièmement, nous faisons une différence entre les pronoms personnels (spécifique *versus* générique), les pronoms impersonnels (Maillard, 1985) ; distinction des impersonnels accidentel et intrinsèque (Corblin, 1991 ; Gaatone, 1991), le pronom *ça* (anaphorique, générique, indistinct, présentatif et phrastique) et les pronoms indéfinis (du pronom *on*, en distinguant ses différentes valeurs (Rey-Debove, 2001), aux formes libres telles que *quelqu'un*). Néanmoins, ces informations certes disponibles ne sont pas exploitées dans ce travail de thèse.

Concernant les cas spécifiques de l'ellipse du sujet, cinq situations ont été recensées (Section 1.3, p. 40). Il existe : (a) l'ellipse due à la modalité de l'oral qui se dispense de réaliser le sujet contraint et référentiellement vide de certains verbes impersonnels ; (b) les impératifs ; (c) le cas où l'expression du sujet d'un verbe n'est pas obligatoire s'il est coréférentiel avec un constituant précédent ; (d) le cas de l'ellipse discursive. Nous avons décidé de regrouper les cas (c) et (d) dans ce que nous avons appelé ellipse du sujet. Le cas (a), qui a été observés dans les productions orales, a été pris en considération. Néanmoins, nous avons différencié les ellipses du sujet telles que (c) et (d) et celles du type de (a), qui ne relèvent du même phénomène. En revanche, les sujets implicites des impératifs (cas (b)) n'ont pas été codés.

Notons que nous avons décidé de considérer la forme *l'on* comme un allomorphe de *on* apparaissant comme étant sa variante soutenue. Pour le *y* lyonnais, qui est utilisé pour les compléments d'objets directs inanimés par opposition au pronom *le* utilisé pour les compléments d'objets directs animés, nous l'avons codé de la même manière que *le*. Néanmoins, notre codage attribué aux SN, permet de retrouver ces deux formes, *l'on* et *y*, aisément, si nous souhaitons l'analyser.

6.2.3. Le statut référentiel de l'unité

Dans un troisième temps, nous avons donné des informations sur le statut référentiel des SNL et des pronoms.

Comme pour beaucoup d'études, nous avons décidé de regrouper sous la catégorie nouvelle information tous référents non mentionnés explicitement précédemment (Allen *et al.*, 2008). En effet, aussi clairs que soient les différents statuts d'accessibilité des référents (que ce soit ceux introduits par Levelt, 1989 ou ceux par Chafe, 1976, 1994)⁷⁵, la concrétisation de ces statuts intermédiaires semble, dans beaucoup de cas, très difficile.

Si la distinction entre ancienne et nouvelle information est évidente, il n'en est pas de même pour la différenciation entre nouvelle information inaccessible *versus* accessible (Prince, 1992). Dans les cas prototypiques, dans lesquels les cadres sont bien définis, la distinction est aisée. Ainsi, dans la clause *dans la classe, le professeur faisait une dictée, le professeur* est une information semi activée par le mot *classe* précédemment énoncé. En effet, dans le cadre prototypique de la classe, le professeur est un référent fortement activé. Néanmoins, nous pouvons nous heurter à certains problèmes pratiques lorsque nous sommes face à des cadres ou des éléments non prototypiques. Ainsi, dans la clause *le professeur passait entre les rangs*, quel est le statut informationnel de *rangs* ? Est-ce un élément réellement activé par le mot *professeur* énoncé précédemment, faisant donc prototypiquement parti du cadre scolaire ou s'agit-il d'une nouvelle information ? Il s'avère que dans beaucoup de cas la question *est-ce une nouvelle information ou une information accessible ?* intervient.

N'ayant pu trouver de base stable, nous décidons de travailler avec les concepts de nouvelle information et d'ancienne information. Pour différencier ces deux statuts informationnels, nous pouvons penser nous baser sur la distinction déterminant défini/indéfini. Un SNL introduit par un déterminant défini encoderait à une information accessible alors qu'un SNL

⁷⁵ Se reporter à la Section 1.2 du Chapitre 1 (p. 32) pour une définition des différents statuts référentiels.

possédant un déterminant indéfini serait un nouveau référent non accessible. Or lors du test de ce critère sur les productions, ce dernier s'est avéré non fiable. C'est dans cette mesure que, comme pour beaucoup de recherches sur le statut informationnel (Allen *et al.*, 2008), nous avons considéré comme *nouvelle information* tous les référents mentionnés explicitement pour la première fois. Sous l'expression *ancienne information*, nous intégrons les pronoms et les SNL représentant une information déjà introduite explicitement dans le texte.⁷⁶

Pour le maintien d'une information sous forme lexicale (substitution lexicale), il a été distingué : (a) les anciennes informations reprenant un référent déjà mentionné sous la forme exacte de la précédente mention c'est-à-dire sous forme de répétition ; et (b) des anciennes informations reprenant un référent sous une forme différente de sa précédente mention. Lorsque nous avons le cas (b), nous avons différencié les substitutions impliquant des modifications portant sur le déterminant du SNL, sur les modificateurs des SNL ou sur le nom du SNL.

Ainsi, un SNL peut avoir le statut de :

a. Nouveau référent, soit une information non connue de l'interlocuteur, il s'agit de la première mention explicite d'un référent (80),

(80) *Un jour, par un bel après-midi de juin, une bande de filles est venue me « parler ».* (3^{ème} n°07, narratif écrit, E/O)

b. un ancien référent ou une réintroduction, soit une information déjà connue de l'interlocuteur donc précédemment, de façon immédiate, mentionnée dans le texte, avec répétition exacte du SNL (81),

(81) *Mon frère m'a conseillé d'acheter le portable d'une fille. Mon frère est venu [...]* (3^{ème} n°01, narratif écrit, O/E)

c. un ancien référent ou une réintroduction avec une modification du déterminant (82),

(82) *Une fille s'était mise avec moi dans une chambre avec d'autres filles aussi, cette fille s'énervait contre moi sans arrêt pour rien.* (CM2 n°06, narratif écrit, E/O)

d. un ancien référent ou une réintroduction avec une modification lexicale (83),

(83) *Donc je suis allé voir ce garçon, l'enfant disait [...]* (5^{ème} n°08, narratif écrit, O/E)

⁷⁶ Dans le codage des données, nous avons distingué ancienne et réintroduction d'information. L'ancienne information renvoie à la mention d'un référent introduit précédemment (de manière assez immédiate). En revanche, la réintroduction correspond à la mention d'un référent introduit certes précédemment dans le texte mais beaucoup plus loin de cette nouvelle mention (critère de codage : si entre deux mentions d'un même référent, il y a plus de deux UT simples ou plus d'une UT complexe (soit plus d'une clause), il s'agit d'une réintroduction). Pour ce travail de thèse, cette différenciation n'est pas exploitée.

e. un ancien référent ou une réintroduction avec une modification de modifieurs (84),

(84) *Tout le monde attendait **cette fête** car c'est **une fête avec plein jeux organisés comme le foot**. (3^{ème} n°46, narratif écrit, E/O)*

Un pronom peut avoir deux valeurs :

a. Ancien référent (85),

(85) *Quand **ma sœur** avait trois ans, / un jour en sortant de son bain, / **elle** glissa sur le placard // et **elle** s'était ouvert la lèvre.// (5^{ème} n°14, narratif écrit, E/O)*

b. Réintroduction d'un référent déjà énoncé (86),

(86) ***Cette fille** s'énervait contre sans arrêt pour rien je disais quelque chose ou je fais quelque chose (...) et **elle** s'énervait. (CM2 n°06, narratif écrit, E/O)*

Pour finir, nous avons ajouté une catégorie pour les utilisations non conventionnelles des pronoms (*pbl*). Nous avons tenu à distinguer les pronoms anaphoriques : (a) utilisés sans SNL introduit précédemment (87) ; (b) pouvant se référer à plus d'un référent préalablement intégré (88) ; ou (c) utilisés sans respect des accords en genre et/ou en nombre avec le référent auquel il se réfère (89).

(87) *Je **lui** dis qu'on peut s'arrêter pour parler mais **lui il** ne veut pas. (CM2 n°15, narratif écrit, E/O)*

(88) *la cantine y avait un garçon là / et il jouait au foot et tout / et y avait un autre garçon dans la même équipe / **pa(r)ce que il** a joué personnel / **il** a commencé à l'insulter. (CM2 n°43, narratif oral, O/E)*

(89) *Quand **une personne** a par exemple perdu un membre important de sa famille et ça **le** rend triste. (CM2 n°19, expositif écrit, E/O)*

Dans (87), le pronom *lui* apparaît une première fois sans que le référent ne soit mentionné au préalable par une forme linguistique indiquant son bas degré d'accessibilité telle que le SNL. Dans (88), l'enfant utilise aisément le pronom de la troisième personne du singulier *il jouait au foot et tout*. Après avoir introduit un nouveau référent masculin, *et y avait un autre garçon*, il utilise à nouveau le pronom *il*. Néanmoins, lors de ces nouvelles utilisations (*pa(r)ce que il a joué personnel / il a commencé à l'insulter*), il est difficile de savoir avec certitude à quel référent renvoie ce pronom (*le garçon* ou *l'autre garçon*). Dans (89), le SNL, *une personne*, est repris par la suite par une forme traduisant le haut degré d'accessibilité du référent, mais le locuteur ne respecte pas le genre du SNL.

Dans un quatrième temps, nous nous sommes demandée si la forme reprenait un référent préalablement introduit en position sujet, objet, autre ou si elle reprenait un énoncé. Étant donné la thématique de ce travail de thèse, ces informations ne sont cependant pas exploitées ici.

6.2.4. La forme verbale de la clause

Dans un cinquième temps, nous avons donné des indications sur l'utilisation du verbe de la clause, nous avons spécifié si l'individu utilisait un verbe avec un fonctionnement transitif (90), intransitif ou copules (91).

(90) *Quand j'ai vu une dame qui **achetait** des babioles sur la chanteuse, comme des tee+shirts, des affiches. (5^{ème} n°47, narratif écrit, E/O)*

(91) *Dans un collège les sujets tels que la triche, les mises à l'écart, envers certaines personnes, les bastons **sont** plutôt fréquents. (3^{ème} n°06, expositif écrit, O/E)*

6.3. Schéma du codage

Le document 1 récapitule les informations données pour chaque forme pronominale et lexicale nous intéressant pour ce travail de thèse.

Document 1 : Récapitulation du codage des syntagmes nominaux

Pronom/SNL	« Pour une forme donnée »	<i>pro/lex</i>
<hr/>		
1. Sujet – Objet – Autre <i>Suj obj aut</i>		
2. Pour les SNL Nombre de nœuds lexicaux		<i>noe1-2, etc.</i>
3. Pour les SNL et pronoms anaphoriques		
a. Nouvelle information		<i>nou</i>
b. Ancienne information (répétition pour les SNL)		<i>anc</i>
c. Ancienne information avec modification lexicale (pour les SNL)		<i>ancml</i>
d. Ancienne information avec modification du déterminant (pour les SNL)		<i>ancmd</i>
e. Ancienne information avec modification des modifieurs (pour les SNL)		<i>ancmm</i>
f. Réintroduction d'une information (pour les SNL)		<i>rei/-ml-mm-md</i>
<hr/>		
5. Verbe transitif/intransitif/copule		<i>vtran/vintr/copul</i>

Les exemples ci-dessous illustrent l'application du système de codage sous *CLAN* à deux textes d'un même enfant de CM2. Le premier exemple est une production narrative orale transcrite et codée de l'individu n°04 de CM2, le second exemple est la version orale de ce même individu.

(92) @BG:TU1

*SBJ: Chérifa@cab et moi on était de bonne@err [: bonnes] copine@err [:copines][^].

%ref:\$Chérifa|lex:sujdis:mod1:dif1:noe1:nou:ras:copul:être|:\$

\$on|pro:sujtra:cli:per:spe:ras:ras: copul:être|:\$

\$copine|lex:obj:mod1:dif1:noe1:nou:ras:copul:être|:\$

@EG:TU1

@BG:TU2

*SBJ:un jour on sais@err [: s'est] disputé.

%ref:\$jour|spt:ras:ras:ras:ras:ras:ras:vintr:se+disputer|:\$

\$on|pro:suj:cli:per:spe:anc:rnsuj:vintr:se+disputer|:\$

*SBJ: parce@err [: parce+que] j' avais un caractère de cochon [^].

%ref:\$j'|pro:suj:cli:per:spe:dei:ras:vtran:avoir|:\$

\$caractère|lex:obj:mod1:dif1:noe2:nou:ras:vtran:avoir|:\$

@EG: TU2

@End

(CM2 n°04, narratif écrit, E/O)

(93) @BG: TU1

*SBJ: #1457 ben@fp j'étais avec une copine.

%ref: \$j'|pro:suj:cli:per:spe:dei:ras:copul:être|:\$

\$copine|lex:aut:mod0:dif0:noel:nou:ras:copul:être|:\$

@EG: TU1

@BG: TU2

*SBJ: et euh@fp #593 on s'est engueulé.

%ref: \$on|pro:suj:cli:indf:ras:ras:ras:vintr:s'+engueuler|:\$

*SBJ: pa(r)ce+que j'avais <un carac@iw> [///] <un:> [/] #799 un méchant caractère.

%ref: \$j'|pro:suj:cli:per:spe:dei:ras:vtran:avoir|:\$

\$caractère|lex:obj:mod1:dif1:noel:nou:ras:vtran:avoir|:\$

@EG: TU2

@End

(CM2 n°04, narratif oral, E/O)

Cette deuxième partie (Chapitre 4 à Chapitre 6) était donc consacrée à plusieurs aspects méthodologiques, de la collecte des données au codage des productions en passant par les logiciels utilisés. Ces points présentés (Chapitre 4 à Chapitre 6) et les fondements théoriques posés (Chapitre 1 à Chapitre 3), nous pouvons désormais nous pencher sur notre partie résultats (Chapitre 7 à Chapitre 12).

Partie III

Résultats

Les Résultats

Informations préliminaires

Différentes analyses sont présentées et discutées dans cette partie. Dans le Chapitre 7, sont présentés des résultats concernant les indicateurs généraux sur la longueur des textes, à savoir *le nombre moyen de mots* (Section 7.1, p. 163), *de clauses* (Section 7.2, p. 164) et *d'UT* (Section 7.3, p. 165) par texte ainsi que *le nombre moyen de mots par clause* (Section 7.4, p. 167) et *de clauses par UT* (Section 7.5, p. 167).

Nous commençons à nous intéresser à *la proportion des SNL et des pronoms* dans le Chapitre 8 ; nous observons ces proportions d'une manière générale et ensuite selon la position syntaxique des SNL et des pronoms. Ces premières analyses nous donnent un indice de la lexicalité des textes.

Partant du postulat que la production de SNL est un indice de maturité, le degré de lexicalité des SNL est étudié dans le Chapitre 9. Nous nous concentrons alors sur le nombre de nœuds lexicaux les composants. Deux VD sont étudiées : les *SNL constitués d'un nœud lexical* et *ceux constitués de plus d'un nœud*. Ces SNL sont ensuite observés selon leur position syntaxique.

Puis, le statut informationnel des SN est étudié dans le Chapitre 10, ce qui permet de voir ce qu'il en est de la cohésion référentielle. Il s'agit de l'observation de *la proportion des SN introduisant une nouvelle information* et de celle *des SN maintenant une information*.

Cette partie Résultats se poursuit avec le Chapitre 11 concernant *les utilisations non standard des pronoms* (*réintroduction pronominale, référent manquant, référents multiples, et problème d'accord*), qui peuvent révéler si les enfants savent utiliser à bon escient les pronoms et indirectement les SNL.

Partant du postulat que les enfants ont une faible proportion *d'utilisation non conventionnelle des pronoms*, cela peut signifier que les individus savent que dans les cas de *réintroduction pronominale*, *réfèrent manquant*, *référents multiples*, ils doivent utiliser des SNL. Nous décidons donc de nous intéresser à la forme utilisée pour maintenir l'information, puisque celle pour introduire l'information est acquise ; d'autant plus que l'encodage de l'ancienne information peut se réaliser à la fois sous forme lexicale et pronominale.

C'est dans cette mesure que nous nous concentrons dans le Chapitre 12 sur la forme encodant l'ancienne information : deux VD, celle de *l'ancienne information sous forme pronominale* et celle de *l'ancienne information sous forme lexicale*, sont examinées. Nous nous focalisons ensuite sur les SNL qui maintiennent une information en prenant en considération leurs constitutions linguistiques. Le maintien de l'information sous forme lexicale est alors observé ; deux maintiens sont distingués : *celui n'impliquant pas de changement linguistique avec la mention précédente* et *celui impliquant des changements linguistiques avec la mention précédent*.

Enfin, l'ensemble des résultats est repris dans la conclusion générale de ce travail de recherche qui se termine par une proposition d'éventuelles perspectives relative à cette recherche.

Dans chacun de ces chapitres, le même type de test statistique est employé. Une analyse de variance (ANOVA) est effectuée avec : (a) deux facteurs inter-individuels, le NIVEAU SCOLAIRE (trois niveaux : CM2, 5^{ème} et 3^{ème})⁷⁷ et l'ORDRE DE PRODUCTION (deux ordres : oral/écrit et écrit/oral) ; et (b) deux facteurs intra-individuels, le TYPE DE TEXTE (deux types : expositif et narratif) et la MODALITÉ (deux modalités : orale et écrite). Dans notre étude, quatre variables sont utilisées : (a) le NIVEAU SCOLAIRE, qui est une VI dite invoquée ; (b) la MODALITÉ DE PRODUCTION ; (c) le TYPE DE TEXTE ; (d) l'ORDRE DE PRODUCTION. Ces trois dernières VI sont des variables dites provoquées.⁷⁸

Concernant l'ORDRE DE PRODUCTION, il a été indiqué dans la partie Méthodologie, que notre protocole avait prévu quatre ordres de passation : A, B, C et D. Ainsi, par exemple, un individu appartenant à l'ordre de passation B, lors de la première semaine, regarde la vidéo puis produit un texte narratif écrit. Ensuite, un expérimentateur lui fait remplir un questionnaire et le sujet produit un texte narratif oral. Une semaine après, il revient et produit

⁷⁷ CM2, âge moyen = 10;9 ; 5^{ème}, âge moyen = 12;7 ; 3^{ème}, âge moyen = 15;2.

⁷⁸ Nous avons également des variables secondaires telles que le sexe des individus (masculin ou féminin) ou le niveau d'expression langagière (avec ou sans difficulté langagière), qui n'ont pas été testées.

un texte expositif écrit, remplit un autre questionnaire et fait un texte expositif écrit. Lors de nos études pilotes, nous nous sommes aperçue que le fait de produire un type de texte avant l'autre n'engendre pas de différences significatives, alors que la modalité, oui. De ce fait, pour les analyses présentées dans ce travail de thèse, nous avons pris la décision de cumuler les ordres A et C ainsi que les ordres B et D. Nous avons alors deux types de groupes : les individus ayant débuté par une production écrite et ceux ayant débuté par une production orale. Nous parlons alors de l'ordre écrit/oral et de l'ordre oral/écrit.

Dans le but de compléter ces analyses de variance à quatre facteurs, NIVEAU SCOLAIRE X ORDRE DE PRODUCTION X TYPE DE TEXTE X MODALITÉ, des tests PLSD de Fisher, comme test post-hoc, ont été effectués, notamment pour le NIVEAU SCOLAIRE.

Afin d'illustrer les résultats significatifs, nous présentons des graphiques dont les barres d'erreur représentent l'erreur standard et des tableaux présentant les valeurs des moyennes (\bar{x}), l'amplitude (A) et l'écart-type (σ). Pour chaque analyse, seuls les résultats et les interactions significatifs sont présentés. Enfin, quand nous parlons des positions syntaxiques, nous entendons les positions sujet (S), objet (O) et autre (Au). Pour finir, les échelles des graphiques varient entre les VD, qui ne sont pas comparées entre elles.

Des analyses ont été faites pour les différents types de sujets (position sujet de verbes transitifs, de verbes intransitifs et de verbes copules). Néanmoins, les analyses de variance sur les différents types de sujet n'ont apporté que peu d'information. C'est dans cette mesure que nous avons décidé de ne pas les présenter.

Enfin, sans entrer dans le débat langue et classe sociale,⁷⁹ nous tentons tout de même de présenter succinctement, dans les discussions, des résultats concernant des populations plus favorisées (classe moyenne supérieure). Ces études sont faites sur population contenant un groupe d'adultes. Ainsi, même s'il s'agit d'adultes de classe socio-professionnelle supérieure, nous pouvons voir comment se situent les enfants de notre étude (de CSP défavorisée), par rapport aux adultes de ces études antérieures.

Pour finir, des exemples issus des productions des participants de notre étude sont régulièrement utilisés pour illustrer nos résultats. Ces exemples, qui sont rapportés tels que les individus les ont formulés, sont souvent mis en confrontation. Ainsi, par exemple, si nous souhaitons marquer la différence entre la modalité orale et la modalité écrite, nous choisissons

⁷⁹ Cf. Introduction générale.

un extrait d'un texte écrit et un autre d'un texte oral d'un même individu et ayant un même sémantisme. Ceci dans le but de montrer, que : (a) pour un même individu, la MODALITÉ est un facteur de variation ; et (b) pour un même sémantisme, un empaquetage différent de l'information est utilisé selon le contexte. Pour la plupart des exemples, nous suivons cette démarche.

Chapitre 7

Indicateurs généraux, Indices de longueur de textes

Nous commençons cette partie par la présentation et l'analyse d'indicateurs généraux, à savoir : (a) *le nombre moyen de mots par texte* ; (b) *le nombre moyen de clauses par texte* ; (c) *le nombre moyen d'Unités Terminales (UT) par texte* ; (d) *le nombre moyen de mots par clause* ; et (e) *le nombre moyen de clauses par UT*. Notons que *le nombre moyen de mots par clause* et *de clauses par UT* sont les mesures reflétant le mieux la complexité d'un texte (Berman, 1998 ; Ravid, 2005 ; Viguié-Simon, 2001). Toutes ces analyses permettent d'avoir une vision globale sur la longueur des textes (94) et (95).

(94) Les gens se bagarrent entre eux pour un rien // et ensuite ça cause des problèmes // et ils ne se parlent plus. // Il y a des gens / qui se causent pas / parce qu'ils sont de différentes peaux et de différentes races. // (CM2 n°20, expositif écrit, O/E)⁸⁰

L'exemple ci-dessus est un texte d'un sujet de CM2 comportant 40 mots, six clauses et quatre UT.

(95) Rassurez-moi, // pas tout le monde n'est comme ça ?! // Je trouve dommage l'attitude de certaines personnes dans certains cas. // C'est surtout au niveau de l'injustice, / comme par exemple, aimer une personne / parce qu'elle est de la même religion que vous / ou parce qu'elle a les mêmes idées, les mêmes goûts, choix... // Mais s'il faut pour ça laisser les autres de côté, / // je dis non ! // Dans ce genre de cas, je ne sais vraiment pas comment réagir // alors je me tais // et je réfléchis / sans dire aucun mot. // Puis, j'essaie de mettre d'accord tout le monde / pour que l'auteur se son « acte » puisse réagir à son

⁸⁰ Nous utilisons / pour indiquer une frontière de clause et // pour une frontière d'UT.

comportement. / // Sinon, si je vois / que la situation n'est pas très grave, / j'essaie d'être neutre, / de ne pas trop rentrer dans la vie des autres. / // Je trouve anormal / de critiquer, / de menacer, / de voler, / de faire peur à d'autres gens / ou même d'être hypocrite avec eux. / // J'ai déjà été le sujet de discussion / qui ne jouait pas en ma faveur // et cela m'a horriblement blessée / // car j'étais seule, / sans moyen de comprendre / pourquoi les autres prenaient du plaisir à me faire souffrir. / // Mais j'ai fini par comprendre / que toute ma vie, ça sera comme ça / et que tout ce qui a pu me faire souffrir / finira enfin à m'endurcir. / // Donc pour moi, il faut savoir ignorer les gens, / // car, comme on me l'a dit souvent, / « l'ignorance tue » ! // J'en ai plus qu'assez de voir / chaque jour des gens pleurer / ou même apeurés à cause d'une dizaine de personnes / qui sont là, / à se moquer, / à critiquer // car, pour moi, les mots sont pires que les maux. / // Après dans la tête des gens, ça travaille sur / « qu'est-ce qu'ils vont me dire demain au collègue ? / Qu'est-ce que je leur ai fait ? / Pourquoi c'est moi / qu'il embête chaque jour ? » // C'est si dur de voir des gens souffrir / alors qu'on ne peut rien faire / pour empêcher ce « massacre ». / // Sans vouloir me venter, / c'est pour cela que, / dès que je vois des nouveaux élèves au collègue, / j'essaie de faire connaissance avec eux (pas tous) / et comme je connais beaucoup de personnes dans ce collègue, / et bien, tout le monde va se mettre à parler avec ces nouvelles « recrues ». / // J'espère / qu'il y a autant de personnes / qui font autant que moi / (je fais confiance à la justice (lol)) // car peut-être qu'un jour on pourra refaire un monde meilleur. // P.S : Je n'ai pas toujours été un ange moi aussi. / // Parfois, il m'est arrivé de « parler » sur les gens / que je n'aime pas. // (3^{ème} n°31, expositif écrit, O/E)

L'exemple présenté ci-dessus est un texte d'un individu de 3^{ème} comportant 436 mots, 77 clauses et 33 UT.

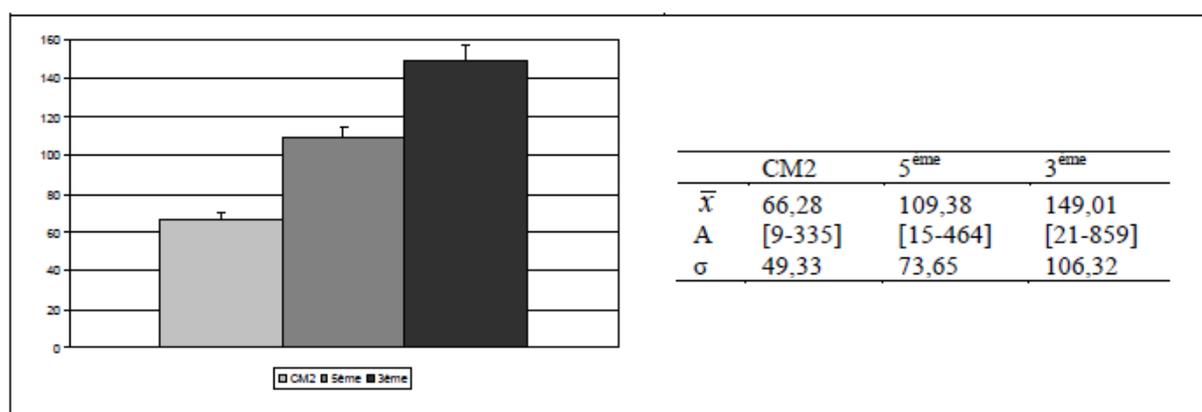
Pour le nombre de mots par texte, nous avons comptabilisé tous les mots de la production mais écarté certains types d'unités. Ainsi, les pauses pleines telles que *euh*, *ben*, *et tout*, *mhm* n'ont pas été comptabilisées. Les mots composés tels que *quelque chose* ou *belle-mère* ont été considérés comme une seule unité. Enfin, pour les phénomènes de révisions, à l'écrit, nous avons pris en considération le résultat de la révision. Ainsi, si dans un premier temps l'enfant écrivait/disait *les grosses disputes* mais revenait dessus pour écrire/dire *les disputes*, nous comptions alors deux mots et non trois. Pour le nombre de clauses par texte, nous avons également écarté certaines clauses telles que *comment je pourrais dire*.

7.1. Nombre moyen de mots par texte

Les figures qui suivent présentent *le nombre moyen de mots par texte* en fonction du NIVEAU SCOLAIRE, de la MODALITÉ et de l'ORDRE DE PRODUCTION.

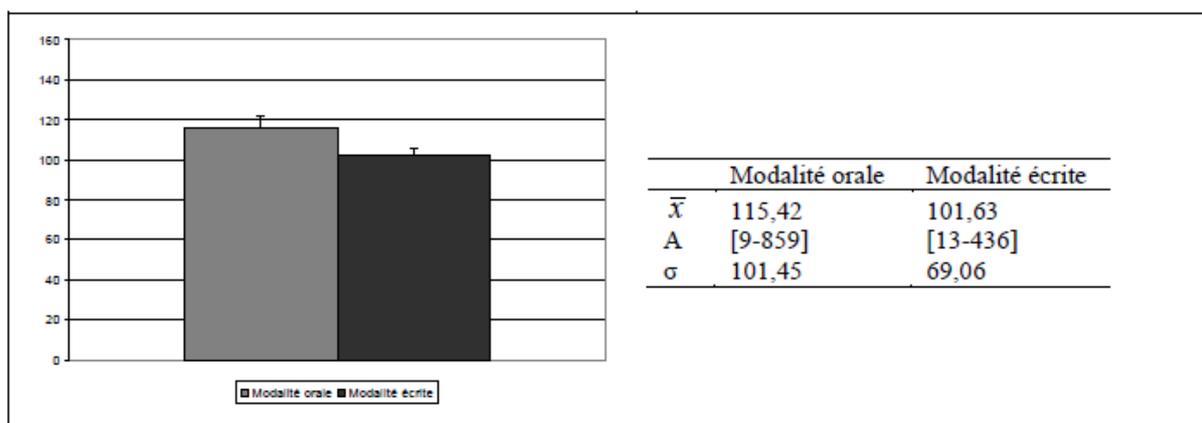
Le nombre de mots par texte augmente avec le NIVEAU SCOLAIRE ($F_{(2,504)} = 48,579$, $p < 0,0001$) ; les 3^{ème} produisent significativement plus de mots que les 5^{ème} qui en produisent plus que les CM2 (Figure 17) ($p < 0,0001$).

Figure 17 : Nombre moyen de mots par texte selon le NIVEAU SCOLAIRE



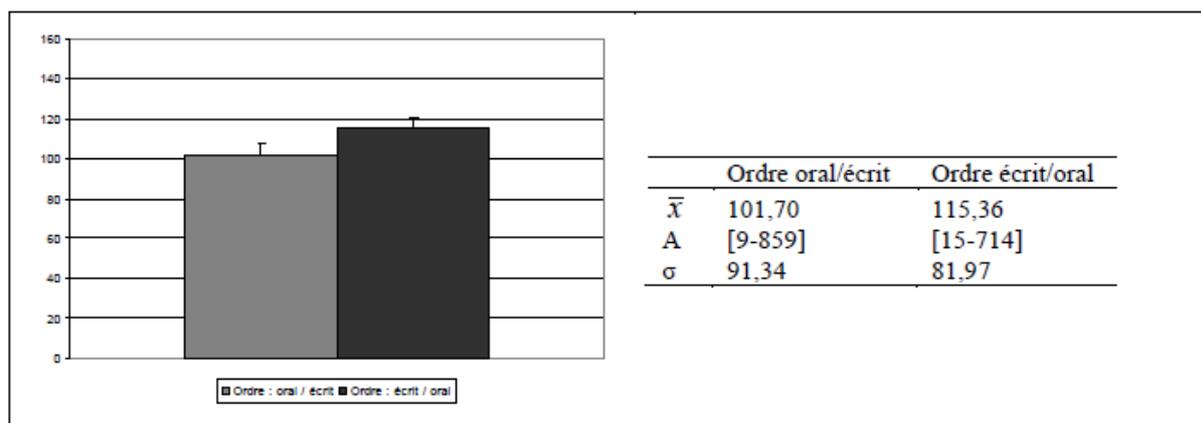
Les différences en fonction de la MODALITÉ sont significatives ($F_{(1,504)} = 4,021$, $p < 0,0455$). Les sujets produisent davantage de mots dans les textes oraux que dans les textes écrits (Figure 18).

Figure 18 : Nombre moyen de mots par texte selon la MODALITÉ



Enfin, l'ORDRE DE PRODUCTION affecte *le nombre moyen de mots par texte* qui est plus élevé dans les textes d'individus ayant produit dans l'ordre écrit/oral que dans l'ordre oral/écrit (Figure 19) ($F_{(1,504)} = 4,307, p < 0,0385$).

Figure 19 : Nombre moyen de mots par texte selon l'ORDRE DE PRODUCTION



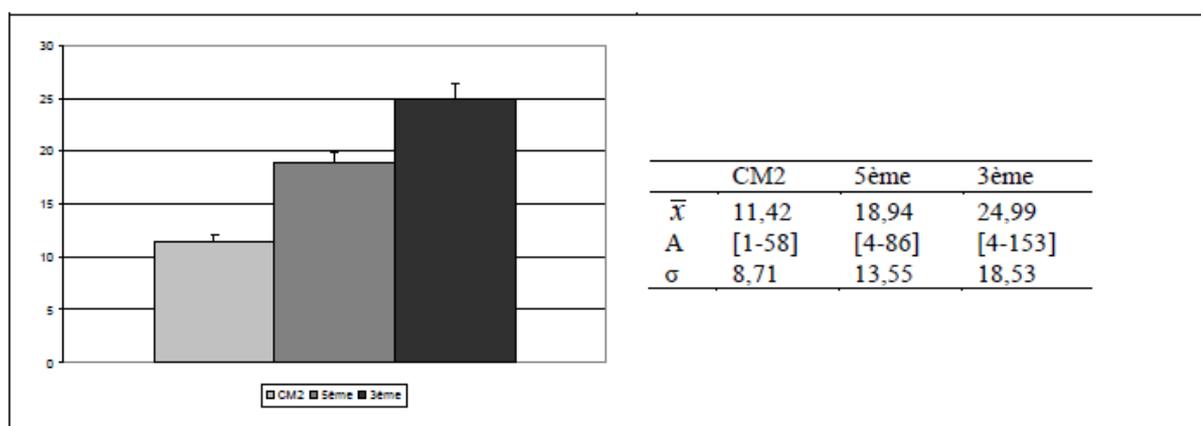
Le facteur TYPE DE TEXTE n'affecte pas significativement cette VD.

7.2. Nombre moyen de clauses par texte

Les figures 20 et 21 présentent *le nombre moyen de clauses par texte* en fonction du NIVEAU SCOLAIRE et de la MODALITÉ.

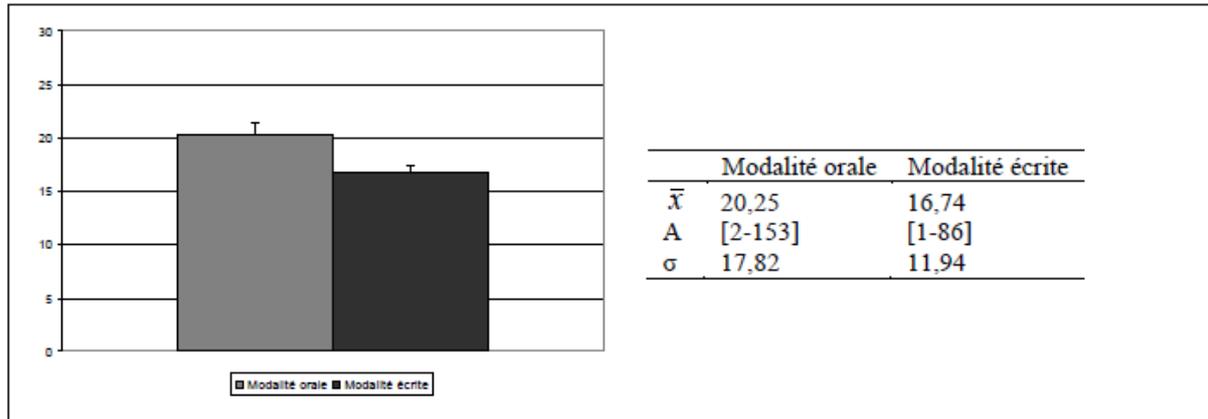
Le nombre de clauses par texte augmente avec le NIVEAU SCOLAIRE ($F_{(2,504)} = 42,372, p < 0,0001$) ; les collégiens de 3^{ème} produisent significativement plus de mots que ceux de 5^{ème} qui en produisent plus que les écoliers de CM2 (Figure 20) ($p < 0,0001$).

Figure 20 : Nombre moyen de clauses par texte selon le NIVEAU SCOLAIRE



Les différences selon la MODALITÉ sont significatives ($F_{(1,504)} = 8,346$, $p < 0,0040$) : les individus produisent davantage de clauses lors de productions orales que lors de productions écrites (Figure 21).

Figure 21 : Nombre moyen de clauses par texte selon la MODALITÉ



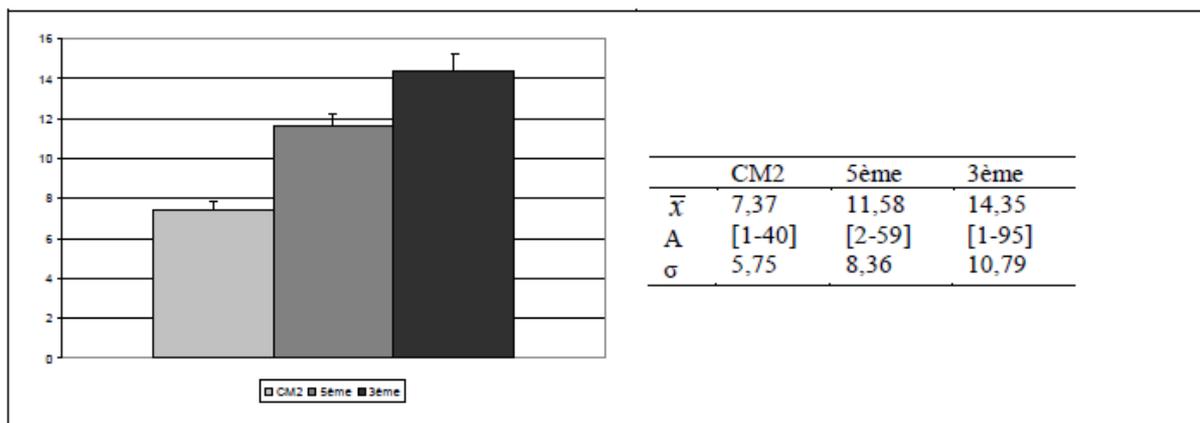
Les facteurs TYPE DE TEXTE et ORDRE n'ont pas d'effet significatif sur *le nombre moyen de clauses par texte*.

7.3. Nombre moyen d'Unité Terminale par texte

Les figures suivantes présentent *le nombre moyen d'UT par texte* en fonction du NIVEAU SCOLAIRE et de la MODALITÉ.

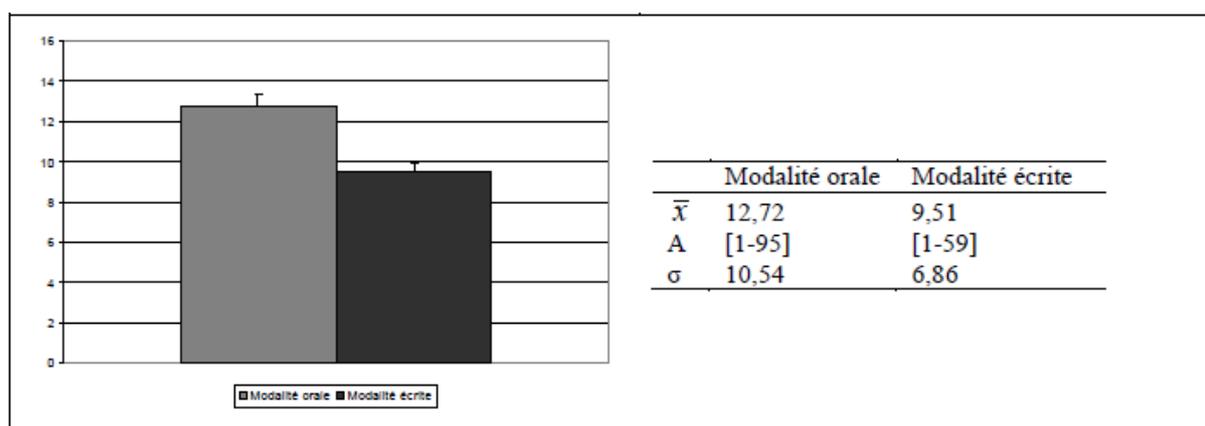
Les différences en fonction du NIVEAU SCOLAIRE sont significatives ($F_{(2,504)} = 31,541$, $p < 0,0001$). *Le nombre d'UT par texte* augmente avec le NIVEAU SCOLAIRE ; les 3^{ème} en produisent plus que les 5^{ème} ($p = 0,0019$), qui en produisent plus que les CM2 (Figure 22) ($p < 0,0001$).

Figure 22 : Nombre moyen d'Unité Terminale par texte selon le NIVEAU SCOLAIRE



Cette VD varie également selon la MODALITÉ ($F_{(1,504)} = 19,423$, $p < 0,0001$) ; en effet, les sujets produisent davantage d'UT dans les textes oraux que dans les textes écrits (Figure 23).

Figure 23 : Nombre moyen d'Unité Terminale par texte selon la MODALITÉ

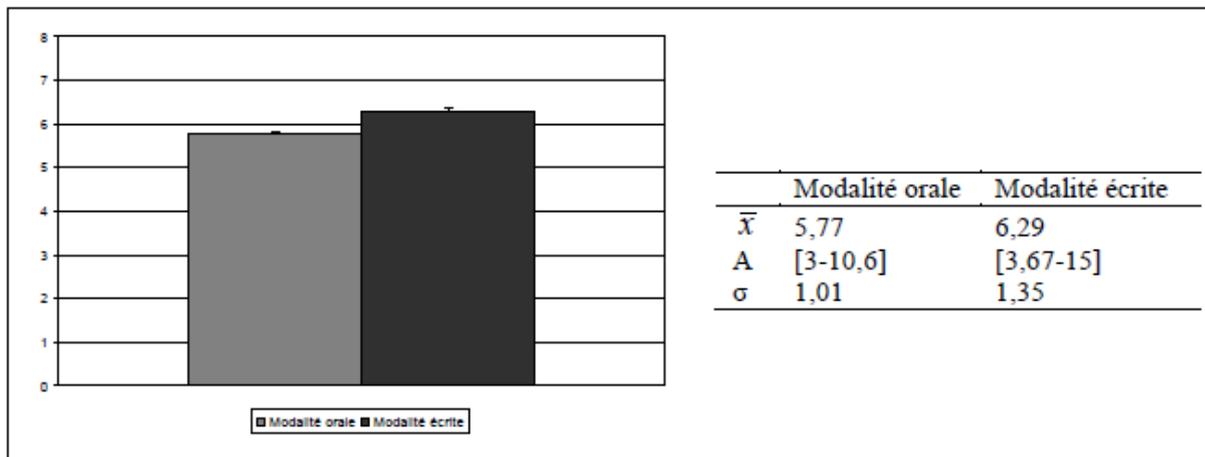


Les facteurs TYPE DE TEXTE et ORDRE n'ont pas d'effet significatif sur *le nombre moyen d'UT par texte*.

7.4. Nombre moyen de mots par clause

La Figure 24 illustre *le nombre moyen de mots par clause* en fonction de la MODALITÉ. Les différences selon la MODALITÉ sont significatives ($F_{(1,504)} = 25,725$, $p < 0,0001$) : les sujets produisent davantage de mots par clause dans les textes écrits que dans les textes oraux (Figure 24).

Figure 24 : Nombre moyen de mots par clause selon la MODALITÉ



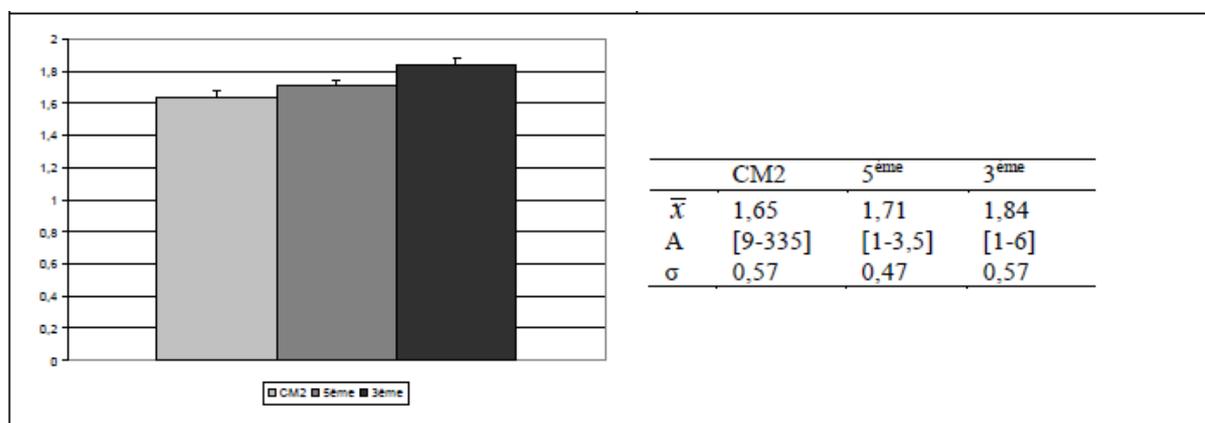
Les facteurs NIVEAU SCOLAIRE, TYPE DE TEXTE et ORDRE n'affectent pas cette VD.

7.5. Nombre moyen de clauses par Unité Terminale

Les figures qui suivent présentent *le nombre moyen de clauses par UT* en fonction du NIVEAU SCOLAIRE, du TYPE DE TEXTE et de la MODALITÉ.

Les différences en fonction du NIVEAU SCOLAIRE sont significatives ($F_{(2,504)} = 6,994$, $p < 0,0010$). Les individus de 3^{ème} produisent significativement plus de clauses dans une même UT que ceux de 5^{ème} ($p = 0,0148$) mais ces derniers en produisent pratiquement autant que les écoliers de CM2 ($p = 2325$) (Figure 25).

Figure 25 : Nombre moyen de clauses par Unité Terminale selon le NIVEAU SCOLAIRE



Le TYPE DE TEXTE ($F_{(1,504)} = 62,928, p < 0,0001$) et la MODALITÉ ($F_{(1,504)} = 12,818, p < 0,0004$) affectent également le nombre moyen de clauses par UT. Les individus produisent davantage de clauses dans une même UT dans les textes expositifs et dans les textes écrits (Figure 26 et Figure 27).

Figure 26 : Nombre moyen de clauses par Unité Terminale selon le TYPE DE TEXTE

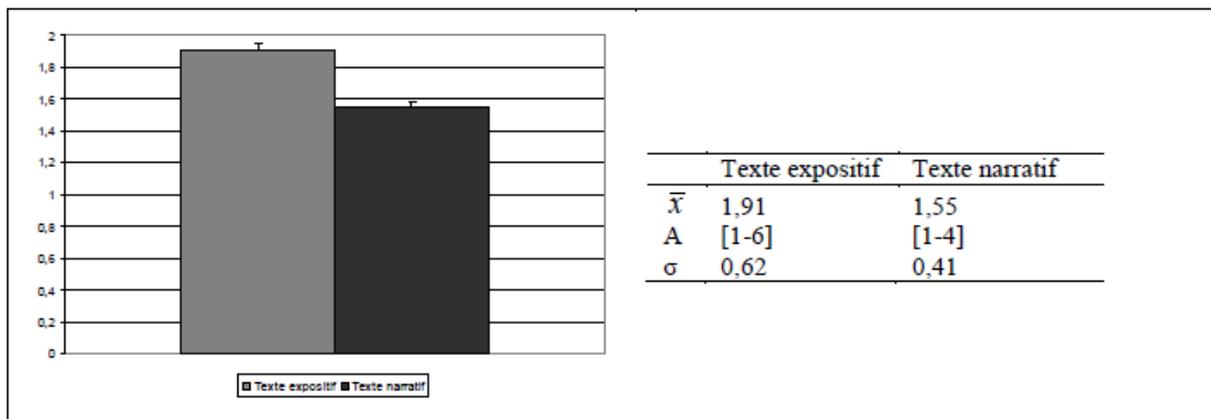
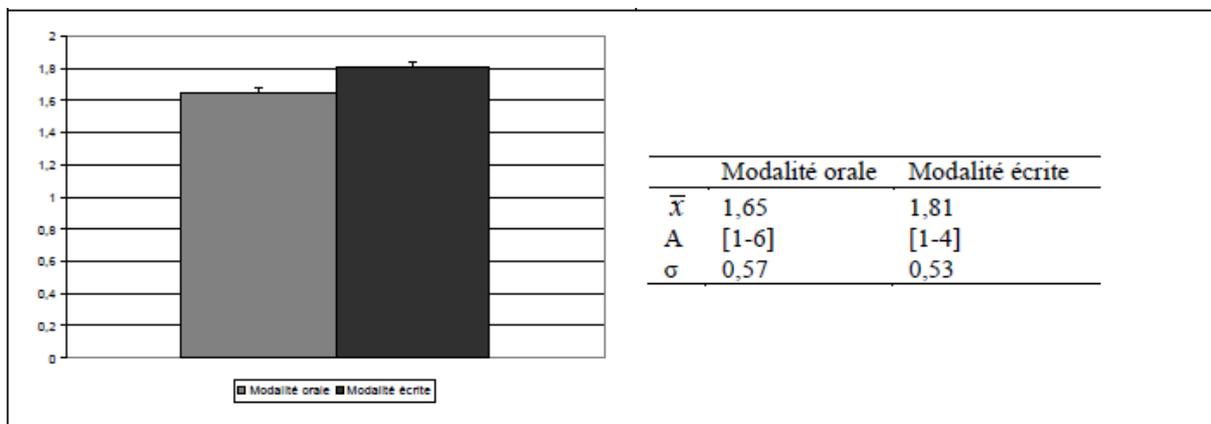


Figure 27 : Nombre moyen de clauses par Unité Terminale selon la MODALITÉ



La variable ORDRE n'a pas d'effet significatif sur le nombre moyen de clauses par UT.

7.6. Discussion du Chapitre 7

Le but de ce chapitre était de voir si *le nombre moyen de mots, de clauses et d'UT par texte* ainsi que *le nombre moyen de mots par clause et celui de clauses par UT* variaient selon les facteurs : NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE.⁸¹

Nous nous attendions à ce que ces variables, d'une manière générale, augmentent avec le NIVEAU SCOLAIRE dans la mesure où les enfants, avec l'expérience de l'école, produiraient des textes de plus en plus longs. L'analyse des données confirme ces attentes. *Le nombre de mots, de clauses, d'Unités Terminales par texte* ainsi que *le nombre moyen de clauses par UT* augmentent significativement non seulement entre les enfants de CM2 et les collégiens de 5^{ème} mais aussi entre les individus de 5^{ème} et ceux de 3^{ème}. Seul *le nombre moyen de mots par clause* ne varie pas selon le NIVEAU SCOLAIRE.

Nous nous attendions également à un effet du contexte de production (TYPE DE TEXTE et MODALITÉ) sur certaines de ces VD. Ainsi, les deux mesures – *nombre moyen de mots par clause et de clauses par UT* – seraient plus élevées dans les textes expositifs et écrits que dans les textes narratifs et oraux. Néanmoins, *les nombres moyens de mots, de clauses et d'UT par texte* seraient sûrement plus importants dans les textes narratifs et oraux que dans les textes expositifs et écrits. En effet, rappelons que *le nombre moyen de mots par clause et de clauses par UT* sont les mesures reflétant la complexité d'un texte (Berman, 1998 ; Ravid, 2005 ; Viguié-Simon, 2001).

Les résultats révèlent que le TYPE DE TEXTE affecte seulement *le nombre moyen de clauses par UT*, qui est plus important dans les textes expositifs que dans les textes narratifs. *Le nombre moyen de clauses par UT* est une des deux mesures reflétant au mieux la complexité, la seconde étant *le nombre moyen de mots par clause*. Ce résultat est compréhensif. En effet, le type de texte expositif implique une pression importante ce qui a pour effet de produire des UT possédant plus de clauses. Les deux exemples représentatifs sont deux extraits de productions écrites, une expositive (96) et une narrative (97), d'un même individu.

(96) *Il y a des personnes / qui ignorent d'autres personnes / parce qu'ils ne les apprécient pas / ou Ø trouvent / que leur comportement n'est pas normal de par leurs tenues vestimentaires, leurs habitudes, ou leurs goûts / ou peut être parce qu'ils n'ont tout simplement pas envie de rester avec elles. // (3^{ème} n°23 expositif écrit, E/O)*

⁸¹ Les attentes, hypothèses et problématiques sont davantage explicitées au Chapitre 3.

(97) Mon père avait dû appeler des dizaines de fois à « SFR » / et Ø dû leur envoyer plusieurs courriers électroniques. /// (3^{ème} n°23 narratif écrit, E/O)

En exposition (96), l'individu produit une UT plus complexe (constituée de 6 clauses), impliquant des subordonnées et des coordinations avec ellipse du sujet. En revanche, en narration (97), le même individu produit une UT composée de deux clauses. Les enfants s'adaptent au contexte académique du texte expositif en produisant des UT plus complexes, impliquant des relations syntaxiques telles que les subordonnées.

La MODALITÉ DE PRODUCTION entraîne également des différences significatives. *Les nombres de mots, de clauses et d'UT par textes* sont plus importants dans les textes oraux. Ceci se comprend dans la mesure où la modalité orale impose un rythme de production rapide, ne possède pas la contrainte de la transcription que l'écrit implique. L'oral se caractérise également par des répétitions d'une même idée sous un encodage linguistique différent et une proportion importante de marqueurs de discours. En revanche, *les nombres moyens de mots par clause et de clauses par UT*, mesures reflétant plus fidèlement la complexité d'un texte (Ravid, 2005), sont plus importants dans les textes écrits. Ce résultat va dans le sens d'études précédentes (Berman et Slobin, 1994 ; Ravid, 2005). Ces comportements reflètent les circonstances de transmission de l'information et de production bien spécifiques des deux modalités du langage. Les deux exemples représentatifs ci-dessous sont deux extraits de productions expositives, une écrite (98) et une orale (99), d'un même individu de 3^{ème}.

(98) Je vais parler des rapports sociaux / qu'ont les adolescents / quand ils sont au collège./// (3^{ème} n°01, expositif écrit, O/E)

(99) Y a les rapports euh entre les personnes au collège euh classe par exemple./// (3^{ème} n°01, expositif oral, O/E)

Pour exprimer une même idée, le sujet choisit d'emballer son message différemment selon la modalité de production. À l'écrit, le locuteur/scripteur produit une UT complexe de trois clauses (98), alors qu'à l'oral (99), il produit une UT constituée d'une seule clause. La modalité écrite permet aux enfants, par le temps de planification important qu'elle offre et par l'absence de pression communicative, de produire des textes plus longs en mots et des UT impliquant plusieurs structures dépendantes d'une clause principale.

L'ORDRE DE PRODUCTION entraîne des différences significatives seulement pour *le nombre moyen de mots par texte* qui est plus élevé dans les textes ayant été produits par des individus appartenant à l'ordre de passation écrit/oral. Les exemples suivants illustrent ce phénomène.

(100) *Les histoires commencent quand des gens ne s'aiment pas, ils s'insultent et ça dégénère et ils se bagarrent. (CM2 n°11, expositif écrit, O/E)*

(101) *La tricherie pour les contrôles ou pour d'autres choses n'est jamais bien parce qu'on n'apprend rien après et qu'on pourrait toujours tricher. Pour pas que ton camarade de classe ne te dise les réponses, dis lui non ou si il veut te faire montrer les réponses, dis lui non parce que peut-être il a faux, mais il ne faut jamais tricher parce que on n'arrivera rien à la vie. (CM2 n°15, expositif écrit, E/O)*

Le texte de l'individu appartenant à l'ordre oral/écrit (100) compte 20 mots. En revanche, le texte de l'enfant appartenant à l'ordre écrit/oral (101) compte 72 mots.

Sans entrer dans le débat langue/CSP,⁸² nos résultats peuvent en partie être comparés aux analyses effectuées sur la population francophone native plus favorisée du projet Spencer (classe moyenne supérieure). Quatre groupes sont observés : des enfants de 9/10 ans (scolarisés en CM2), de 12/13 ans (scolarisés en 5^{ème}), de 15/16 ans (scolarisés en 2^{nde}) et des adultes. Viguié-Simon (2001) montre que *les nombres de clauses* et *d'UT par texte* ainsi que *le nombre de clauses par UT* augmentent également avec les années de scolarisation.⁸³ Le Tableau 8 résume ces informations selon l'âge et les met en comparaison avec notre population ReFlex défavorisée.

[Tableau 8 : Indicateurs généraux selon l'âge et la Classe Socio-Professionnelle]

Tableau 8 : Indicateurs généraux selon l'âge et la Classe Socio-Professionnelle

	Clauses par texte		UT par texte		Clauses par UT	
	Spencer favorisé	ReFlex défavorisé	Spencer favorisé	ReFlex défavorisé	Spencer favorisé	ReFlex défavorisé
	e	e	e	e	e	e
9/10 ans et CM2	11,34	11,42	6,83	7,37	1,78	1,65

⁸² Rappelons que CSP correspond à la Catégorie Socio-Professionnelle.

⁸³ *Le nombre de mots par texte* ainsi que *le nombre de mots par clause* ne sont pas testés.

12/13 ans et 5^{ème}	14,30	18,94	8,10	11,58	1,9	1,71
15/16 ans et 3^{ème}	18,21	24,99	10,03	14,35	1,88	1,84
Adultes	38,07	/	19,57	/	2,1	/

Nos participants produisent autant voire davantage de *clauses et d'UT par texte*. Pour le *nombre de clauses par UT*, les participants ReFlex produisent des UT un peu moins longues que les enfants du projet Spencer. Nos participants de 3^{ème} se rapprochent des adultes ; en sachant que nos 3^{ème} ont une année de scolarisation de moins que les 15/16 ans du projet Spencer, qui sont en classe de 2nde. Cette comparaison montre que la courbe développementale des locuteurs/scripteurs défavorisés de notre recherche n'est pas différente de celle des individus favorisés du projet Spencer.

Néanmoins, concernant *le nombre de clauses par UT*, mesure reflétant la complexité syntaxique d'un texte (Berman, 1998 ; Ravid, 2005 ; Viguié-Simon, 2001), les résultats de Viguié-Simon révèlent que la complexité des UT n'est pas « *un développement progressif [...] mais plutôt un développement par palier, [...] un développement extrêmement tardif* » (Viguié-Simon, 2001:124). Concernant notre population, nous notons des différences significatives entre les trois groupes d'enfants. Notre conclusion est alors que la complexité des UT relève certes d'une capacité maîtrisée totalement tardivement mais qu'il s'agit tout de même d'un développement progressif.

Le fait que *le nombre de clauses par UT* soit plus important dans les textes expositifs que dans les textes narratifs est également révélé par Viguié-Simon (2001). De ce résultat, Viguié-Simon (2001) confirme une de ses hypothèses de travail selon laquelle cette VD varierait selon le TYPE DE TEXTE mais pas selon la MODALITÉ ; ceci irait dans le sens des travaux antérieurs d'O'Donnell, Griffin et Norris (1967) et Loban (1976).

Concernant l'impact de la MODALITÉ, nous ne pouvons comparer *le nombre de mots, de clauses et d'UT par textes* ainsi que *le nombre moyen de mots par clause* avec la population des travaux de recherche de Viguié-Simon (2001). En effet, les tests sur ces VD ne sont pas effectués. Néanmoins, elle teste *le nombre de clauses par UT*. La chercheuse montre que la MODALITÉ DE PRODUCTION n'a pas autant d'impact sur *le nombre de clauses par UT* pour ses participants que pour les nôtres. Sa conclusion confirmerait les données d'O'Donnell *et al.* (1967) et Loban (1976) selon lesquelles *le nombre de clauses par UT* ne varierait seulement

par le type de textes. Néanmoins, Viguié-Simon (2001) nuance en disant que la MODALITÉ DE PRODUCTION interagit avec le TYPE DE TEXTE.

Les résultats de ce Chapitre 7 confirment en partie nos hypothèses : toutes les VD varient en fonction des facteurs de notre étude. Néanmoins, selon la VD, les facteurs impliquant des variations ne sont pas les mêmes. Ainsi, si pour *le nombre de mots par texte*, le NIVEAU SCOLAIRE, la MODALITÉ et l'ORDRE ont un impact, pour *le nombre moyen de clauses par texte* seuls les facteurs NIVEAU SCOLAIRE et MODALITÉ de production ont un effet significatif.

Chapitre 8

Lexicalité des productions : Proportions des syntagmes nominaux

Dans ce chapitre, les résultats relatifs à *la proportion des SNL et des pronoms* sont présentés. Nous entendons par SNL toute unité ayant comme tête un nom ; par formes pronominales, à tous types de pronoms pris en considération dans notre étude.⁸⁴ Par syntagme nominal (SN), nous entendons l'addition des SNL et des pronoms.⁸⁵

Dans un premier temps, des analyses générales sur les proportions des SNL et des pronoms sont présentées.

Dans un second temps, nous nous intéressons au nombre moyen de ces deux unités par clause (102).

(102) *J'ai acheté un nouveau téléphone portable sur internet.* (3^{ème} n°23, narratif écrit, E/O)

Dans cet exemple, nous comptabilisons deux SNL, *un nouveau téléphone portable* et *sur internet*, et un pronom *j'*.

Dans un troisième temps, des analyses selon les positions syntaxiques sont proposées. Nous observons alors *la proportion des SNL et des pronoms en position sujet, objet et autre*.

⁸⁴ Aussi bien les déictiques que les anaphoriques.

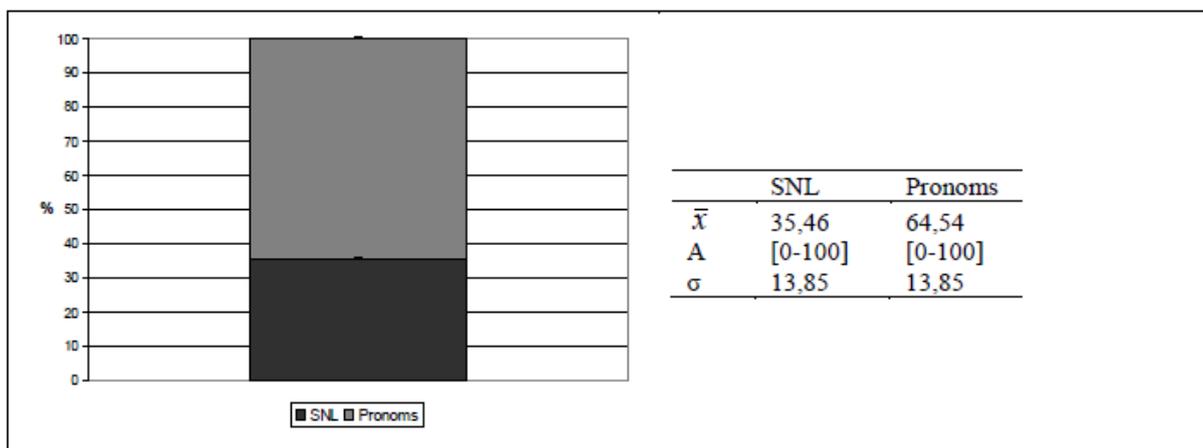
⁸⁵ Se reporter au Chapitre 6, Section 6.2 (p.143) pour le codage des SNL et des pronoms.

8.1. Analyses générales

8.1.1. Proportion des syntagmes nominaux lexicaux et des pronoms

La Figure 28 présente les moyennes de pourcentages *des pronoms* et *des SNL*. Elle montre que les individus, tous facteurs confondus (NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE), utilisent significativement plus de pronoms que de SNL dans leurs productions (Test T : $t_{(527)} = -24,183$, $p < 0,0001$).

Figure 28 : Proportions des SNL et pronoms par texte par rapport au nombre total de SN

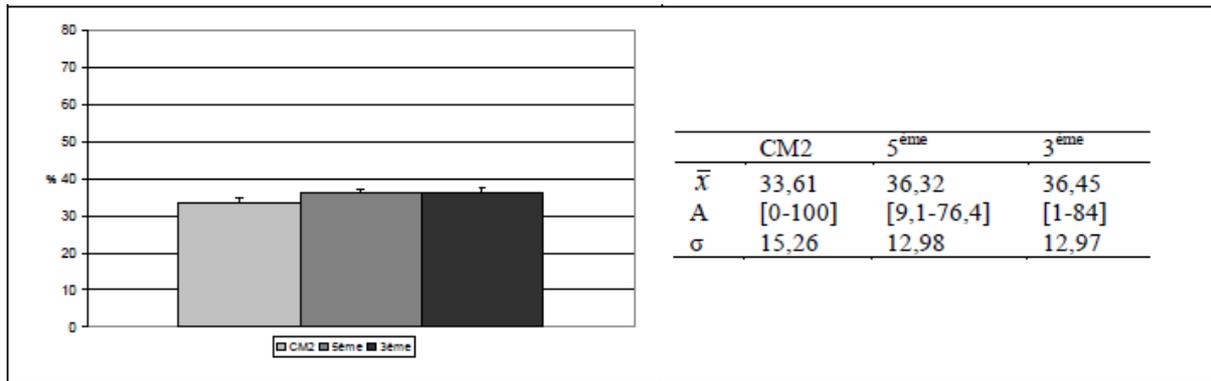


Les figures 29 à 34 présentent *les proportions des SNL et des pronoms* par rapport au nombre total de SN en fonction du NIVEAU SCOLAIRE, du TYPE DE TEXTE, de la MODALITÉ et de l'ORDRE DE PRODUCTION. Les différences selon le NIVEAU SCOLAIRE ($F_{(2,504)} = 2,844$, $p < 0,05$), le TYPE DE TEXTE ($F_{(1,504)} = 7,362$, $p < 0,0069$), la MODALITÉ ($F_{(1,504)} = 72,506$, $p < 0,0001$) et l'ORDRE ($F_{(1,504)} = 13,144$, $p < 0,0003$) sont significatives.

Les figures 29 à 31 présentent *les proportions des SNL* par rapport au nombre total de SN.

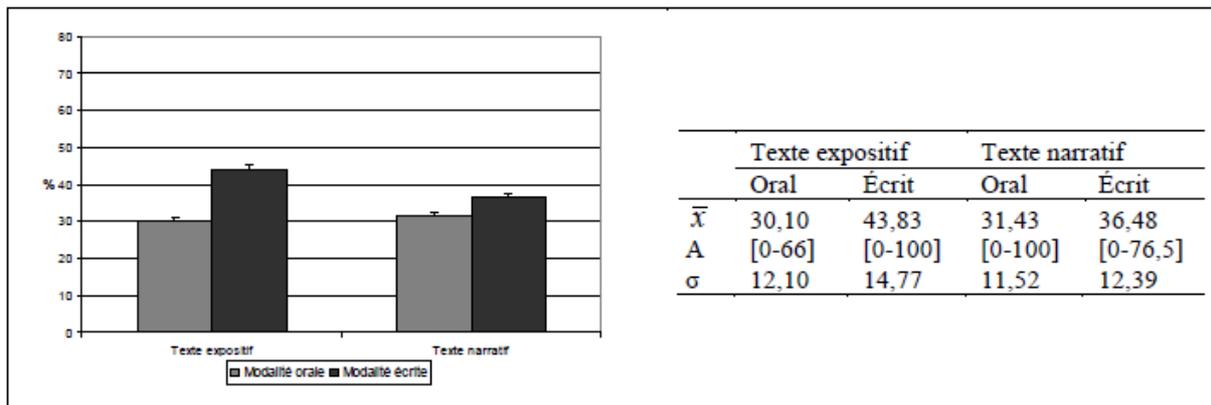
Les proportions des SNL augmente avec le NIVEAU SCOLAIRE (Figure 29). Si les écoliers de CM2 se distinguent significativement des individus de 5^{ème} ($p = 0,0461$) et de ceux de 3^{ème} ($p = 0,0345$), ces deux dernières populations produisent une proportion de SNL quasi équivalente ($p = 0,9231$).

Figure 29 : Proportions des SNL par texte selon le NIVEAU SCOLAIRE



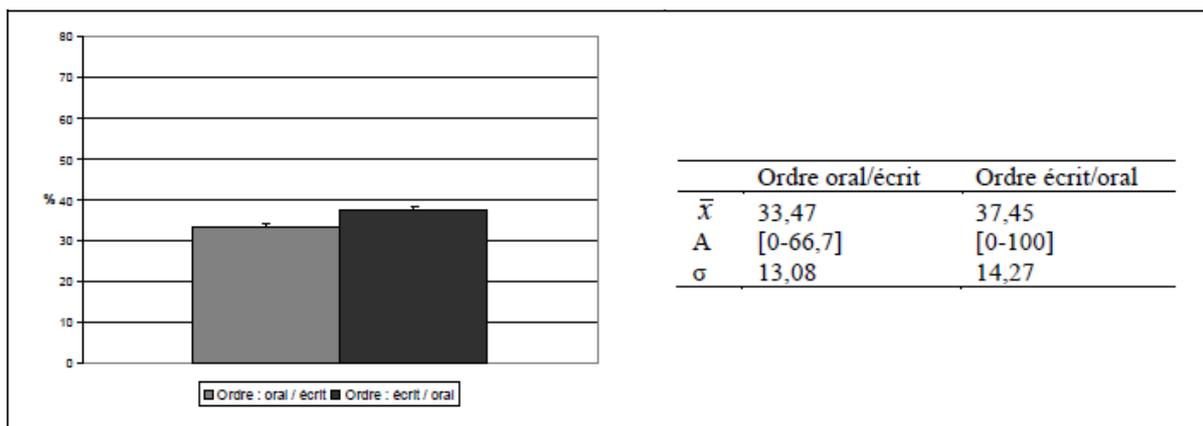
De plus, un fort taux de SNL semblent davantage caractériser les textes expositifs et les textes écrits que les productions narratives et orales (Figure 30).

Figure 30 : Proportions des SNL par texte selon le TYPE DE TEXTE et la MODALITÉ



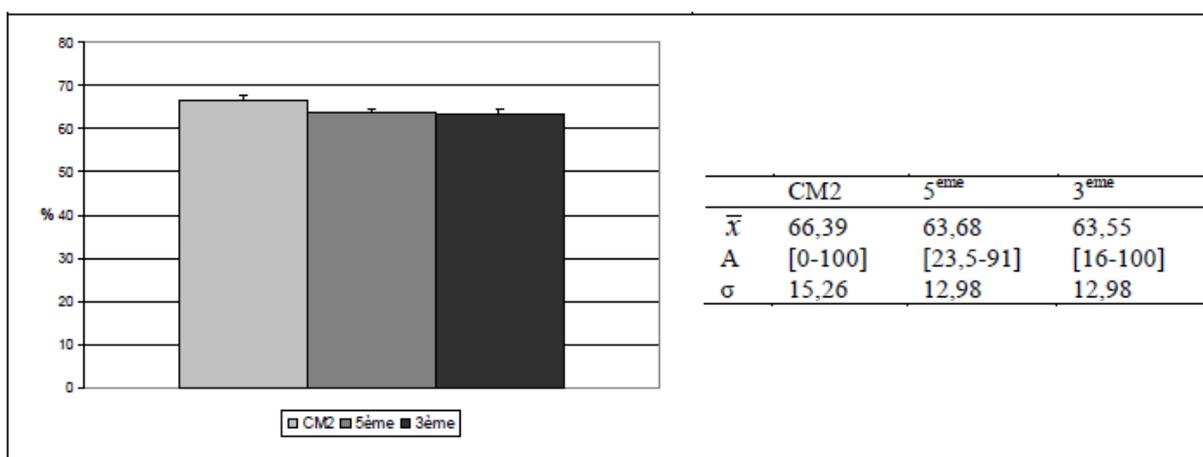
Enfin, les individus ayant produit dans l'ordre écrit/oral, produisent davantage de SNL que ceux ayant produit dans l'ordre inverse (Figure 31).

Figure 31 : Proportions des SNL par texte selon l'ORDRE DE PRODUCTION



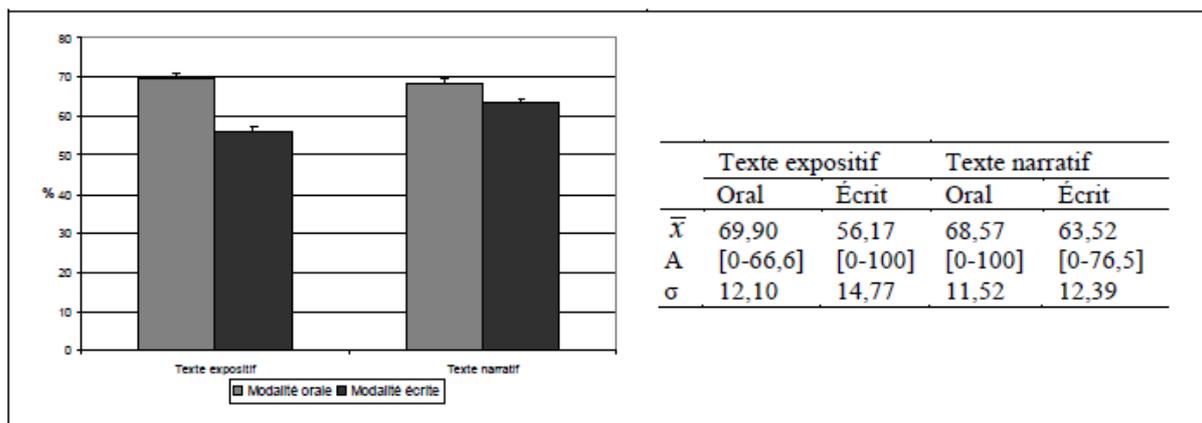
Les figures 32 à 34 présentent *les proportions des pronoms* par rapport au nombre total de SN. *La proportion des pronoms* diminue avec le NIVEAU SCOLAIRE (Figure 32). Si les individus de CM2 se distinguent significativement des collégiens de 5^{ème} ($p = 0,0461$) et de ceux de 3^{ème} ($p = 0,0345$), ces deux dernières populations produisent une proportion proche de pronoms ($p = 0,9231$).

Figure 32 : Proportions des pronoms par texte selon le NIVEAU SCOLAIRE



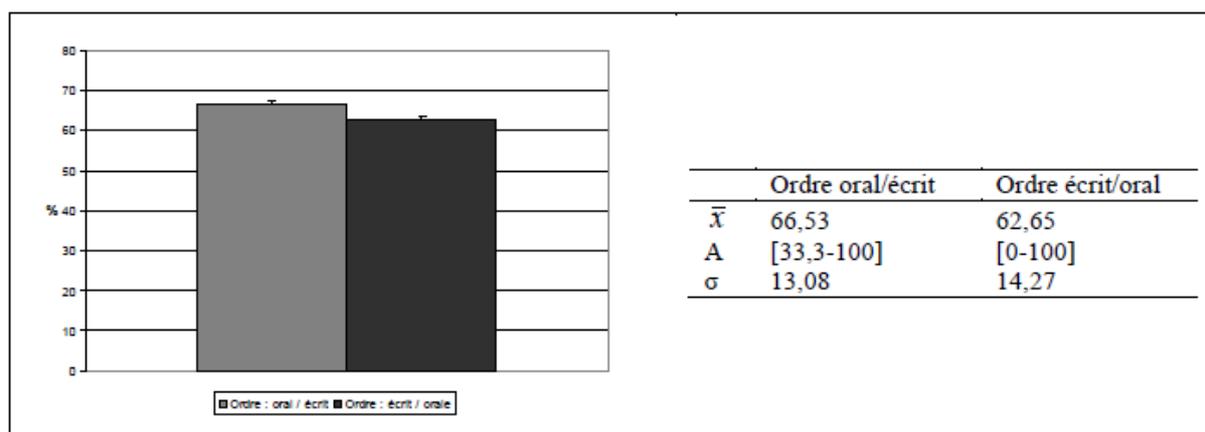
Les pronoms semblent plus caractériser les textes narratifs et oraux que les productions expositives et écrites (Figure 33).

Figure 33 : Proportions des pronoms par texte selon le TYPE DE TEXTE et la MODALITÉ



De plus, les individus produisant en premier lieu un texte oral puis en second lieu un texte écrit, ont plus recours aux pronoms que les individus commençant par un texte écrit et finissant par un texte oral (Figure 34).

Figure 34 : Proportions des pronoms par texte selon l'ORDRE DE PRODUCTION



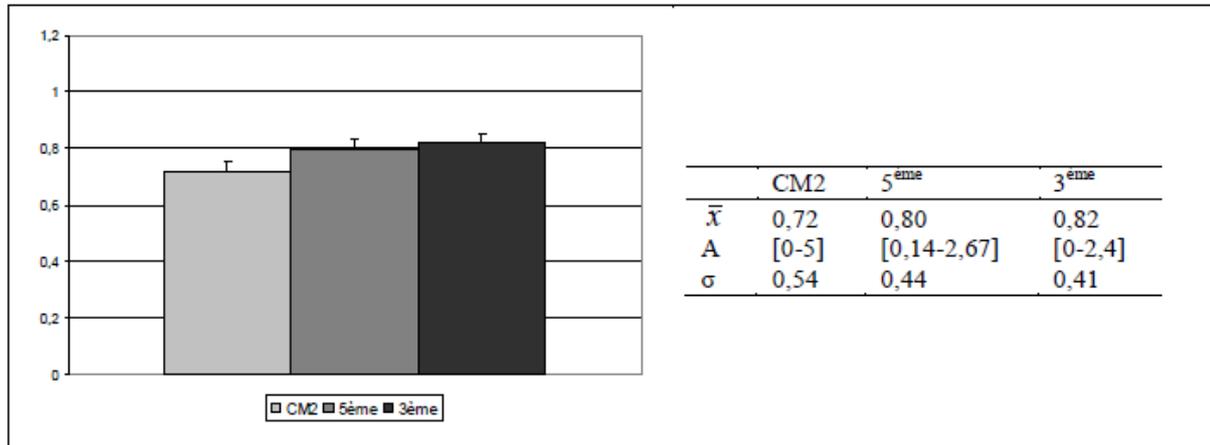
Les résultats révèlent deux interactions entre le TYPE DE TEXTE et la MODALITÉ ($F_{(1,504)} = 15,724$, $p < 0,0001$) (Figure 30 et Figure 33). En production expositive écrite, la proportion de SNL atteint son maximum. Dans les autres conditions – narratif écrit, expositif oral et narratif oral – il s'agit du nombre de pronoms qui est au maximum. De plus, notons qu'en contexte narratif, la différence entre les deux modalités est moindre, contrairement au contexte expositif.

8.1.2. Nombre moyen d'items lexicaux par clause

Les figures 35 à 37 présentent le nombre moyen d'items lexicaux par clause en fonction du NIVEAU SCOLAIRE, du TYPE DE TEXTE, de la MODALITÉ et de l'ORDRE DE PRODUCTION.

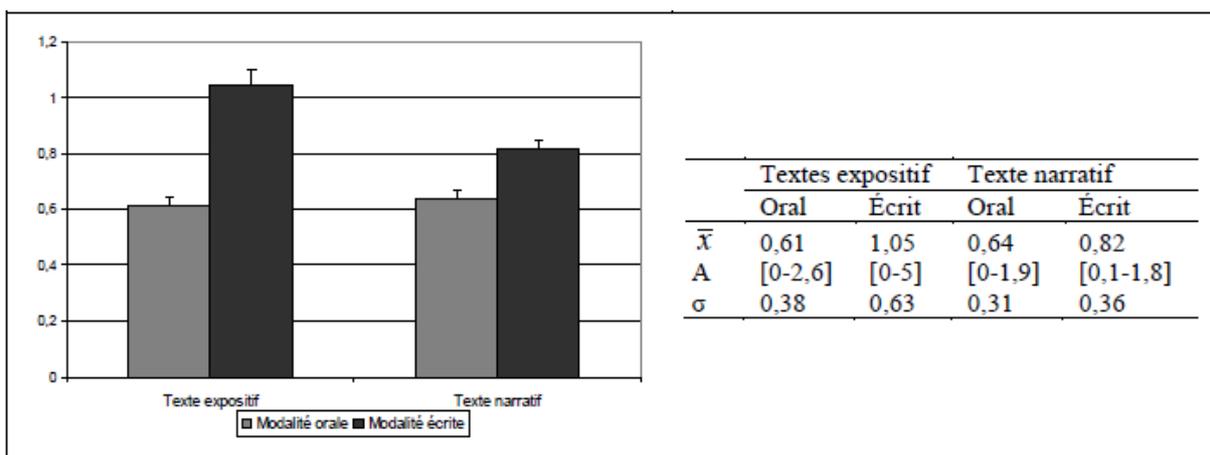
Les différences en fonction du NIVEAU SCOLAIRE sont significatives ($F_{(2,504)} = 2,874, p < 0,05$) : le nombre moyen d'items lexicaux par clause augmente avec le NIVEAU SCOLAIRE (Figure 35). Seuls les individus de CM2 et de 3^{ème}, se distinguent significativement ($p = 0,0238$).

Figure 35 : Nombre moyen d'items par clause selon le NIVEAU SCOLAIRE



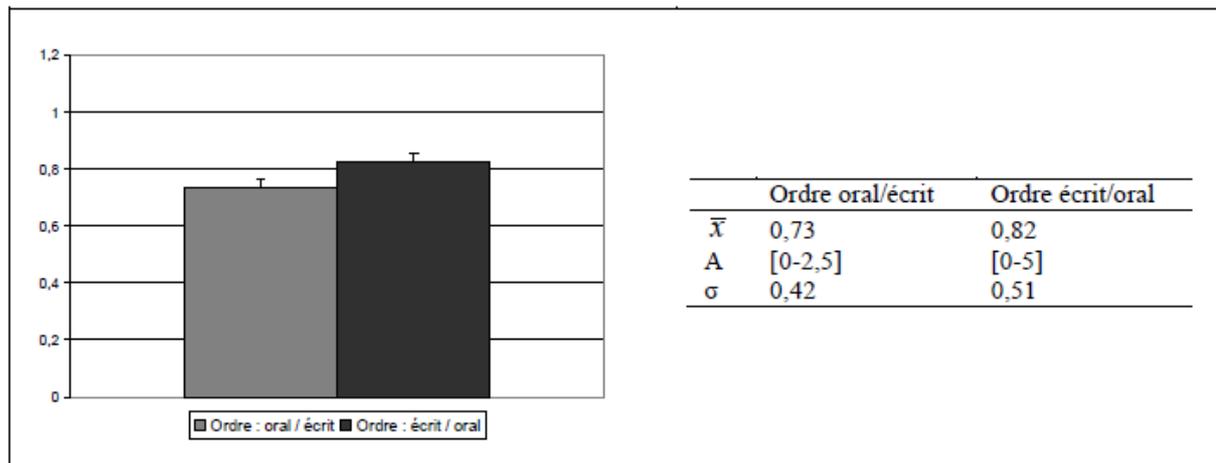
Le contexte de production fait également varier le nombre moyen d'items lexicaux par clause (TYPE DE TEXTE ($F_{(1,504)} = 7,095, p < 0,0080$) et MODALITÉ ($F_{(1,504)} = 65,045, p < 0,0001$)). Le nombre moyen d'items lexicaux par clause est plus élevé dans les textes expositifs et écrits que dans les productions narratives et orales (Figure 36).

Figure 36 : Nombre moyen d'items lexicaux par clause selon le TYPE DE TEXTE et la MODALITÉ



Enfin, les individus ayant produit dans l'ordre écrit/oral, produisent plus de noms dans une même clause que les individus ayant produit dans l'ordre oral/écrit ($F_{(1,504)} = 5,814, p < 0,0163$) (Figure 37).

Figure 37 : Nombre moyen d'items lexicaux par clause selon l'ORDRE DE PRODUCTION



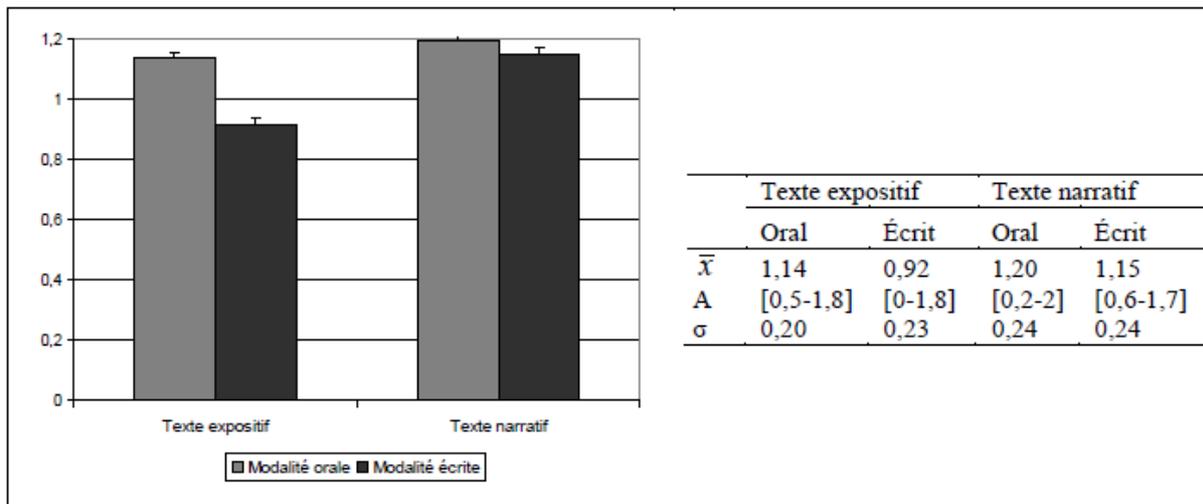
Une interaction entre le TYPE DE TEXTE et la MODALITÉ ($F_{(1,504)} = 11,854, p < 0,0006$) est à noter (Figure 36). En production expositive écrite, le nombre moyen d'items lexicaux par clause atteint son maximum. De plus, en contexte narratif, l'effet de la modalité est moindre (contrairement au contexte expositif).

8.1.3. Nombre moyen de pronoms par clause

Les figures 38 à 39 présentent le nombre moyen de pronoms par clause en fonction du TYPE DE TEXTE, de la MODALITÉ et de l'ORDRE DE PRODUCTION.

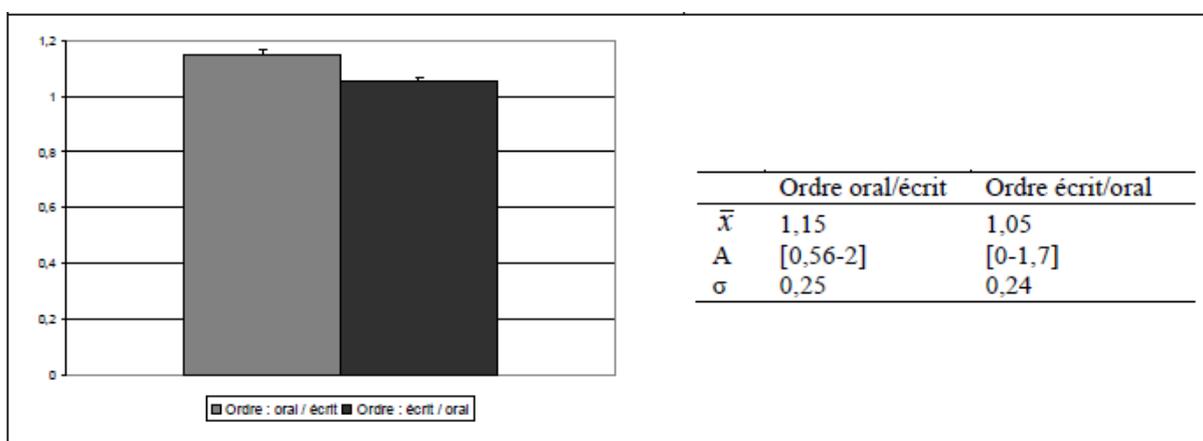
Les différences en fonction du contexte de production sont significatives (TYPE DE TEXTE ($F_{(1,504)} = 58,429, p < 0,0001$) MODALITÉ ($F_{(1,504)} = 48,326, p < 0,0001$)). Le nombre moyen de pronoms par clause est plus élevé dans les textes narratifs et oraux que dans les productions expositives et écrites (Figure 38).

Figure 38 : Nombre moyen de pronoms par clause selon le TYPE DE TEXTE et la MODALITÉ



L'ORDRE a également un effet ($F_{(1,504)} = 26,941$, $p < 0,0001$). Les individus ayant produit dans l'ordre oral/écrit tendent à employer plus de pronoms dans une même clause que les individus ayant produit dans l'ordre inverse (Figure 39).

Figure 39 : Nombre moyen de pronoms par clause selon l'ORDRE DE PRODUCTION



Le NIVEAU SCOLAIRE n'a pas d'effet significatif.

Les résultats révèlent une interaction entre le TYPE DE TEXTE et la MODALITÉ ($F_{(1,504)} = 22,064$, $p < 0,0001$) (Figure 38). Si en production expositive écrite, le nombre moyen de SNL par clause atteint son maximum, en production expositive orale et narrative orale, le nombre de pronoms par clause est à son maximum. De plus, en contexte narratif, encore une fois, l'effet de la modalité est moins importante qu'en contexte expositif.

8.1.4. Discussion des résultats de la Section 8.1

Le but de cette partie (8.1) était de voir si *les proportions des SNL et des pronoms* variaient en fonction du NIVEAU SCOLAIRE, du TYPE DE TEXTE, DE la MODALITÉ et de l'ORDRE.⁸⁶

Nous supposons que les plus jeunes des individus auraient une préférence pour l'utilisation des pronoms, puis ils en useraient moins au profit de SNL. L'analyse montre, en effet, que le NIVEAU SCOLAIRE affecte significativement *les productions des SNL et des pronoms* et confirme nos hypothèses. Les individus, avec les années de scolarisation, utilisent de moins en moins de pronoms et ont davantage recours aux noms lexicaux (103)/(104). Ces résultats vont dans le sens de travaux précédents (entre autres, Blanche-Benveniste, 1990, 1995 ; Argerich *et al.*, 2000 ; Berman, 2008 ; Berman et Verhoeven, 2002 ; Ravid, 2000, 2005, 2006 ; Ravid *et al.*, 2002 ; Strömquist *et al.*, 2002).

(103) *Il faudrait se parler c'est mieux. (CM2 n°05, expositif écrit, O/E)*

(104) *Les problèmes de communication chez les jeunes sont aussi liés au fait qu'ils [les adolescents] ne veulent pas communiquer comme dans la vidéo. (3^{ème} n°30, expositif écrit, O/E)*

Si l'écolier (103) préfère exprimer une idée sous forme pronominale, le collégien (104), par contraste, encode son message par le biais de SNL.

De plus, les participants de notre étude produisent de plus en plus de noms lexicaux dans une même clause (105) et (106).

(105) *On devrait essayer de se clamer. (CM2 n°02, expositif écrit, E/O)*

(106) *Au sein d'un établissement privé ou public, l'intégration des handicapés sera beaucoup plus dure que pour une personne « normale ». (3^{ème} n°07, expositif écrit, E/O)*

Si l'écolier n'utilise aucun nom lexical pour encoder son message, le collégien utilise cinq items lexicaux dans une même clause.

L'analyse des données montre que si les écoliers de CM2 se distinguent des individus de 3^{ème} et ceux de 5^{ème}, ces deux dernières populations ont un comportement assez similaire lorsqu'elles ont à choisir entre les *pronoms* et les *SNL*. Concernant le nombre moyen de noms par clause seules les populations extrêmes se distinguent : les CM2 produisent significativement moins d'items lexicaux dans une même clause que les collégiens de 3^{ème}.

⁸⁶ Les attentes, hypothèses et problématiques sont davantage explicitées au Chapitre 3 (Section 3.2.2.1, p. 99).

Une autre de nos attentes était l'influence du contexte de production (TYPE DE TEXTE et MODALITÉ) sur l'emploi des SNL et des pronoms. Pour des raisons évoquées dans les Chapitres 2 et 3, *les pronoms* seraient utilisés davantage dans les textes narratifs et oraux alors que *les SNL* seraient préférentiellement utilisés en contexte expositif et écrit.

Nos résultats vont dans le sens de nos prédictions : *les proportions des SNL et des pronoms* varient effectivement significativement avec le TYPE DE TEXTE. Les individus produisent davantage de SNL dans les textes expositifs que dans les textes narratifs, dans lesquels ils utilisent plus de pronoms. *Le nombre de noms lexicaux par clause* est également plus élevé dans les textes expositifs que dans les textes narratifs qui se caractérisent comme ayant *un nombre moyen de pronoms par clause* plus élevé. Ces tendances vont dans le sens des conclusions de travaux précédents selon lesquels le type de texte expositif est associé à un registre plus académique que le texte narratif engendrant une production plus importante de SNL (entre autres, Argerich *et al.*, 2000 ; Berman, 2008 ; Berman et Verhoeven, 2002 ; Ravid, 2000, 2005, 2006 ; Ravid *et al.*, 2002 ; Strömquist *et al.*, 2002). Pour illustrer ces changements, les exemples suivants sont proposés.

(107) Dans un *collège* les *sujets* tels que la *triche*, les *mises à l'écart* envers certaines *personnes*, les *bastons* sont plutôt fréquentes. (3^{ème} n°06, expositif écrit, O/E)

(108) J'ai connu un garçon / *qui*, dès qu'*il* parlait / ou dès qu'*il* arrivait au collège, / souffrait / et subissait des moqueries (3^{ème} n°06, narratif écrit, O/E)

En contexte d'exposition, le collégien utilise plus de noms lexicaux que de pronoms dans une même clause (*collège, sujets, triche, mises à l'écart, personnes, bastons*) alors qu'en narration ce même locuteur/scripteur encode son message plus par le biais de pronoms (*j', qui, il*). Si en exposition l'individu utilise sept noms lexicaux dans une même clause, en narration ce même enfant produit trois items lexicaux en cinq clauses (*garçon, collège, moqueries*).

Ces résultats peuvent être expliqués en nous référant aux définitions de ces deux types de textes. Comme nous l'avons expliqué dans la Section 2.4 (p. 70), les types de textes expositif et narratif impliquent des modes de pensée bien spécifiques. Ceci se concrétise par des choix linguistiques différents. En observant la proportion des SNL, nous concluons que les enfants font des choix en adéquation avec le contexte de production. Le texte expositif impose, entre autres, aux locuteurs/scripteurs une pression informative importante, l'utilisation d'un registre académique, qui se traduit par une utilisation plus importante de SNL.

Le texte narratif, dans notre étude, correspond à la narration d'expériences personnelles avec des informations spécifiques que les locuteurs relatent et maintiennent ; d'où une utilisation plus massive des pronoms.

Si nous ne pouvons observer la planification, nous pouvons nous concentrer sur les traces de la planification, dans ce travail de recherche, à travers les formes linguistiques. De par leurs choix linguistiques, nous pouvons penser que selon le contexte de production, les individus de notre étude mobilisent davantage un type de planification plutôt qu'un autre.

Ainsi, le type de texte expositif, par son exigence cognitive, fait plus appel à *la planification par connaissances transformées*, qui implique un effort d'adéquation entre deux espaces : celui des contenus et celui de la rhétorique. En revanche, le type de texte narratif fait plus appel à *la planification par connaissances rapportées*, qui correspond à un travail de récupération des contenus, des informations dans la *mémoire à long terme*, reformulés au fur et à mesure qu'ils sont récupérés.

La MODALITÉ implique, comme nous l'avons montré dans la Section 2.2 (p. 54), des circonstances de production très spécifiques (Berman, 2000 ; Berman *et al.*, 2002 ; Blanche-Benveniste, 1990, 1995, 2000 ; Cahana-Amitay et Berman, 1999 ; Fayol, 1997 ; Gayraud, 2000 ; Gayraud *et al.*, 2000 ; Jisa, 1998, 2000, Halliday, 1989 ; Johansson 1999 ; Ravid *et al.*, 2002). La modalité écrite se caractérise principalement par deux aspects : une contrainte temporelle assouplie et un temps de planification supérieur (Bonin, 2003 ; Fayol, 1997) ce qui permet la production de formes syntaxiques dont le traitement plus rares (Gayraud *et al.*, 2001 ; Jisa, 2004). La modalité orale se différencie de la modalité écrite par son temps de planification moindre et sa pression communicative supérieure (Blanche-Benveniste, 1990, 1995, 2000 ; Fayol, 1997 ; Gayraud, 2000). Ainsi, par les caractéristiques de ces deux modalités, nous comprenons que *les SNL* caractérisent plus les productions écrites, et que *les pronoms* caractérisent plus les textes oraux (110). De plus, si *le nombre moyen de noms par clause* est supérieur dans les textes écrits (111), la tendance s'inverse pour *le nombre moyen de pronoms par clause*, qui est plus élevé dans les oraux (112).

(109) Les **bagarres** commencent par des **regards** bizarres entre les **personnes**. (3^{ème} n°20, expositif écrit, E/O)

(110) **On** se regarde bizarrement. (3^{ème} n°20, expositif oral, E/O)

Ces exemples illustrent la situation : à l'écrit, le collégien utilise plus de SNL que de pronoms alors qu'à l'oral ce même locuteur/scripteur encode un message similaire par le biais de pronoms.

(111) À la **fin** du **contrôle**, mon **amie** avait copié la **moitié** de ma **feuille**. (3^{ème} n°42, narratif écrit, E/O)

(112) En fait **elle** avait presque **tout** copié en fait. (3^{ème} n°42, narratif oral, E/O)

À l'écrit et à l'oral, l'individu, pour une même information, effectue des choix linguistiques différents. (111) est une clause contenant cinq noms lexicaux alors que (112) est une clause ne contenant aucun nom mais deux pronoms.

Les résultats révèlent que l'ORDRE affecte les proportions des SNL et des pronoms ainsi que leur nombre moyen par clause, ce qui va dans le sens de nos prédictions. Le fait de produire d'abord à l'écrit et ensuite à l'oral implique une proportion d'autant plus importante de SNL par texte et par clause, ce qui est illustré avec (113) et (114), qui sont deux extraits de production d'un même individu appartenant à l'ordre de production écrit/oral.

(113) **Le meilleur moyen de résoudre ce problème est la paix** ou l'ignorance de l'autre. (5^{ème} n°06, expositif écrit, E/O)

(114) **Et euh le meilleur moyen de résoudre ce problème et ben ben c'est de faire la paix**. (5^{ème} n°06, expositif écrit, E/O)

Ces exemples montrent le transfert de formes plus typiques de l'écrit (un SNL lourd en position sujet constitué d'un infinitif) dans la production orale. La collégienne transfère les SNL activés et produits à l'écrit vers l'oral ; ainsi, *le meilleur moyen*, *ce problème* et *la paix* sont retrouvés à l'oral. En revanche, le fait de produire d'abord à l'oral et ensuite à l'écrit implique une proportion plus importante de pronoms par texte et par clause.

Le fait de produire de plus en plus de SNL (et de noms par clause) au détriment des pronoms avec les années de scolarisation est une des conclusions de travaux portant sur des populations francophones natives non défavorisées (Gayraud, 2000 ; Jisa et Mazur, 2005 ; Mazur, 2005 ; Salas, Tolchinsky, Mazur-Palandre, Jisa et Chenu, 2008). Les travaux de Gayraud (2000) concernant une population dite de classe moyenne se concentrent sur la proportion des SNL dans les textes narratifs, expositifs et argumentatifs, à l'écrit et à l'oral, de cinq groupes : des enfants de 7 ans, de 9 ans, de 12 ans, de 15 ans et d'adultes. Cette proportion augmente significativement avec l'âge. Les travaux de Jisa et Mazur (2005) et Mazur (2005) se

focalisent sur une population de classe moyenne supérieure (Projet Spencer). Quatre groupes sont observés : des enfants de 9/10 ans, de 12/13 ans, de 15/16 ans et des adultes. Les analyses révèlent que la proportion des SNL dans les textes augmente significativement avec l'âge. Néanmoins, cette population se scinde en deux, nous avons : (a) les 9/10 ans et les 12/13 ans qui se ressemblent ; (b) et les 15/16 ans et les adultes.

En regardant les proportions de plus près et en les comparant aux participants de ce travail de thèse (Tableau 9), les individus de CM2 (âge moyen = 10,9) et de 5^{ème} (âge moyen = 12,7) ne se distinguent guère des individus favorisés de 9/10 ans, de 12/13 ans.

[Tableau 9 : Proportion en pourcentage des SNL et des pronoms dans les textes selon l'âge et la Classe Socio-Professionnelle]

Tableau 9 : Proportion en pourcentage des SNL et des pronoms dans les textes selon l'âge et la Classe Socio-Professionnelle⁸⁷

	Population favorisée (Spencer)		Population défavorisée (ReFlex)	
	Pronom	SNL	Pronom	SNL
9/10 ans et CM2	70	30	66	34
12/13 ans et 5^{ème}	65	35	64	36
15/16 ans et 3^{ème}	52	48	64	36
Adultes	51	49	/	/

Néanmoins, l'écart se creuse quant à notre population de 3^{ème} (âge moyen = 15,2) et celle des 15/16 ans favorisés. Notons tout de même que les 15/16 ans du projet Spencer étaient en classe de 2^{nde} au lycée général, donc une classe de plus que nos 3^{ème}. Si les plus jeunes ne se différencient guère, l'écart se creuse plus tard : nos 3^{ème} se comportent plus comme des 5^{ème} que comme des 15/16 ans ou encore des adultes.

Le fait de produire des SNL davantage dans les textes expositifs et écrits que dans les textes narratifs et oraux est également un résultat trouvé lors de travaux antérieurs sur des populations de classe moyenne supérieure (Gayraud, 2000 ; Jisa et Mazur, 2005 ; Mazur, 2005). Les analyses de Gayraud (2000) montrent que pour tous les âges, les SN lexicaux sont plus fréquents dans les textes expositifs (et argumentatifs) et écrits que dans les textes narratifs et oraux ; ce qui est confirmé par les travaux de Jisa et Mazur (2005) et Mazur

⁸⁷ Les proportions sont arrondies.

(2005). Ainsi, notre population défavorisée n'a pas un comportement différent des populations de classe moyenne supérieure de Gayraud (2000) et de Jisa et Mazur (2005) et Mazur (2005). Nos hypothèses ont été en partie confirmées. En effet, les deux VD, *la proportion des SNL et des pronoms*, sont affectées par les facteurs étudiés – NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE–.

Ces analyses nous révèlent d'intéressantes tendances. Mais qu'en est-il si nous observons *la proportion des SNL et des pronoms* en prenant en considération les positions syntaxiques ? Est-ce que la proportion des SNL augmente avec l'âge dans les trois positions syntaxiques ? Est-ce que la proportion des SNL et des pronoms dans les différentes positions varie également selon les variables TYPE DE TEXTE, MODALITÉ et ORDRE ? La fonction syntaxique influence-t-elle la proportion de ces unités linguistiques ?

Afin de répondre à ces questions, nous allons, dans la prochaine section, présenter les analyses de *la proportion des SNL et des pronoms en position sujet, objet et autre*.

8.2. Proportion des SN selon les positions syntaxiques

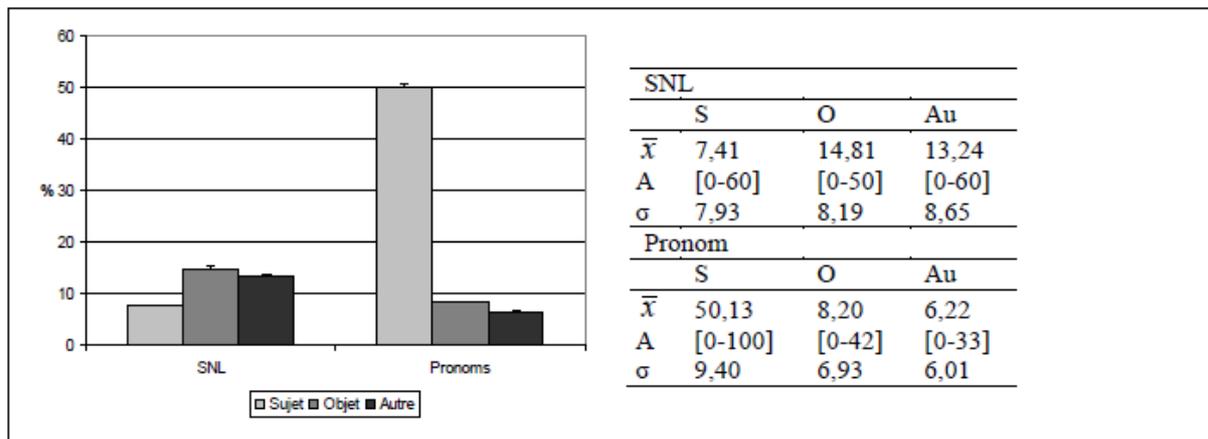
Dans cette partie, sont présentés les résultats concernant *les proportions des SNL et des pronoms* selon le nombre total de SN mais en distinguant les positions syntaxiques. Nous regardons, par exemple, *la proportion des SNL en position sujet selon le nombre total de SN*.

8.2.1. Formes et positions syntaxiques : comparaison

Les informations regroupées en Figure 40 résument *les proportions des SNL et des pronoms en position sujet, objet et autre* selon le nombre total de SN. Ainsi :

- 7,41% des SN sont des SNL en position sujet et 50,13% des pronoms en position sujet ;
- 14,81% des SN sont des SNL en position objet et 8,20% des pronoms dans cette même position ;
- 13,24% des SN sont des SNL en position autre et 6,22% des pronoms dans cette même position.

Figure 40 : Proportions des SNL selon les positions syntaxiques (par rapport au nombre total de SN)



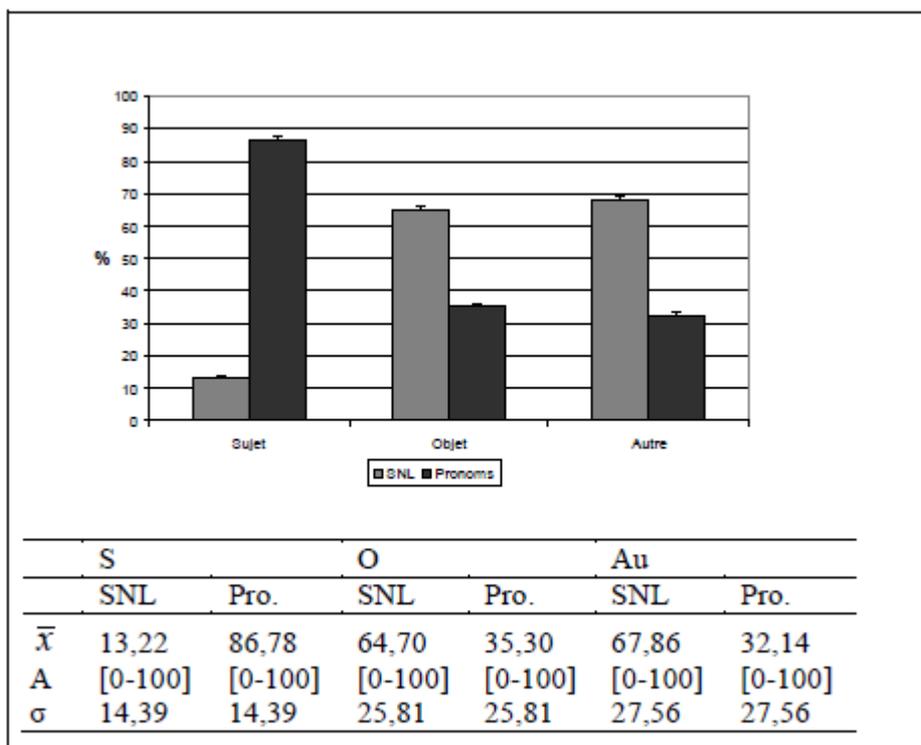
Une tendance peut être relevée : les SNL sont plus fréquents en position non sujet alors que les pronoms sont plus fréquents en position sujet.

Un autre type d'analyse a été effectué concernant *la proportion des SNL et des pronoms* selon leur position syntaxique. Nous avons observé : (a) *la proportion des SNL et des pronoms en position sujet* selon le nombre de SN total en sujet ; (b) *la proportion des SNL et des pronoms en position objet* selon le nombre de SN total en position objet ; et (c) *la proportion des SNL et des pronoms en position autre* selon le nombre de SN total en position autre.

Ces résultats confirment les tendances observées préalablement. Les informations regroupées en Figure 41 résument les moyennes des pourcentages des deux types des SN observés pour chaque position syntaxique (selon le nombre de SN total dans chaque position) tous facteurs confondus.

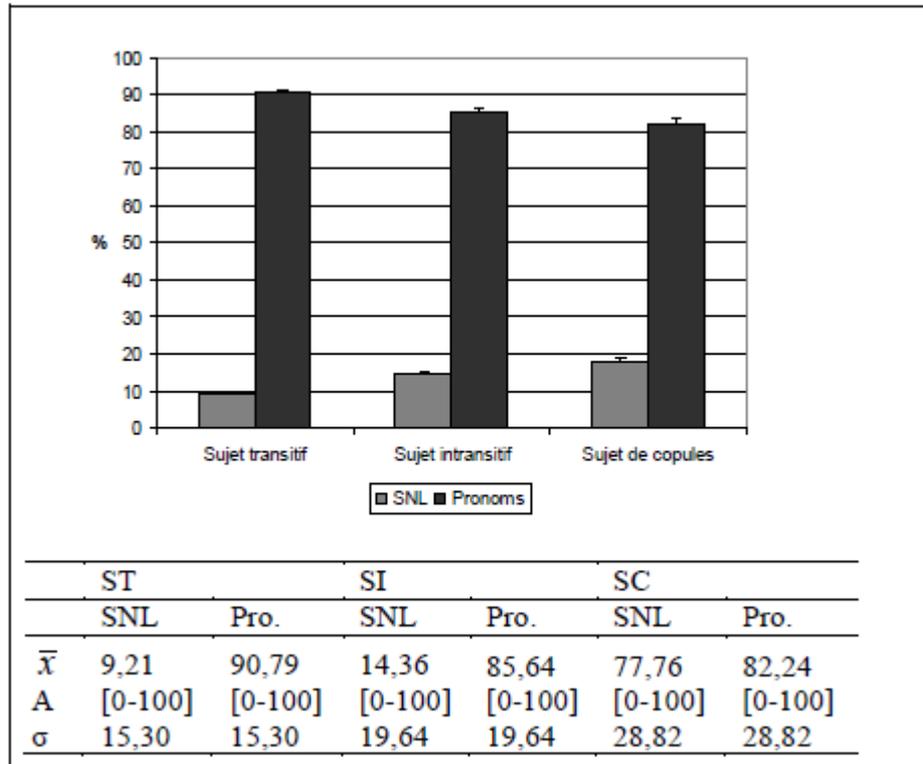
La position sujet est remplie davantage par des pronoms que par des SNL (13,22% des sujets sont lexicaux contre 86,78% pronominaux) (Test T : $t_{(527)} = 58,746$, $p < 0,0001$). La tendance s'inverse pour les positions objet et autres qui contiennent plus de SNL que de pronoms (64,70% des SN en position objet sont lexicaux contre 35,30% pronominaux (Test T : $t_{(513)} = 12,913$, $p < 0,0001$) ; 67,86% des SN autre sont lexicaux contre 32,14% pronominaux (Test T : $t_{(500)} = 14,505$, $p < 0,0001$)).

Figure 41 : Les positions syntaxiques, SNL versus pronoms



La tendance observée pour la position sujet est également vérifiée pour les trois types de sujets (de verbes transitifs, intransitifs et copules). La position sujet transitif (Test T : $t_{(514)} = 60,49$, $p < 0,0001$), la position sujet intransitif (Test T : $t_{(509)} = -40,97$, $p < 0,0001$) et la position sujet de copules (Test T : $t_{(406)} = 22,57$, $p < 0,0001$) sont plus pronominaux que lexicaux (Figure 42). La position sujet intransitif est plus lexicale que la position sujet transitif.

Figure 42 : La position sujet, SNL versus pronoms



8.2.2. Proportion des SNL selon les positions syntaxiques

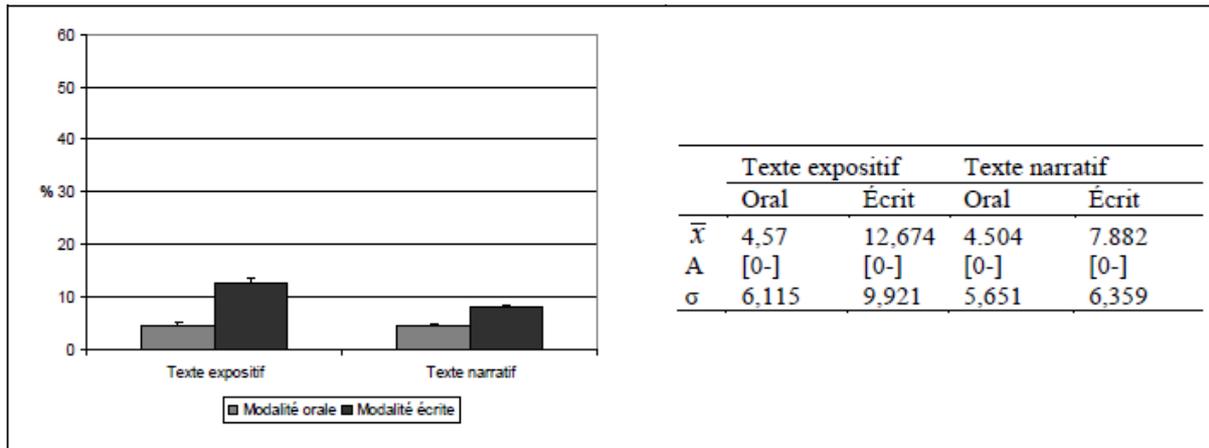
8.2.2.1. Proportion des SNL en position sujet

Les figures 43 et 44 présentent *les proportions des SNL en position sujet* en fonction du TYPE DE TEXTE et de la MODALITÉ.

Les différences en fonction du TYPE DE TEXTE sont significatives ($F_{(1,504)} = 15,097$, $p < 0,0001$). *La proportion des SNL en position sujet* est plus importante dans les productions expositives que dans les productions narratives (Figure 43).

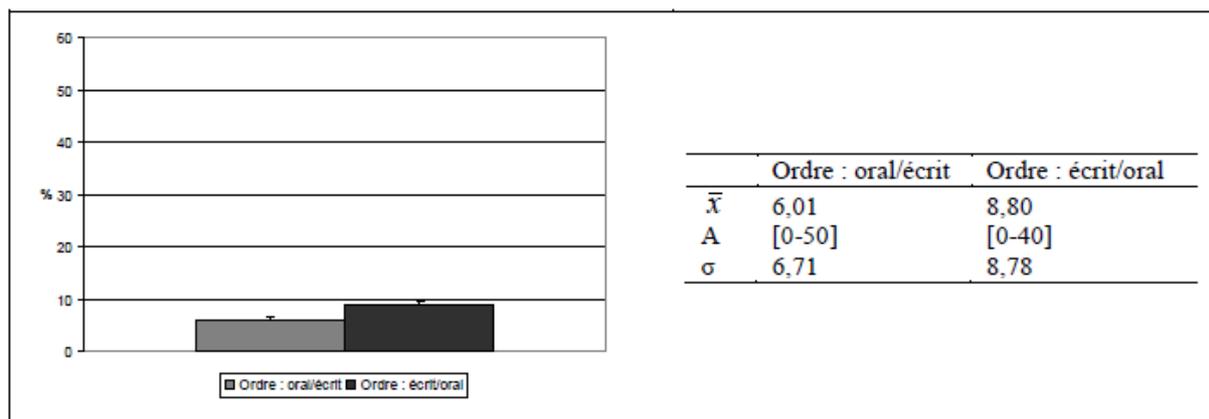
La MODALITÉ a également un effet ($F_{(1,504)} = 84,868$, $p < 0,0001$) : les SNL en position sujet semblent davantage caractériser les textes écrits que les textes oraux (Figure 43).

Figure 43 : Proportions des SNL en position sujet selon le TYPE DE TEXTE et la MODALITÉ (par rapport au nombre total de SN)



Les individus ayant produit dans l'ordre écrit/oral produisent davantage de SNL en position sujet que ceux ayant produit dans l'ordre oral/écrit ($F_{(1,504)} = 20,215$, $p < 0,0001$) (Figure 44).

Figure 44 : Proportions des SNL en position sujet selon l'ORDRE DE PRODUCTION (par rapport au nombre total de SN)



Le facteur NIVEAU SCOLAIRE n'a pas d'effet significatif sur cette VD.

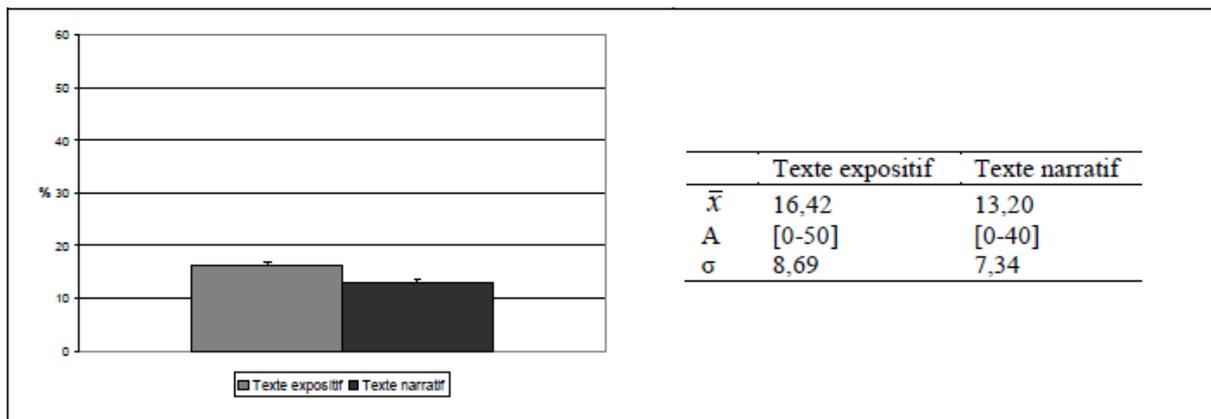
Les analyses révèlent également une interaction entre les VI TYPE DE TEXTE et MODALITÉ ($F_{(1,504)} = 14,441$, $p < 0,0002$) (Figure 43). La proportion des SNL en position sujet atteint son maximum en production expositive écrite.

En contexte expositif, la différenciation des modalités est davantage effectuée qu'en contexte narratif. Enfin, en contexte écrit, la différenciation entre les deux types de textes est davantage visible qu'en contexte oral.

8.2.2.2. Proportion des SNL en position objet

La Figure 45 renseigne sur *les proportions des SNL en position objet* en fonction du TYPE DE TEXTE. Seules les différences en fonction du TYPE DE TEXTE sont significatives ($F_{(1,504)} = 20,969$, $p < ,0001$). *La proportion des SNL en position objet* est plus importante en production expositive qu'en production narrative (Figure 45).

Figure 45 : Proportions des SNL en position objet selon le TYPE DE TEXTE (par rapport au nombre total de SN)



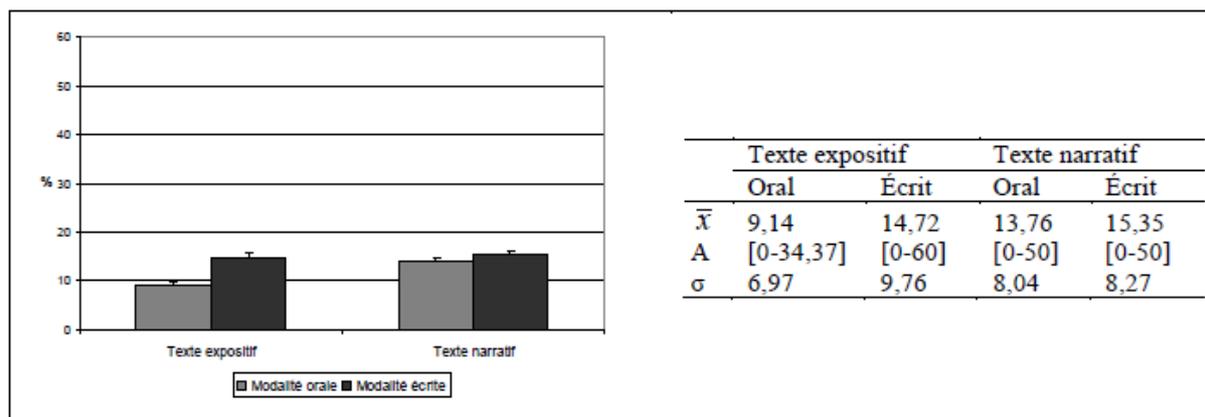
8.2.2.3. Proportion des SNL en position autre

La Figure 46 représente *les proportions des SNL en position autre* en fonction du TYPE DE TEXTE et de la MODALITÉ.

Les différences en fonction du TYPE DE TEXTE ($F_{(2,504)} = 13,692$, $p < ,0002$) et de la MODALITÉ ($F_{(1,504)} = 24,407$, $p < 0,0001$) sont significatives. *La proportion des SNL en position autre* est plus importante dans les productions narratives et les productions écrites que dans les textes expositifs et les textes oraux (Figure 46).

Les variables NIVEAU SCOLAIRE et ORDRE n'ont pas d'effet significatif.

Figure 46 : Proportions des SNL en position autre selon le TYPE DE TEXTE et la MODALITÉ (par rapport au nombre total de SN)



Les analyses révèlent également une interaction entre le TYPE DE TEXTE et la MODALITÉ ($F_{(1,504)} = 7,812, p < 0,0054$) (Figure 46). En contexte expositif, la différenciation des deux modalités du langage est davantage marquée : *la proportion de SNL en position autre* est plus importante à l'écrit qu'à l'oral. En production orale, la distinction entre les deux types de textes est plus visible, ce qui se concrétise par une production *de SNL en position autre* plus importante en narration.

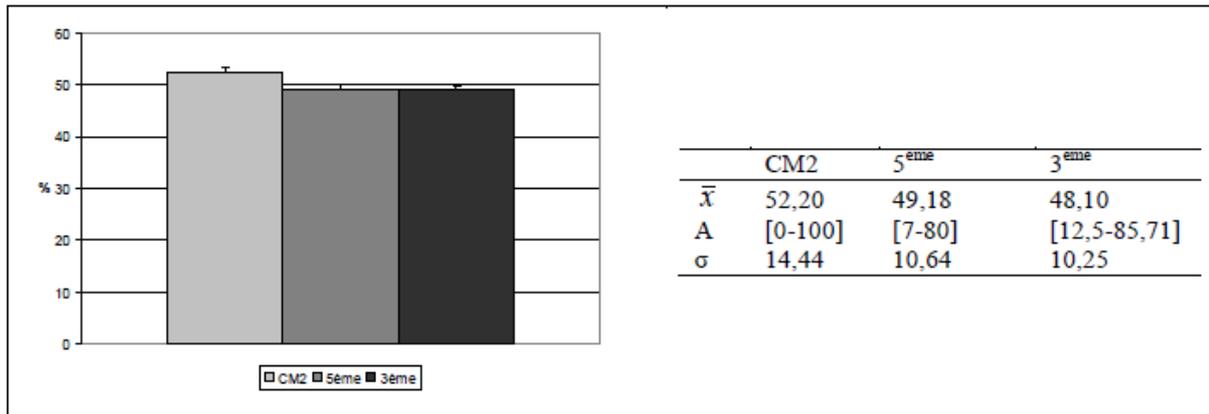
8.2.3. Proportion des pronoms selon les positions syntaxiques

8.2.3.1. Proportion des pronoms en position sujet

Les figures 47 à 49 présentent *les proportions des pronoms en position sujet* en fonction du NIVEAU SCOLAIRE, du TYPE DE TEXTE, de la MODALITÉ et de l'ORDRE DE PRODUCTION.

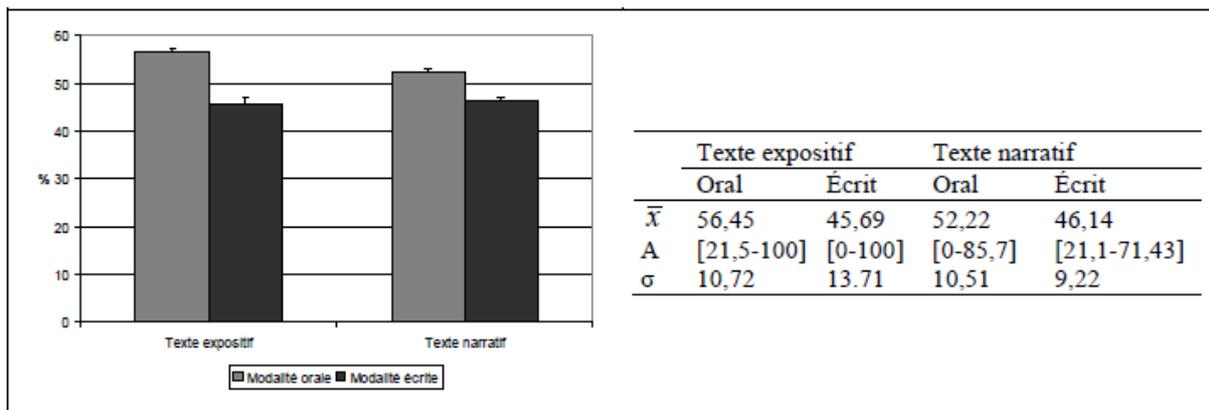
Les différences en fonction du NIVEAU SCOLAIRE sont significatives ($F_{(2,504)} = 4,808, p < 0,0085$). En effet, *la production de pronoms en position sujet* diminue avec les années de scolarisation (Figure 47). Le test post-hoc révèle que si les écoliers de CM2 produisent significativement moins de pronoms dans cette position que les collégiens de 5^{ème} ($p = 0,0102$) et de 3^{ème} ($p = 0,0060$), ces deux dernières populations ont des proportions très proches ($p = 0,8779$).

Figure 47 : Proportions des pronoms en position sujet selon le NIVEAU SCOLAIRE (par rapport au nombre total de SN)



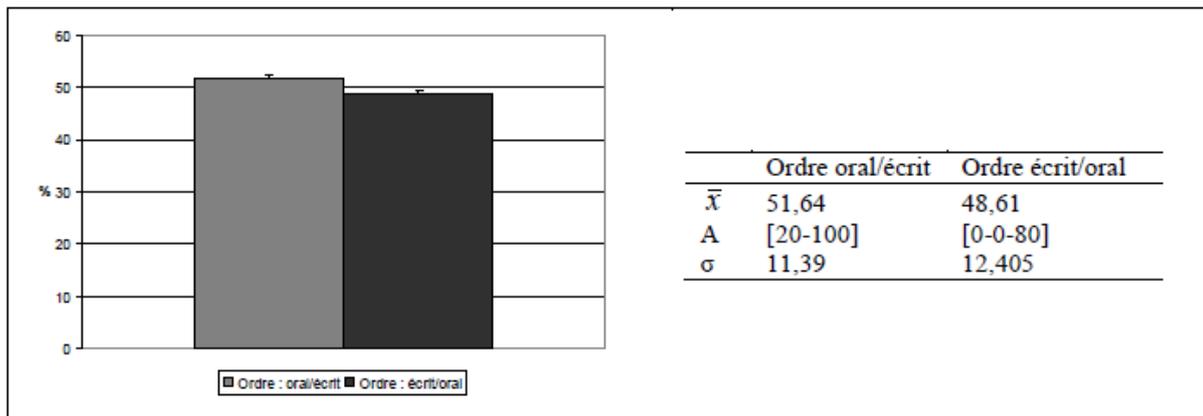
Le TYPE DE TEXTE et la MODALITÉ font également varier *les pronoms en position sujet* qui sont plus importants dans les textes expositifs et oraux que dans les textes narratifs et écrits ($F_{(1,504)} = 4,027$, $p < 0,0453$, $F_{(1,504)} = 77,312$, $p < 0,0001$) (Figure 48).

Figure 48 : Proportions des pronoms en position sujet selon le TYPE DE TEXTE et la MODALITÉ (par rapport au nombre total de SN)



Enfin, l'ORDRE a un effet sur *la proportion des pronoms en position sujet* ($F_{(1,504)} = 10,036$, $p < 0,0016$). Les individus ayant produit dans l'ordre oral/écrit ont tendance à utiliser davantage de pronoms en position sujet que les individus ayant produit dans l'ordre inverse (Figure 49).

Figure 49 : Proportions des pronoms en position sujet selon l'ORDRE DE PRODUCTION (par rapport au nombre total de SN)



Les analyses révèlent également une interaction entre le TYPE DE TEXTE et la MODALITÉ ($F_{(1,504)} = 6,117$, $p < 0,0137$) (Figure 48). *La proportion des pronoms en position sujet* atteint son maximum en contexte expositif oral. En contexte expositif, la différence de proportion entre les deux modalités de production est plus importante qu'en contexte narratif.

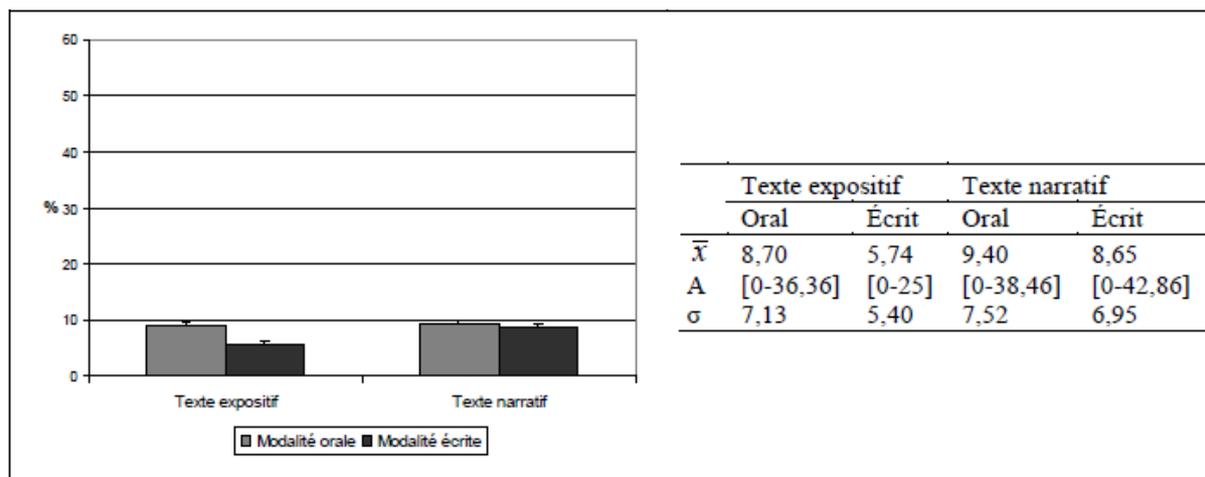
8.2.3.2. Proportion des pronoms en position objet

La Figure 50 renseigne sur *les proportions des pronoms en position objet* en fonction du TYPE DE TEXTE et de la MODALITÉ.

Les différences selon le contexte de production sont significatives (TYPE DE TEXTE ($F_{(1,504)} = 7,910$, $p < 0,0051$) et MODALITÉ ($F_{(1,504)} = 11,575$, $p < 0,0001$)). *La proportion des pronoms en position objet* semble davantage caractériser les textes oraux et narratifs que les textes écrits et expositifs (Figure 50).

Les deux autres facteurs de cette étude, NIVEAU SCOLAIRE et ORDRE, n'interviennent pas.

Figure 50 : Proportions des pronoms en position objet selon le TYPE DE TEXTE et la MODALITÉ (par rapport au nombre total de SN)



Les analyses révèlent également une interaction entre le TYPE DE TEXTE et la MODALITÉ ($F_{(1,504)} = 4,524$, $p < 0,0339$) (Figure 50). En contexte expositif, la différence de proportion entre les modalités du langage est plus importante à l'écrit qu'à l'oral. Cette différenciation en contexte narratif est moindre. Enfin, en contexte écrit, la différence de proportion entre les deux types de textes est plus importante qu'en contexte oral.

8.2.3.3. Proportion des pronoms en position autre

Les figures 51 et 52 présentent *les proportions des pronoms en position autre* en fonction du TYPE DE TEXTE et de la MODALITÉ.

Les différences selon le contexte de production sont significatives (TYPE DE TEXTE ($F_{(1,504)} = 41,512$, $p < 0,0001$) et MODALITÉ ($F_{(1,504)} = 4,249$, $p < 0,0398$)). *Les pronoms en position autre* caractérise plus les textes écrits et narratifs que les textes oraux et expositifs (Figure 51 et Figure 52).

Le NIVEAU SCOLAIRE et l'ORDRE ne font pas varier significativement cette VD.

Figure 51 : Proportions des pronoms en position autre selon le TYPE DE TEXTE (par rapport au nombre total de SN)

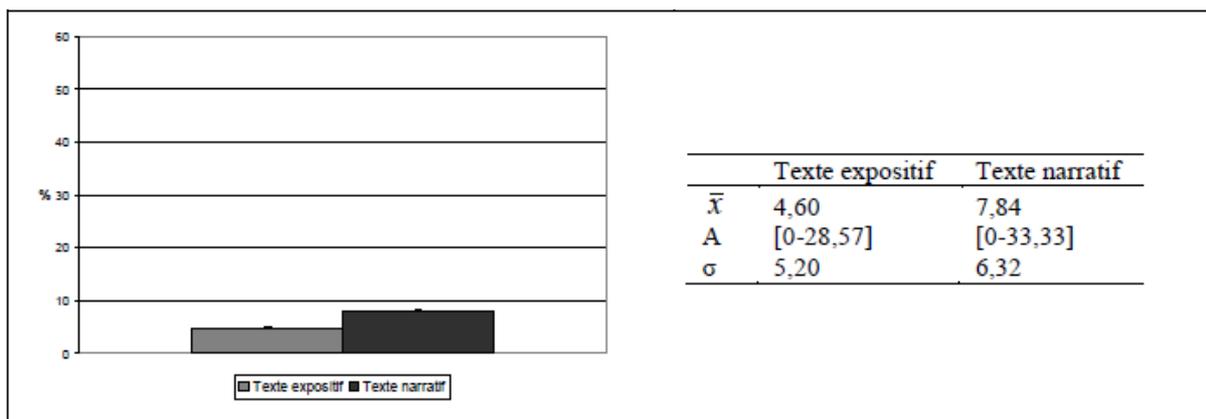
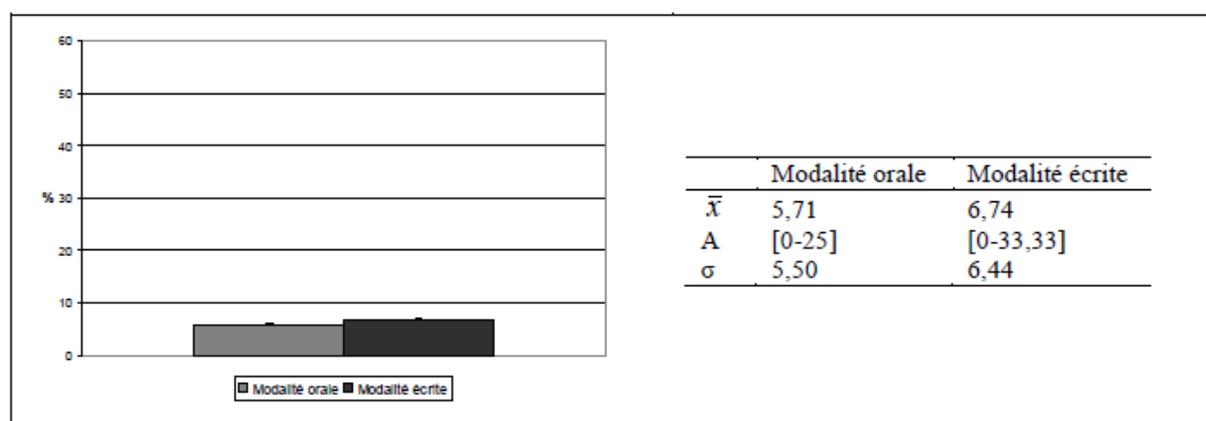


Figure 52 : Proportions des pronoms en position autre selon la MODALITÉ (par rapport au nombre total de SN)



8.2.4. Discussion des résultats de la Section 8.2

Pour des raisons évoquées dans la Section 2.2.2 (p. 99), nous présumons que : (a) *les proportions des SNL et des pronoms* selon les positions syntaxiques varieraient également selon les facteurs de cette étude ; et (b) effectuer des analyses en prenant en considération les positions syntaxiques (Figure 40 à Figure 52) compléterait les résultats concernant *la proportion des SNL et des pronoms* sans prendre en considération les positions (Figure 28 à Figure 39).⁸⁸

Rappelons que *la proportion des SNL* toutes positions confondues augmente avec le NIVEAU SCOLAIRE, est plus élevée dans les textes expositifs et plus importante dans les textes écrits. Pour finir, cette proportion est également plus importante dans les textes des individus appartenant à l'ordre de passation écrit/oral. De plus, *la proportion des pronoms* diminue avec

⁸⁸ Les attentes, hypothèses et problématiques sont davantage explicitées au Chapitre 3 (Section 3.2.2.1, p. 99).

les années de scolarisation, sont plus fréquents dans les textes narratifs et oraux que dans les textes expositifs et écrits. Enfin, les textes ayant été produits par des individus appartenant à l'ordre de passation oral/écrit ont une plus grande proportion de pronoms.

Le NIVEAU SCOLAIRE affecte significativement *la proportion des pronoms en position sujet* qui diminue avec les années de scolarisation.⁸⁹ Néanmoins, si les écoliers de CM2 (115) se distinguent significativement des 5^{ème} et des 3^{ème} (116), ces deux dernières populations produisent une proportion de pronoms quasi équivalente. Les exemples qui suivent illustrent cette situation.

(115) *Plus **on** se bagarre, plus **ça** fait de grosses disputes.* (CM2 n°18, expositif écrit, O/E)

(116) ***Les facteurs des problèmes** sont souvent la jalousie, l'impulsivité.* (3^{ème} n°14, expositif écrit, O/E)

Si l'écolier utilise un pronom en position sujet, *on* ou *ça*, le collégien, par contraste, remplit cette position avec un SNL, *les facteurs des problèmes*.

Les proportions des SNL et des pronoms selon leurs positions syntaxiques varient également de manière significative selon le TYPE DE TEXTE. Les analyses révèlent que *les proportions des SNL en position sujet et objet* confirment la tendance générale en étant plus élevées dans les textes expositifs (117) que dans les textes narratifs (118).

(117) ***La dispute** éclate lorsque **plusieurs personnes** ne sont pas d'accord. **Le meilleur moyen de résoudre ce problème, est la paix ou l'ignorance** de l'autre parfois, **les gens concerné** font appel à la violence.* (5^{ème} n°06, expositif écrit, E/O)

(118) *Elles sont plus âgées que moi comme nous sommes très complices nous avons prévu **une partie de cache-cache** au bout d'un quart d'heure, je n'ai plus voulu jouer.* (5^{ème} n°06, narratif écrit, E/O)

Ces deux exemples sont révélateurs de l'impact du type de texte sur la proportion des SNL dans ces deux positions : en effet, (117) est beaucoup plus lexical que (118).

Néanmoins, une précision apparaît concernant *la proportion des SNL en position autre* qui est plus importante dans les textes narratifs que dans les textes expositifs. L'utilisation de *SNL en position autre* sert à l'introduction de compléments circonstanciels consistant à compléter des descriptions de situation, aspects caractérisant davantage les productions narratives.

⁸⁹ Nous tenons à signaler que les analyses de variance concernant la proportion des SNL en position sujet transitif laissent apparaître un effet principal de l'âge : la production de SNL en position sujet transitif augmente avec l'âge.

De plus, si *les proportions des pronoms en position objet et autre* vont dans le sens des résultats de la proportion des pronoms toutes positions confondues, *la proportion des pronoms en position sujet* se distingue. En effet, *les proportions des pronoms en position objet et autre* sont plus élevées dans les textes narratifs que dans les textes expositifs. En revanche, *la proportion des pronoms en position sujet* est plus importante dans les textes expositifs que dans les textes narratifs.

Les proportions des SNL et des pronoms selon leurs positions syntaxiques varient également selon la MODALITÉ DE PRODUCTION. Les analyses distinguant les positions syntaxiques mettent en avant que *les proportions des SNL en position sujet et autre* sont plus importantes en production écrite (119) qu'en production orale (120). Ceci confirme la tendance observée lors des tests plus généraux. Cependant, *la proportion des SNL en position objet* ne varie pas selon la MODALITÉ.

(119) Mais **cette personne** l'a très mal pris. **Sa première réaction** a été de frapper. (3ème n°45, expositif écrit, E/O)

(120) **I(ls)** cherchent pas pourquoi, **i(ls)** s'enflamment, puis après **i(ls)** s(e) frappent. (3ème n°45, expositif oral, E/O)

En production écrite, la position sujet est remplie par un item lexical (*cette personne / sa première réaction*). En revanche, en production orale, le même individu remplit cette position avec un pronom (*i(ls)*).

(121) Une élève demande **la réponse** à son camarade. (5ème n°40, expositif écrit)

(122) Alors après euh ils demandent \emptyset à leurs camarades. (5ème n°40, expositif oral)

Nous remarquons qu'en production écrite (121), le collégien encode son message en produisant un SNL en position objet (*la réponse*) alors qu'en production orale (122), la position objet n'est pas remplie.

Les résultats de la proportion des pronoms selon les positions syntaxiques dévoilent que *les proportions des pronoms en position sujet et objet* sont effectivement plus importantes dans les textes oraux. Néanmoins, la tendance s'inverse pour *la proportion des pronoms en position autre* qui est davantage élevée dans les productions écrites que dans les productions orales.

La position sujet pose problème aux individus lorsqu'il s'agit de la remplir par des formes lexicales et ce notamment en production orale. En revanche, en production écrite ils arrivent à

pallier cette difficulté. *La position autre* est davantage remplie par des SNL en production écrite. À l'écrit, les individus répondent à la pression informative de cette modalité en enrichissant son texte d'informations secondaires (compléments circonstanciels), ce qui suppose une production de SNL. En revanche, produire des *SNL en position objet* ne semble pas poser de problème et être le choix linguistique le plus évident et le plus naturel et ce, quelle que soit la modalité.

L'ORDRE DE PRODUCTION affecte seulement deux variables : *les proportions des SNL et des pronoms en position sujet*. La proportion des SNL en position sujet est plus élevée dans les productions produites par des individus appartenant à l'ordre écrit/oral ce qui confirme la tendance relevée par les analyses toutes positions confondues. Seule *la proportion des pronoms en position sujet* confirme la tendance observée lors des analyses toutes positions confondues.

8.3. Conclusion du Chapitre 8

Les analyses prenant en considération les positions syntaxiques complètent les résultats *des proportions des SNL et des pronoms* (Section 8.1, p. 176) toutes positions confondues. Les tests ne prenant pas en compte les positions syntaxiques, nous renseignent sur la faible proportion des SNL par rapport à celle des pronoms. Ces analyses dévoilent alors qu'il est plus fréquent de produire des pronoms que des SNL. Les tests selon les positions syntaxiques complètent cette conclusion en montrant que, en plus, il est plus évident pour les locuteurs/scripteurs de produire des SNL dans des positions non sujet. Les positions syntaxiques ne sont pas traitées de la même façon : pour une même unité, selon sa position, la proportion varie selon des facteurs différents.

Nous pouvons conclure en mettant ces résultats en lien avec les travaux de Du Bois sur les principes de transmission de l'information de la langue orale (1987, 2003). La transmission de l'information est différente selon les modalités de production.⁹⁰ À l'oral, cette transmission est contrainte, entre autres, par le rythme de production rapide. Une des conséquences est la production d'*un seul élément lexical par clause*, une des contraintes de la PAS.

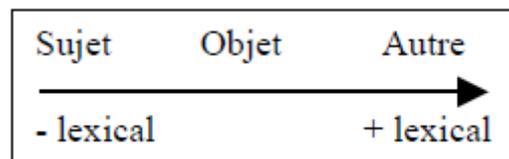
Cette contrainte se vérifie pour notre population de francophones natifs notamment pour les textes oraux et les textes narratifs. Or les individus, lorsqu'ils sont face à des contraintes

⁹⁰ Se reporter à la Section 2.2 du Chapitre 2 (p. 54) dans laquelle ces notions sont davantage travaillées.

contextuelles différentes, telles que le texte expositif et la modalité écrite, apprennent à aller à l'encontre de cette contrainte naturelle. *Le nombre moyen d'items lexicaux par clause* pour les textes expositifs et les textes écrits est significativement plus élevé. En production écrite, la transmission de l'information se réalise différemment, les choix syntaxiques en sont donc affectés. Cette contrainte de la PAS ne semble pas résister à la modalité écrite ainsi qu'au texte expositif.

La PAS de Du Bois (1987, 2003) propose une autre contrainte selon laquelle *les éléments lexicaux ne sont pas en position sujet*, et notamment en position sujet transitif. Cette seconde contrainte est respectée : la position sujet est la position la moins lexicale, ce qui rappelle la notion de *light subject constraint* de Chafe (1994) et les conclusions de divers travaux (Blanche-Benveniste, 1990 ; François, 1974 ; Jeanjean, 1980 ; Lambrecht, 1984, 1987 ; Mazur, 2005 ou encore Ravid et Shlesinger, 1995). Pour les positions non sujet la tendance s'inverse ; les positions objet et autre sont plus souvent remplies par des SNL que par des pronoms (Figure 41). Un continuum peut alors être tracé (Figure 53).

Figure 53 : Schématisation de la lexicalité des positions



Des analyses de variance différenciant trois positions sujets ont été réalisées : sujet de verbes transitifs, sujet de verbes intransitifs et sujet de verbes copules.⁹¹ En français, les positions sujet de verbes transitifs, sujet de verbes intransitifs et sujet de verbes copules sont traitées de la même façon.⁹² Même si la position sujet intransitif semble être tout de même un peu plus lexicale que la position sujet transitif, la position sujet intransitif se rapproche plus de la position sujet transitif que de la position objet ou autre.

De plus, si pour la position sujet la différence entre la proportion de SNL et celle de pronoms est assez importante, elle diminue pour les positions non sujet. La position sujet pose davantage problème aux individus. Ceci se traduit en produisant (a) moins de SNL, plus coûteux cognitivement, et (b) plus de pronoms, moins coûteux, en positions sujet. Les

⁹¹ Ces statistiques ne révélant rien de plus que celles confondant ces trois positions sujet (présentés dans ce travail), nous avons décidé de ne pas les présenter.

⁹² Contrairement aux langues étudiées par Du Bois et ses collègues où il apparaît que la position sujet de verbes intransitifs se rapproche plus de la position objet que de la position sujet de verbes transitifs.

individus sont capables de produire des SNL mais de préférence en position objet et autre : la position sujet est plus délicate à gérer cognitivement. Ainsi, cumuler position sujet et SNL semblent être trop coûteux et afin d'éviter une surcharge cognitive, les formes lexicales sont préférentiellement en position non sujet.

Pour finir, le fait de produire davantage de SNL en position non sujet qu'en position sujet est également un résultat des travaux sur des populations de classe moyenne supérieure (Gayraud, 2000 ; Jisa et Mazur, 2005 ; Mazur, 2005 ; Salas *et al.*, 2008).

Après avoir observé *la proportion des SNL et des pronoms* dans les textes, il peut également être intéressant de se concentrer plus précisément sur le degré de lexicalité des SNL. En effet, la production du SNL, notamment en position sujet, semble être un indice de maturité linguistique. C'est dans cette mesure que le chapitre suivant est alors dédié au degré de lexicalité du SNL.

Chapitre 9

Le degré de lexicalité des syntagmes nominaux lexicaux

Lorsque nous étudions le SNL, nous pouvons nous intéresser à sa complexité linguistique. Nous avons vu au Chapitre 1 (section 3, p. 40) que ce concept réfère à la complexité lexicale et l'architecture syntaxique (Ravid, 2005 ; Ravid et Berman, sous presse). Pour notre travail de thèse, nous nous concentrons davantage sur une des mesures pouvant calculer la complexité syntaxique d'un SNL : le nombre de nœud lexicaux.

Ainsi, dans un premier temps, nous nous sommes penchée sur *la proportion des SNL composés d'un nœud lexical* (123) et *de plus d'un nœud lexical*⁹³ (124). Nous parlons alors de SNL simple *versus* SNL complexe.

(123) *Les problèmes seraient résolus comme ça. (3^{ème} n°27, expositif écrit, O/E)*

(124) *Les problèmes entre les gens sont de plusieurs sortes. (3^{ème} n°25, expositif écrit, E/O)*

Dans un second temps, nous nous sommes concentrée sur la proportion des SNL simples et complexes selon leur position syntaxique. Les résultats concernant *la proportion des SNL*

⁹³ La notion de nœud est abordée dans la Section 1.3 du Chapitre 1 (p. 42/43 et 46/47).

constitués d'un nœud (125), de plus d'un nœud (126) en position sujet, objet et autre sont présentés.

(125) *Les disputes peuvent engendrer des bagarres.* (3^{ème} n°01, expositif écrit, O/E)

(126) *Mon professeur de technologie aperçut l'adresse du blog.* (3^{ème} n°01, narratif écrit, O/E)

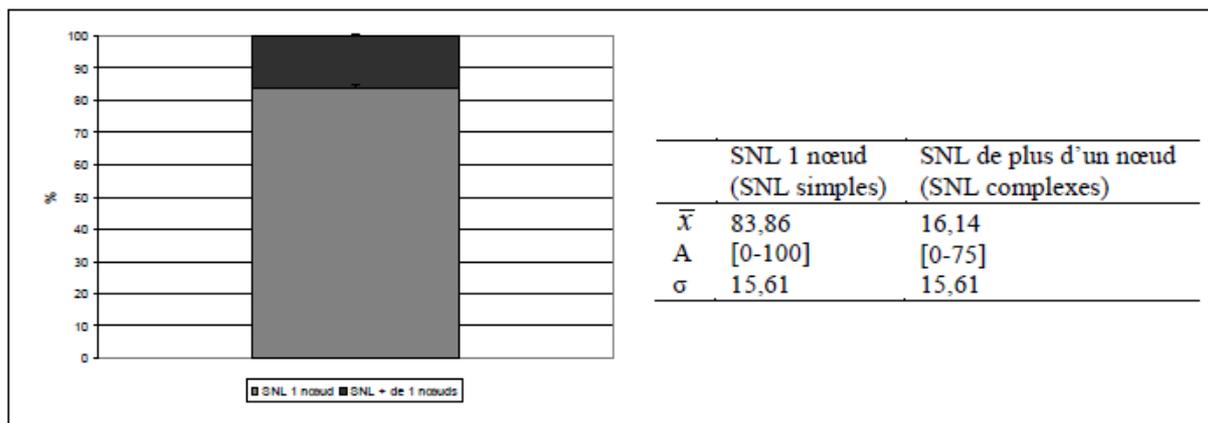
9.1. Analyses générales

Dans cette partie est présentée les proportions des SNL constitués d'un nœud lexical et de plus d'un nœud lexical selon le nombre total de SNL.⁹⁴

9.1.1. Indicateurs préliminaires

Les informations de la Figure 54 présentent les proportions des SNL constitués d'un nœud lexical et de plus d'un nœud lexical. Les SNL sont formés en majorité d'un seul nœud lexical (plus de 80%) alors que les SNL ayant plus d'un nœud sont plus rares (16%) (Test T : $t_{(524)} = 49,71$, $p < 0,0001$).

Figure 54 : Proportions des SNL selon leur degré de lexicalité et leur position syntaxique



⁹⁴ Pour les 525 textes contenant des SNL sur les 528 textes au total.

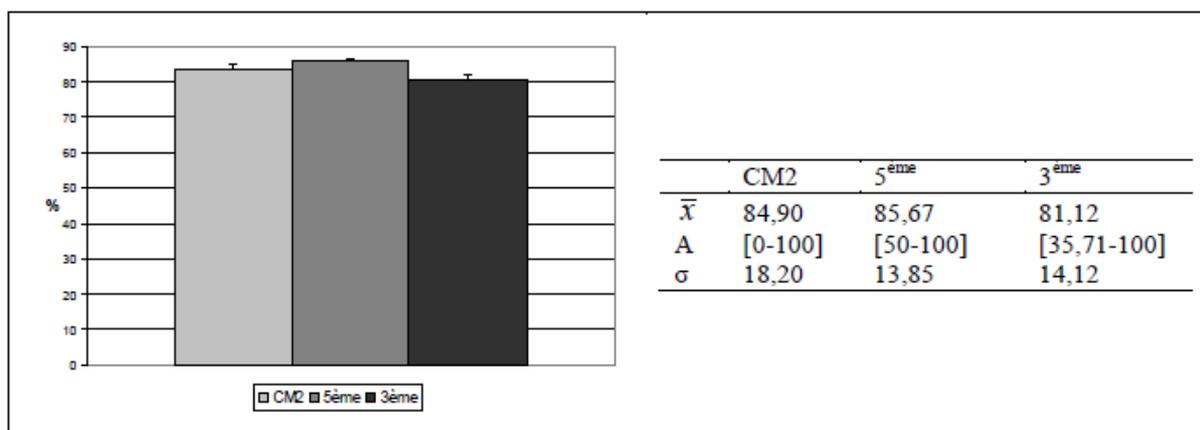
9.1.2. Proportion des SNL composés d'un et de plus d'un nœud lexical

Les analyses qui suivent présentent *les proportions des SNL composés d'un nœud lexical et des SNL composés de plus d'un nœud lexical* par rapport au nombre total de SNL en fonction du NIVEAU SCOLAIRE, du TYPE DE TEXTE, de la MODALITÉ et de l'ORDRE DE PRODUCTION (Figure 55 à Figure 62). Les différences en fonction du NIVEAU SCOLAIRE ($F_{(2,501)} = 4,774$, $p < 0,0088$), du TYPE DE TEXTE ($F_{(1,501)} = 15,402$, $p < 0,0001$), de la MODALITÉ ($F_{(1,501)} = 34,517$, $p < 0,0001$) et de l'ORDRE ($F_{(1,501)} = 3,974$, $p < 0,0468$) sont significatives.

Les figures 55 à 58 présentent *les proportions des SNL composés d'un nœud lexical*.

Les enfants de CM2 et de 5^{ème} produisent significativement plus ce type de SNL que les 3^{ème} (respectivement, $p = 0,0168$ et $p = 0,0042$). Les deux plus jeunes populations ont des proportions dont la différence n'est pas significative ($p = 0,6319$) (Figure 55). Les individus les plus âgés produisent moins de SNL simples pour produire davantage de SNL plus riches en nœud lexical.

Figure 55 : Proportions de SNL constitués d'un nœud selon le NIVEAU SCOLAIRE (par rapport au nombre total de SNL)



Les individus produisent plus ce type de SNL dans les textes narratifs et oraux que dans les textes expositifs et écrits (Figure 56 et Figure 57).

Figure 56 : Proportions de SNL constitués d'un nœud selon le TYPE DE TEXTE (par rapport au nombre total de SNL)

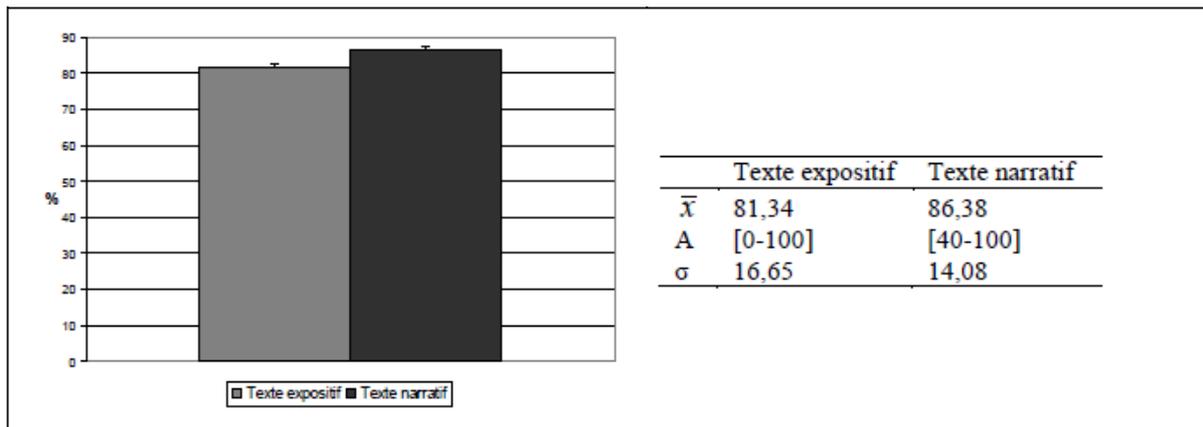
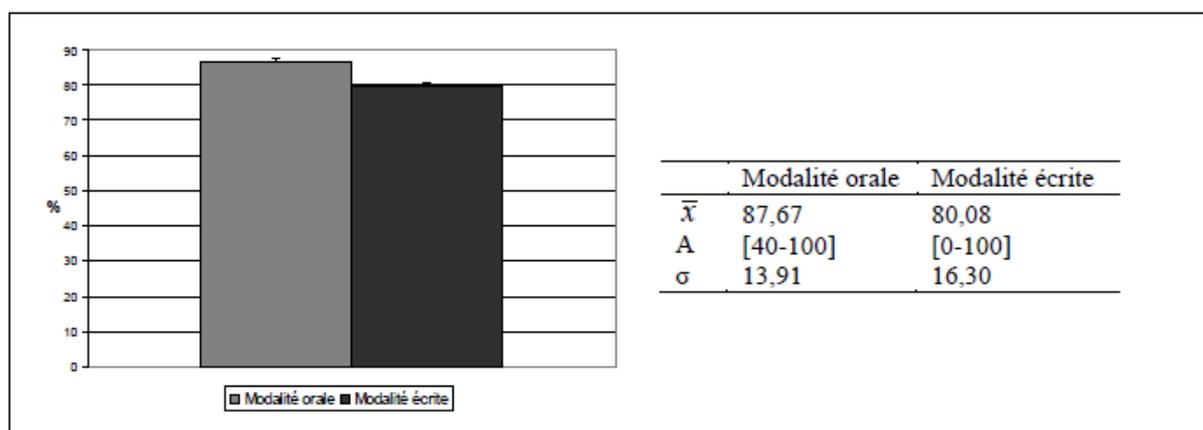
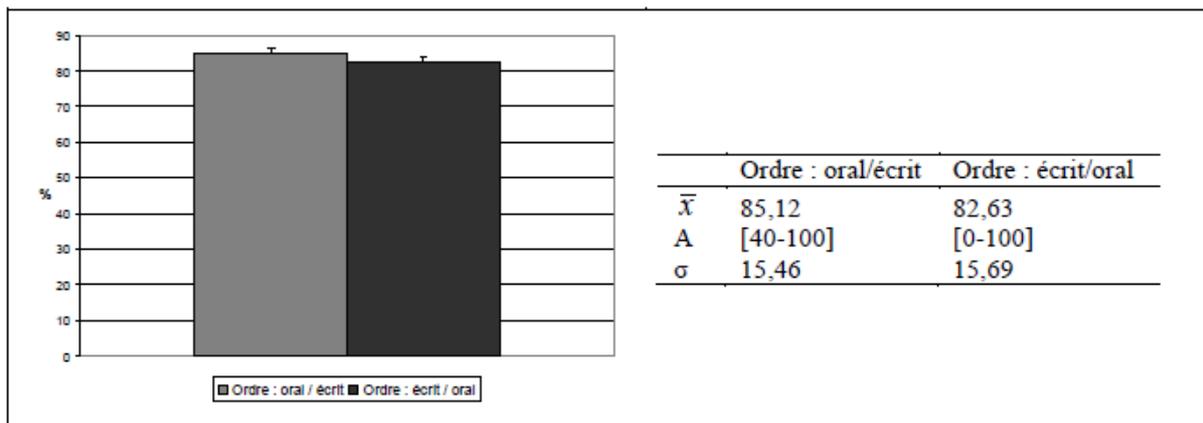


Figure 57 : Proportions de SNL constitués d'un nœud selon la MODALITÉ (par rapport au nombre total de SNL)



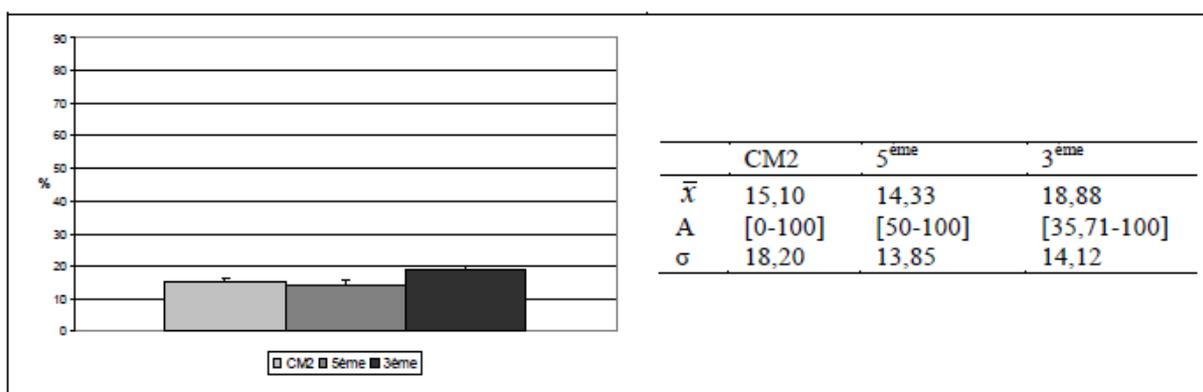
Enfin, les individus appartenant à l'ordre oral/écrit produisent davantage de SNL de ce type que ceux appartenant à l'ordre écrit/oral (Figure 58).

Figure 58 : Proportions de SNL constitués d'un nœud selon l'ORDRE DE PRODUCTION (par rapport au nombre total de SNL)



Les figures 59 à 62 présentent les proportions des SNL composés de plus d'un nœud lexical. Les 3^{ème} produisent significativement plus de SNL complexes que les CM2 ($p = 0,0168$) et que les 5^{ème} ($p = 0,0042$), ces deux dernières populations ont des proportions dont la différence n'est pas significative ($p = 0,6319$) (Figure 59).

Figure 59 : Proportions de SNL constitués de plus d'un nœud selon le NIVEAU SCOLAIRE (par rapport au nombre total de SNL)



Les individus produisent plus ce type de SNL dans les textes expositifs et écrits que dans les textes narratifs et oraux (Figure 60 et Figure 61).

Figure 60 : Proportions de SNL constitués de plus d'un nœud selon le TYPE DE TEXTE (par rapport au nombre total de SNL)

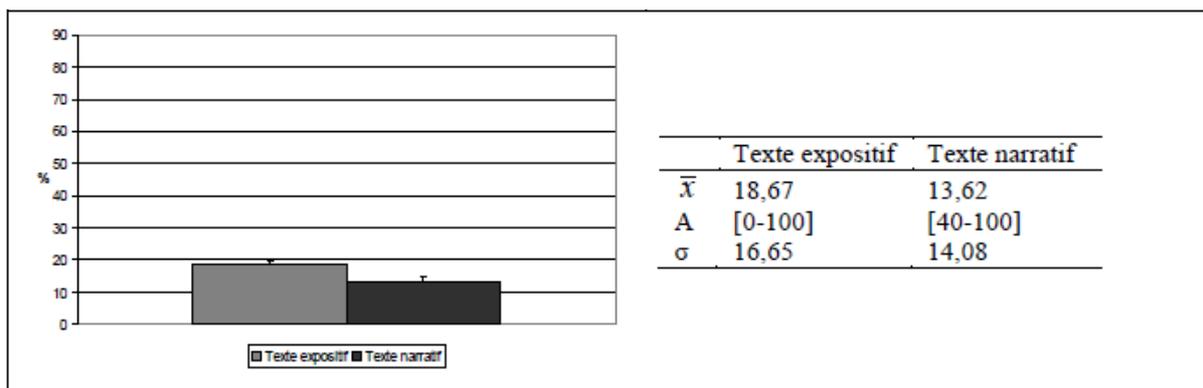
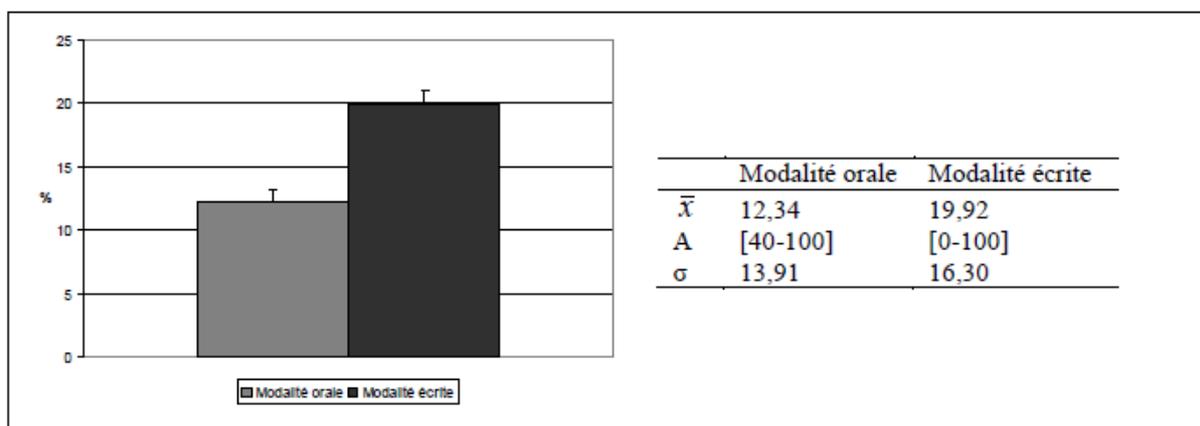
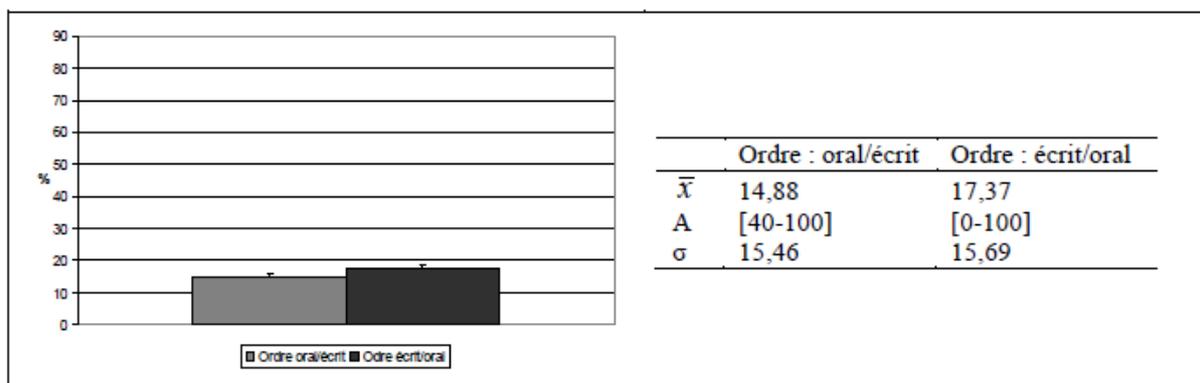


Figure 61 : Proportions de SNL constitués de plus d'un nœud selon la MODALITÉ (par rapport au nombre total de SNL)



Enfin, les participants appartenant à l'ordre écrit/oral produisent davantage de SNL de ce type que ceux appartenant à l'ordre oral/écrit (Figure 62).

Figure 62 : Proportions de SNL constitués de plus d'un nœud selon l'ORDRE DE PRODUCTION (par rapport au nombre total de SNL)



9.1.3. Discussion des résultats de la Section 9.1

Le but de cette partie était de voir si le degré de lexicalité des SNL variait selon les facteurs : NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE.⁹⁵

D'un point de vue général, les individus produisent, en général, en plus grande *proportion des SNL composés d'un seul nœud lexical* ; l'encodage des messages par des *SNL de plus d'un nœud* est plus rare.

Une de nos attentes concernant le degré de lexicalité des SNL était que celui-ci augmentait avec le NIVEAU SCOLAIRE, ce qui irait dans le sens des résultats du Chapitre 8 révélant que *la proportion des SNL* augmentait avec les années de scolarisation. Ceci est confirmé par nos analyses. Les écoliers de CM2 produisent plus *de SNL constitués d'un nœud lexical* que les deux populations de collèges. Ce résultat, à première vue insolite, se comprend du fait que *les SNL de plus d'un nœud*, bien que plus rares et plus complexes, augmentent avec les années de scolarisation. Les adolescents de 3^{ème} produisent moins de *SNL constitués d'un nœud lexical* pour produire davantage de SNL plus complexes (127).

(127) Il provoque aussi de nombreux conflits entre les gens racistes et les gens visés par ce problème. (3^{ème} n°13, expositif écrit, E/O)

L'individu de 3^{ème} produit un SNL complexe de plus de deux nœuds lexicaux (de nombreux conflits entre les gens racistes et les gens visés par ce problème).

L'interprétation concernant *la proportion de SNL constitués de plus d'un nœud* peut être nuancée. En effet, si les CM2 semblent produire autant *de SNL constitués de plus d'un nœud* que les 5^{ème} (Figure 59), la raison en est due à la forte fréquence d'un SNL spécifique dans les productions des CM2 : *les problèmes entre les gens*. Ce SNL était présent dans la consigne du texte expositif. Les CM2 n'ont fait que s'approprier ce SNL. Les individus plus âgés (5^{ème}) produisent des SNL en faisant preuve d'une plus grande flexibilité quant aux choix des lexèmes.

Dans nos attentes et problématiques (Chapitre 3), nous avons mentionné que nous nous attendions également à un effet du contexte de production (MODALITÉ et TYPE DE TEXTE). Plus précisément, les participants de notre étude pourraient produire des *SNL composés de plus d'un nœud lexical* davantage en production expositive et écrite qu'en production narrative et orale.

⁹⁵ Les attentes, hypothèses et problématiques sont davantage explicitées au Chapitre 3 (Section 3.2.2.2, p. 100).

L'effet significatif de la variable TYPE DE TEXTE sur le degré de lexicalité des SNL est confirmé par nos résultats. *Les SNL constitués d'un nœud* sont plus une caractéristique des textes narratifs que des textes expositifs, ces textes se définissant comme ayant plus de SNL complexes (constitués de plus d'un nœud). Ceci va dans le sens de nos prédictions. Les exigences du type de texte expositif sont suivies par les locuteurs/scripteurs ce qui se traduit par plus de SNL plus complexes. Nos sujets, même les plus jeunes, ont la capacité de s'adapter au contexte de production, ce qui est une preuve de leur cheminement vers l'expertise et vers le statut de *proficient speaker* (Berman, 2008). Les exemples ci-dessous, sont deux extraits de deux productions écrites, une narrative (128) et une expositive (129), d'une même adolescente de 3^{ème}.

(128) *Elle l'a mal pris et s'est énervée en employant **des mots vulgaires**. (3^{ème} n°39, narratif écrit, O/E)*

(129) *En effet, un élève attardé ou mal formé subira **des moqueries de ses camarades** et se fera taper dessus. (3^{ème} n°39, expositif écrit, O/E)*

En contexte expositif, la jeune fille produit un SNL complexe alors qu'en contexte narratif elle pronominalise davantage son message et utilise un SNL contenant un nœud lexical.

L'effet de la MODALITÉ est confirmé par les analyses. Les individus produisent davantage de *SNL constitués d'un nœud* à l'oral qu'à l'écrit. En revanche, en production écrite, ces mêmes individus réalisent plus de SNL complexes *composés de plus d'un nœud lexical*. La modalité écrite, de par son temps de planification important et son absence de pression communicative, permet aux locuteurs/scripteurs de produire des formes demandant plus de ressources cognitives (donc des SNL complexes). L'oral, en revanche, exerce une pression communicative telle que les sujets ne peuvent accéder à ce type de forme, d'autant plus que le temps de planification est réduit. Les deux exemples présentés ci-dessous sont des extraits de deux productions expositives, une orale (130) et une écrite (131), d'un adolescent de 3^{ème}.

(130) *Y a **des gens** / i(l) s'aiment pas. (3^{ème} n°12, expositif oral, O/E)*

(131) *On ne peut pas vivre dans la **haine des cultures**. (3^{ème} n°12, expositif écrit, O/E)*

Nous remarquons qu'à l'écrit, l'individu utilise un SNL complexe composé de deux nœuds alors qu'à l'oral, l'utilisation de SNL est moindre.

Pour finir, nous avons également supposé un effet de l'ORDRE DE PRODUCTION sur *la proportion des SNL composés d'un nœud lexical* et celle *des SNL constitués de plus d'un nœud*. Les

résultats dévoilent, en effet, que les locuteurs/scripteurs appartenant au groupe de passation écrit/oral produisent davantage de *SNL constitués de plus d'un nœud* que ceux appartenant à l'ordre de passation oral/écrit, qui produisent davantage de *SNL composés d'un nœud*. Le fait de produire un texte écrit puis un texte oral favorise la production de SNL complexes. En effet, une fois ce type structure activée et appliquée à l'écrit, l'individu le transfère à l'oral. Néanmoins, si le sujet commence par faire un texte oral, il tend à utiliser davantage de SNL simples du fait des contraintes de cette modalité ; ce choix est ensuite maintenu lors de la production écrite. Les exemples ci-dessous illustrent cette différence.

(132) Les **problèmes** entre les **gens** se focalisent particulièrement sur les **scènes** de **racisme**.
(CM2 n°16, expositif écrit, E/O)

(133) Les **gens** se cherchent trop. (CM2 n°17, expositif écrit, O/E)

L'écolier appartenant à l'ordre de passation écrit/oral (132) produit des SNL de deux nœuds alors que le second appartenant à l'ordre de passation oral/écrit (133) en produit un contenant un seul nœud.

Nos résultats vont dans le sens des travaux de Jisa et Mazur (2005) et Mazur (2005) sur des classes socio-professionnelles plus favorisées. En effet, le fait que la production de SNL complexes (à savoir constitués de plus d'un nœud lexical) augmente avec les années de scolarisation, et que les SNL sont plus complexes en production expositive et écrite est l'une des conclusions de ces travaux. Bien que nous ne puissions pas comparer les proportions,⁹⁶ nous pouvons tout de même dire que nos participants ont un comportement guère différent des enfants favorisés.

Dans cette partie, nous avons décidé de nous concentrer sur une des mesures permettant d'observer la complexité linguistique du SNL : le nombre de noms lexicaux. Il a été vu au Chapitre 1 (Section 1.3, p. 40) que la complexité du SNL pouvait également se calculer par rapport aux nombres de modifieurs différents le constituant. Des analyses concernant le nombre de modifieurs différents constituant les SNL ont été effectuées. Les résultats révèlent que les individus, avec les années de scolarisation, produisent de plus en plus de SNL constitués d'au moins deux modifieurs différents. Ces SNL ont une proportion plus importante dans les productions expositives et écrites.⁹⁷

⁹⁶ En effet, dans les travaux sur la population favorisée, les analyses n'étaient pas faites de la même manière.

⁹⁷ Les analyses de variance montrent que les différences observées selon le NIVEAU SCOLAIRE ($F_{(2,501)}=8,090$, $p < 0,0003$), le TYPE DE TEXTE ($F_{(1,501)}=11,749$, $p < 0,0007$) et la MODALITÉ ($F_{(1,501)}=20,266$, $p < 0,0001$) sont significatives. Ces analyses concernent la proportion des SNL constitués d'au moins deux modifieurs différents selon le nombre total de SNL.

Nous pouvons nous demander ce qu'il en est *du degré de lexicalité des SNL* en prenant en considération les positions syntaxiques. Est-ce que les analyses en prenant en considération la fonction syntaxique peuvent compléter les résultats, comme cela a été le cas pour *la proportion des SNL et des pronoms* au Chapitre 8 ? Est-ce que *le degré de lexicalité des SNL* dans les différentes positions varie également selon les variables NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE ? La fonction syntaxique influence-t-elle le degré de lexicalité des SNL ?

Afin de répondre à ces questions, nous allons, dans la prochaine section, présenter les analyses du degré de lexicalité *des SNL en position sujet, objet et autre*.

9.2. Le degré de lexicalité des SNL selon leur position syntaxique

Dans cette partie sont présentés les résultats concernant *les proportions des SNL constitués d'un et de plus d'un nœud lexical* par rapport au nombre total de SNL⁹⁸ et ce, en fonction de leurs positions syntaxiques.

Dans un premier temps, nous rapportons les résultats *des proportions des SNL constitués d'un nœud lexical en position sujet, objet et autre*. Dans un second temps, sont présentés les résultats *des proportions des SNL constitués de plus d'un nœud lexical en position sujet, objet et autre*.

9.2.1. Indicateurs préliminaires

La Figure 63 présente *les proportions des SNL composés d'un nœud et de plus d'un nœud* dans chacune des positions syntaxiques.

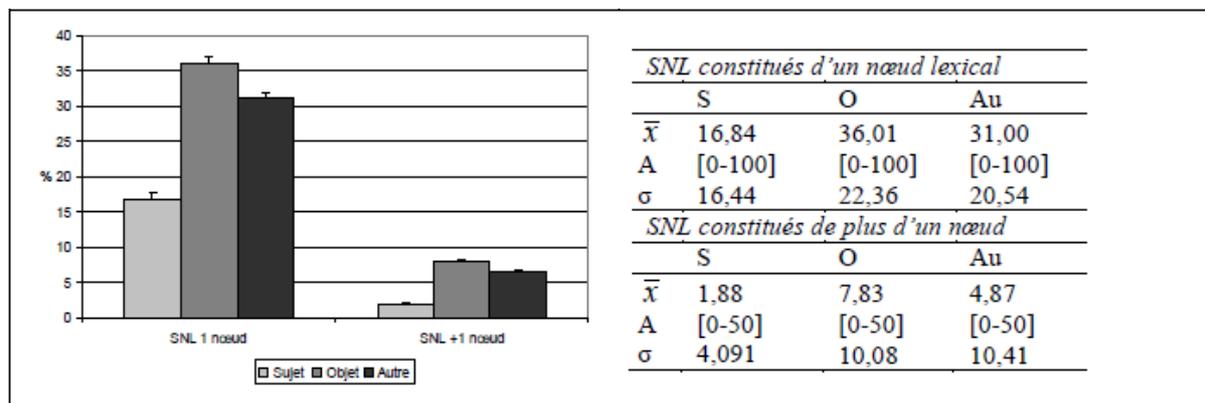
Ainsi, ces analyses révèlent que :

- 16,84% des SNL sont constitués d'un nœud en position sujet ;
- 36,01% des SNL sont constitués d'un nœud en position objet ;
- 31% des SNL sont constitués d'un nœud en position autre ;
- 1,88% des SNL sont constitués de plus d'un nœud position sujet ;

⁹⁸ Pour les 525 textes contenant des SNL sur les 528 textes au total.

- 7,83% des SNL sont constitués de plus d'un nœud en position objet ;
- 4,87% des SNL sont constitués de plus d'un nœud en position autre.

Figure 63 : Proportions des SNL selon leur degré de lexicalité et leur position syntaxique selon le nombre de SNL total



Les SNL d'une façon générale sont plus fréquents en position non sujet, ce qui confirme les résultats antérieurs (Chapitre 8). Les SNL sont également plus complexes en position non sujet.

9.2.2. Les SNL constitués d'un nœud lexical

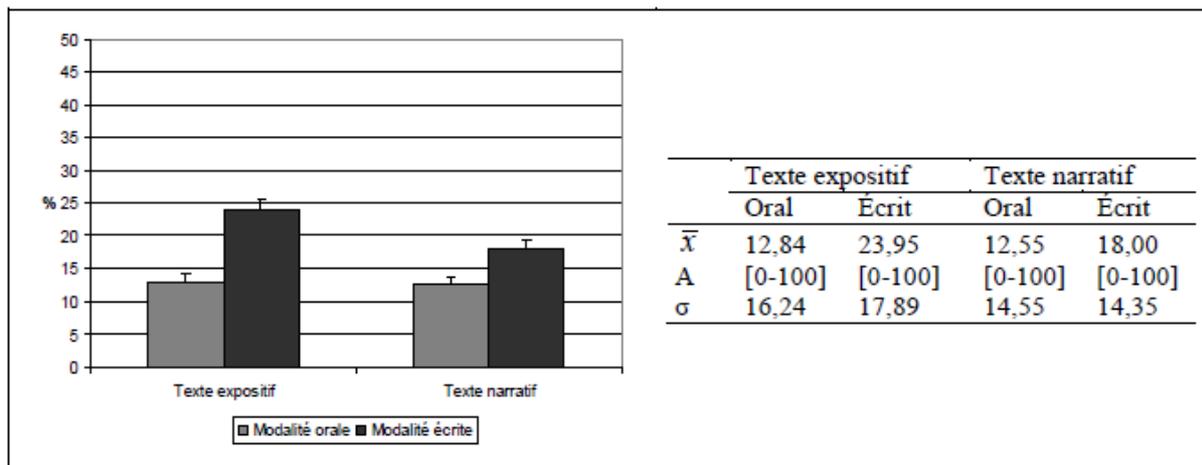
9.2.2.1. Les SNL formés d'un nœud lexical en position sujet

Les figures 64 et 65 présentent *les proportions des SNL composés d'un nœud lexical en position sujet* en fonction du TYPE DE TEXTE, de la MODALITÉ et de l'ORDRE.

Les différences en fonction du TYPE DE TEXTE sont significatives ($F_{(1,501)} = 5,166$, $p < 0,0235$). *Les SNL constitués d'un nœud lexical en position sujet* caractérisent plus les textes expositifs que les textes narratifs (Figure 64).

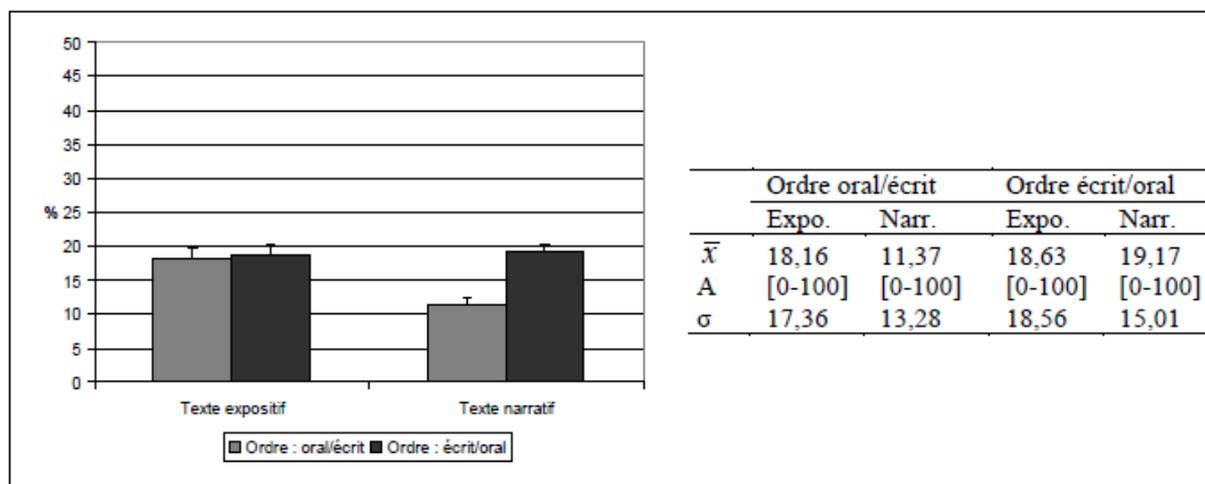
La MODALITÉ a également un effet ($F_{(1,501)} = 36,500$, $p < 0,0001$) : *les SNL composés d'un nœud lexical en position sujet* caractérisent plus l'écrit que l'oral. (Figure 64).

Figure 64 : Proportions des SNL constitués d'un nœud lexical en position sujet selon le TYPE DE TEXTE et la MODALITÉ (par rapport au nombre total de SNL)



Enfin, l'ORDRE fait varier *la proportion des SNL composés d'un nœud lexical en position sujet* ($F_{(1,501)} = 9,147$, $p < 0,0026$). Les individus appartenant à l'ordre écrit/oral utilisent davantage de SNL constitués d'un nœud dans cette position que ceux appartenant à l'ordre inverse (Figure 65).

Figure 65 : Proportions des SNL constitués d'un nœud lexical en position sujet selon le TYPE DE TEXTE et l'ORDRE (par rapport au nombre total de SNL)



La variable NIVEAU SCOLAIRE ne fait pas varier significativement cette VD.

Les résultats ci-dessus révèlent également deux interactions entre les variables TYPE DE TEXTE et MODALITÉ ($F_{(1,503)} = 4,289$, $p < 0,0389$) et entre TYPE DE TEXTE et ORDRE ($F_{(1,503)} = 7,128$, $p < 0,0078$). Plusieurs remarques peuvent être faites. Premièrement, en contexte expositif écrit, *la proportion de SNL constitués composés d'un nœud en position sujet* est la plus importante (versus les trois autres contextes de production : expositif oral, expositif écrit et narratif oral).

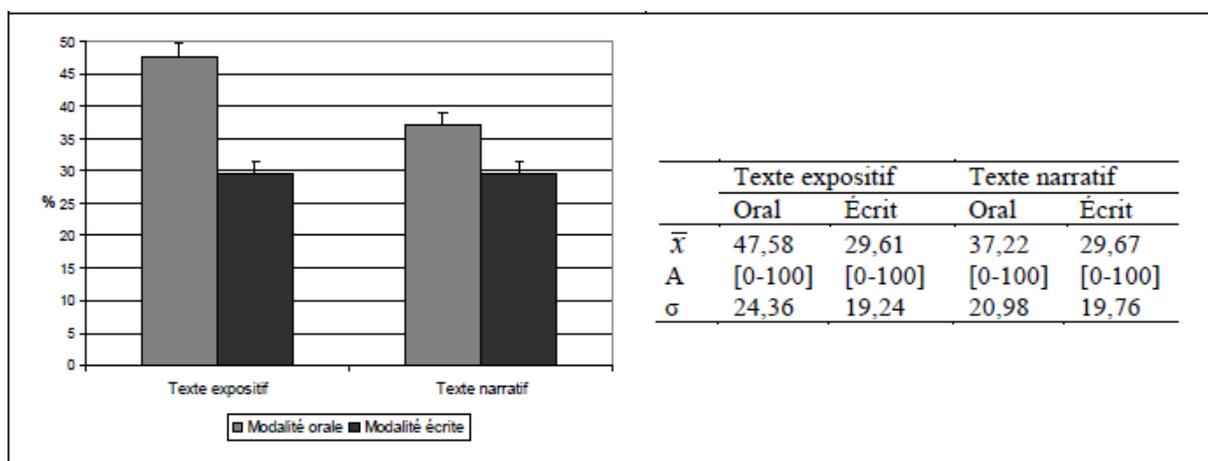
Deuxièmement, en production expositive, les différences entre les deux modalités de production sont plus fortes qu'en production narrative. Troisièmement, en production narrative, les différences entre les ordres de production sont plus fortes qu'en production expositive.

9.2.2.2. Les SNL formés d'un nœud lexical en position objet

Les figures 66 et 67 présentent *les proportions des SNL composés d'un nœud lexical en position objet* en fonction du TYPE DE TEXTE, de la MODALITÉ et de l'ORDRE DE PRODUCTION.

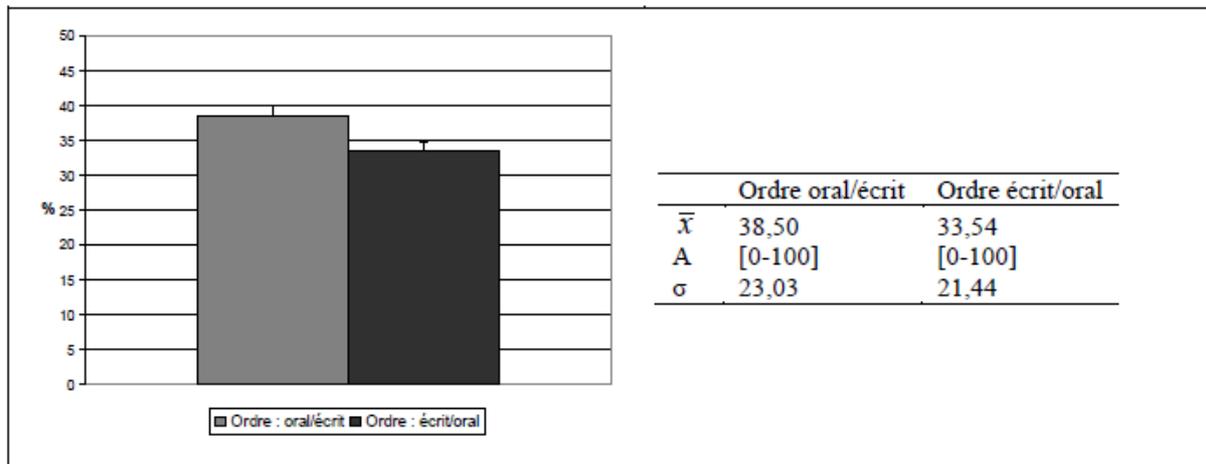
Les différences en fonction du contexte de production sont significatives (TYPE DE TEXTE ($F_{(1,501)} = 8,230$, $p < 0,0043$) et MODALITÉ ($F_{(1,501)} = 48,512$, $p < 0,0001$)). Les individus produisent plus de SNL composés d'un nœud lexical en position objet dans les textes expositifs et oraux que dans les productions narratives et écrites (Figure 66).

Figure 66 : Proportions des SNL constitués d'un nœud lexical en position objet selon le TYPE DE TEXTE et la MODALITÉ (par rapport au nombre total de SNL)



L'ORDRE affecte également *la proportion des SNL composés d'un nœud lexical en position objet* ($F_{(1,501)} = 3,889$, $p < 0,0211$). Les individus appartenant à l'ordre de production oral/écrit produisent davantage de SNL constitués d'un nœud dans cette position que les individus appartenant à l'ordre écrit/oral (Figure 67).

Figure 67 : Proportions des SNL constitués d'un nœud lexical en position objet selon l'ORDRE DE PRODUCTION (par rapport au nombre total de SNL)



La variable NIVEAU SCOLAIRE n'a pas d'effet significatif.

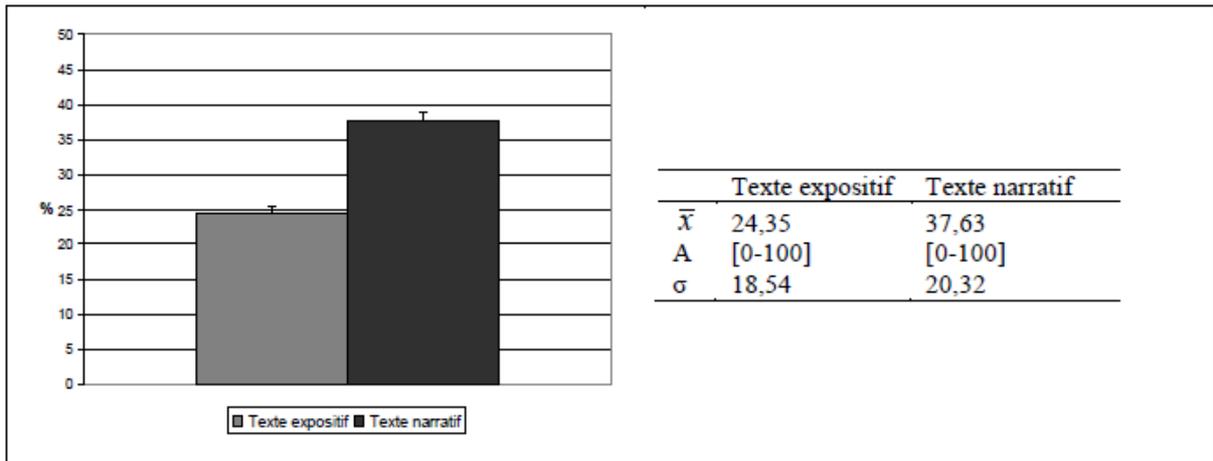
Les résultats révèlent également une interaction entre le TYPE DE TEXTE et la MODALITÉ ($F_{(1,501)} = 8,079$, $p < 0,0047$) (Figure 66). Plusieurs remarques peuvent être faites. Premièrement, en production expositive orale, la proportion de SNL constitués d'un nœud en position objet est la plus élevée (versus les trois autres contextes de production : soit expositif écrit, narratif écrit et narratif oral). Deuxièmement, en contexte narratif, l'effet de la modalité est atténué par rapport au contexte expositif. Troisièmement, en production écrite, l'effet du type de texte est moindre (contrairement à la modalité orale).

9.2.2.3. Les SNL formés d'un nœud lexical en position autre

La Figure 68 présente les proportions des SNL composés d'un nœud lexical en position autre en fonction du TYPE DE TEXTE.

Seules les différences en fonction du TYPE DE TEXTE sont significatives ($F_{(1,501)} = 62,717$, $p < 0,0001$). La proportion des SNL composés d'un nœud lexical en position autre est plus importante dans les textes narratifs que dans les textes expositifs (Figure 68).

Figure 68 : Proportions des SNL constitués d'un nœud lexical en position autre selon le TYPE DE TEXTE (par rapport au nombre total de SNL)



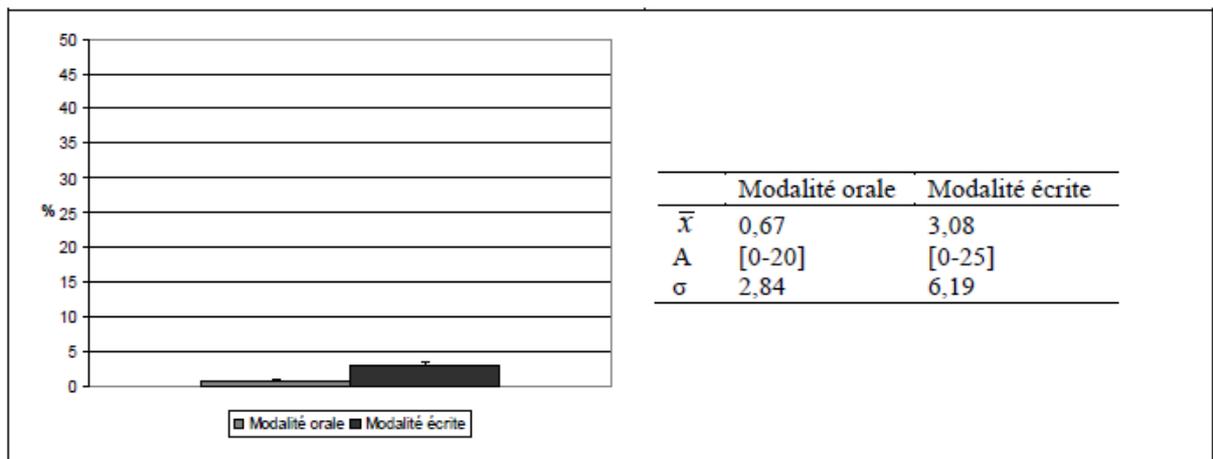
9.2.3. Les SNL constitués de plus d'un nœud lexical

9.2.3.1. Les SNL formés de plus d'un nœud lexical en position sujet

La figure 69 présente les proportions des SNL composés de plus d'un nœud lexical en position sujet en fonction de la MODALITÉ.

Les différences en fonction de la MODALITÉ sont significatives ($F_{(1,501)} = 32,912$, $p < 0,0001$). La proportion des SNL composés de deux nœuds lexicaux en position sujet est plus élevée dans les textes écrits que dans les textes oraux (Figure 69).

Figure 69 : Proportions des SNL constitués de plus d'un nœud lexical en position sujet selon la MODALITÉ (par rapport au nombre total de SNL)



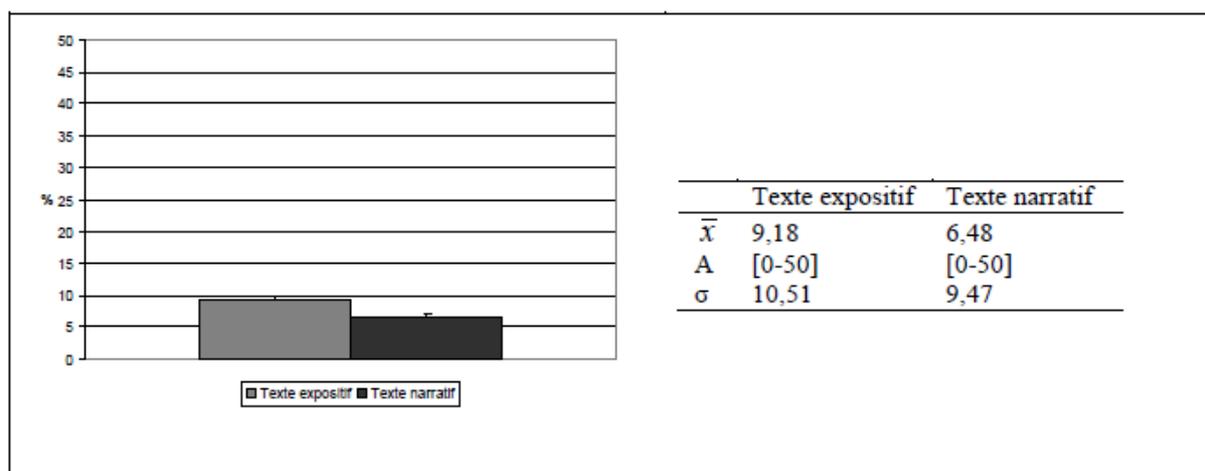
Les autres variables – NIVEAU SCOLAIRE, TYPE DE TEXTE et ORDRE – n’influencent pas la variable dépendante de manière significative.

9.2.3.2. Les SNL formés de plus d’un nœud lexical en position objet

La Figure 70 présente *les proportions des SNL composés de plus d’un nœud lexical en position objet* en fonction du TYPE DE TEXTE et de l’ORDRE DE PRODUCTION.

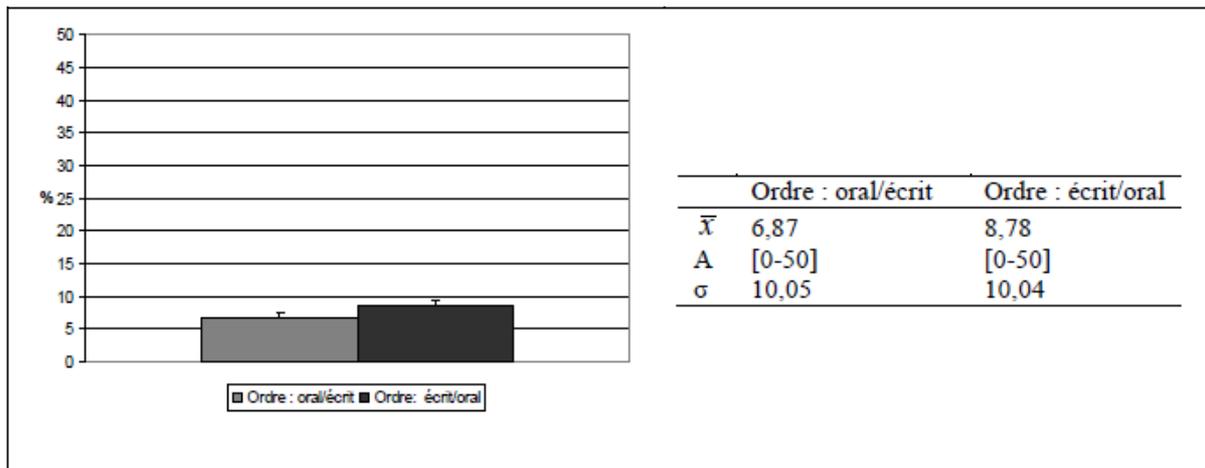
Les différences en fonction du TYPE DE TEXTE sont significatives ($F_{(1,501)} = 9,664, p < 0,0020$). *Les SNL composés de plus d’un nœud lexical en position objet* est plus caractéristique des productions expositives que des productions narratives (Figure 70).

Figure 70 : Proportions des SNL constitués de plus d’un nœud lexical en position objet selon le TYPE DE TEXTE (par rapport au nombre total de SNL)



L’ORDRE fait également varier *la proportion des SNL composés de plus d’un nœud lexical en position objet* ($F_{(1,501)} = 5,012, p < 0,0256$), qui est plus importante dans les textes des individus ayant produit dans l’ordre écrit/oral.

Figure 71 : Proportions des SNL constitués de plus d’un nœud lexical en position objet selon l’ORDRE (par rapport au nombre total de SNL)



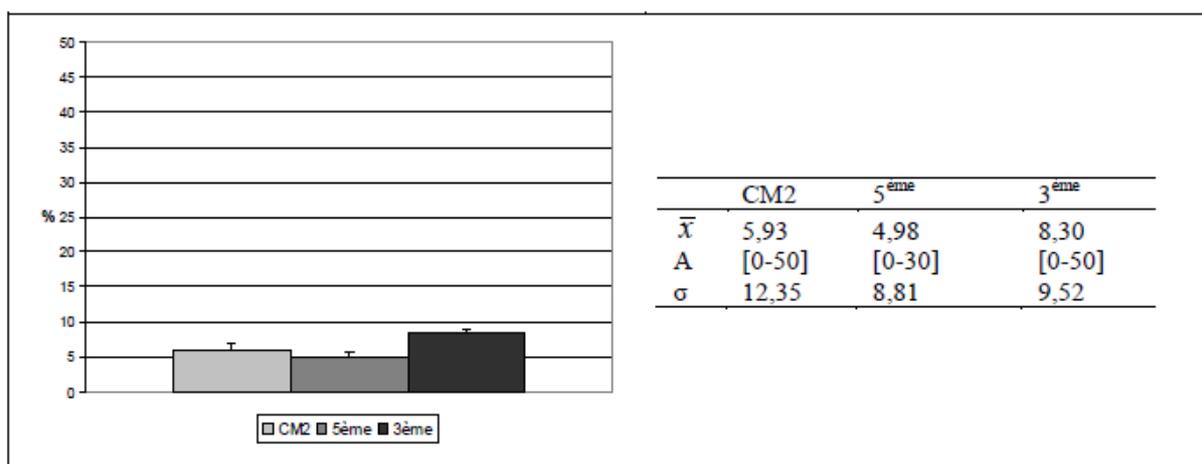
Les deux autres facteurs – NIVEAU SCOLAIRE et MODALITÉ – n’ont pas d’effet significatif.

9.2.3.3. Les SNL formés de plus d’un nœud lexical en position autre

Les figures 72 et 73 présentent *les proportions des SNL composés de plus d’un nœud lexical en position autre* en fonction du NIVEAU SCOLAIRE et de la MODALITÉ.

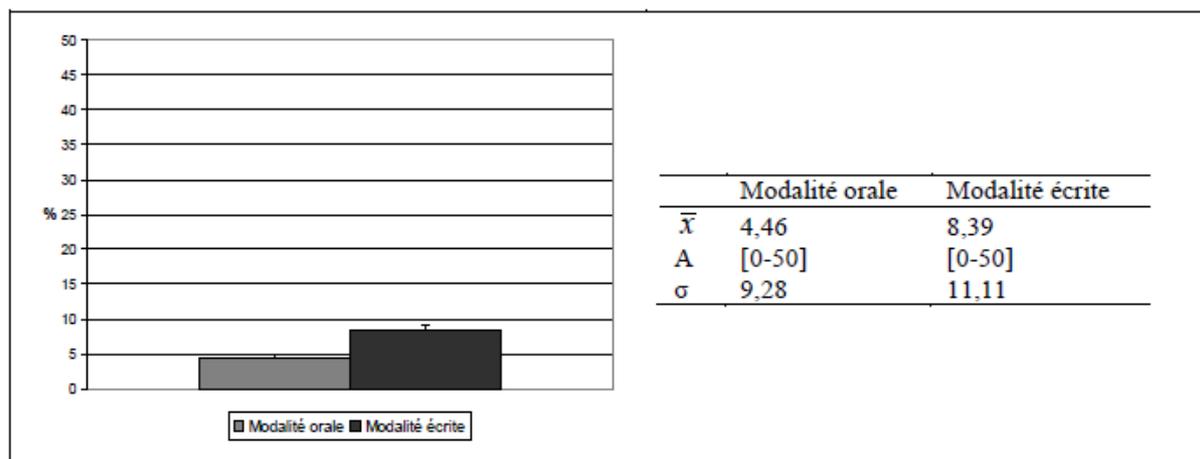
Les différences en fonction du NIVEAU SCOLAIRE sont significatives ($F_{(2,501)} = 4,911, p < 0,0077$). *La proportion des SNL composés de plus d’un nœud lexical en position autre* augmente significativement entre les individus de 5^{ème} et de 3^{ème} ($p = 0,0024$) et entre ceux de CM2 et 3^{ème} ($p = 0,0293$) ; les groupes de CM2 et 5^{ème} ont des proportions relativement proches ($p = 0,3883$) (Figure 72).

Figure 72 : Proportions des SNL constitués de plus d’un nœud lexical en position autre selon le NIVEAU SCOLAIRE (par rapport au nombre total de SNL)



La MODALITÉ a également un effet ($F_{(1,501)} = 19,248, p < 0,0001$). Les individus produisent davantage ces SNL en position autre dans les textes écrits que les textes oraux (Figure 73).

Figure 73 : Proportions des SNL constitués de plus d'un nœud lexical en position autre selon la MODALITÉ (par rapport au nombre total de SNL)



Les variables TYPE DE TEXTE et ORDRE n'ont pas d'effet significatif.

9.2.4. Discussion de la Section 9.2

Une de nos attentes concernant *les proportions des SNL composés d'un nœud et de plus d'un nœud lexical*, étaient que ces dernières varieraient différemment en fonction des facteurs – NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE DE PRODUCTION – si leurs positions syntaxiques sont prises en considération.⁹⁹ Les statistiques incluant les positions syntaxiques (Figure 63 à Figure 73) complètent les tendances observées lors des analyses plus générales (Figure 54 à Figure 62).

Rappelons que *la proportion des SNL constitués d'un nœud* diminue avec les années de scolarisation, et est plus importante dans les textes narratifs et oraux que dans les textes expositifs et écrits. De plus, cette proportion est plus élevée dans les textes des individus appartenant à l'ordre oral/écrit. *La proportion des SNL constitués de plus d'un nœud* augmente avec les années de scolarisation, et est plus importante dans les textes expositifs et écrits que dans les textes narratifs et oraux. De plus, cette proportion est plus élevée dans les textes des individus appartenant à l'ordre écrit/oral.

⁹⁹ Pour des raisons évoquées à la Section 1.3. Les hypothèses et attentes sont davantage explicitées au Chapitre 3 (Section 3.2.2.2, p. 100).

Les analyses selon les positions syntaxiques révèlent que le NIVEAU SCOLAIRE fait varier seulement *la proportion des SNL composés de plus d'un nœud lexical en position autre* : celle-ci augmente avec les années de scolarisation.

Il peut paraître surprenant que le NIVEAU SCOLAIRE ne fasse pas varier les SNL, et notamment ceux constitués de plus d'un nœud en position sujet. Néanmoins, si nous faisons des analyses plus précises, nous remarquons que la proportion des SNL constitués de deux nœuds lexicaux augmente avec l'âge.¹⁰⁰ Au-delà, de deux nœuds, les analyses ne révèlent aucune significativité du NIVEAU SCOLAIRE. Ceci est sans doute lié à leur faible proportion dans les textes.

Les proportions des SNL constitués d'un nœud lexical en position sujet, objet et autre et la proportion des SNL constitués de plus d'un nœud en position objet varient selon le TYPE DE TEXTE. Si les SNL constitués d'un nœud lexical en position autre sont davantage une caractéristique des textes narratifs, les SNL constitués d'un nœud lexical en position sujet et objet et les SNL constitués de plus d'un nœud lexical en position objet caractérisent plus les textes expositifs.

La MODALITÉ affecte différemment la proportion des SNL selon leur degré de lexicalité et leur position syntaxique. *La proportion des SNL constitués d'un nœud lexical en position objet confirme, en étant plus importante dans les textes oraux, la tendance observée par les analyses des SNL d'un nœud toutes positions confondues. Néanmoins, la proportion des SNL constitués d'un nœud lexical en position sujet est plus importante à l'écrit qu'à l'oral tout comme les proportions des SNL constitués de plus d'un nœud en position de sujet et de plus d'un nœud en position autre.*

Produire des SNL complexes et en position sujet se fait relativement plus fréquemment à l'écrit qu'à l'oral. En revanche, produire des SNL simples et en position objet peut se réaliser plus facilement à l'oral. Produire des SNL à l'oral semble être davantage possible, si ces derniers se trouvent en position objet et s'ils sont simplement constitués. Produire un SNL en position sujet est assez lourd : l'enfant doit cumuler une tâche d'accès lexical et une tâche syntaxique. Il doit de plus aller à l'encontre de la structure préférentielle de la langue en produisant du lexique en position sujet (selon les contraintes de la PAS). L'enfant semble

¹⁰⁰ L'analyse de variance montre que les différences observées selon le NIVEAU SCOLAIRE, pour les SNL constitués de deux nœuds, sont significatives : ($F_{(2,501)} = 4,170, p < 0,0160$).

trouver les ressources nécessaires à l'écrit de par son rythme de production plus lent et l'absence de pression communicationnelle.

Pour finir, l'ORDRE DE PRODUCTION affecte *la proportion des SNL composés d'un nœud lexical en position sujet et objet* et celle *des SNL composés de plus d'un nœud lexical en position objet*.

La proportion des SNL constitués d'un nœud lexical en position objet confirme la tendance générale selon laquelle les SNL composés d'un nœud sont plus fréquents dans les productions des individus appartenant à l'ordre oral/écrit. En revanche, *la proportion des SNL constitués d'un nœud lexical en position sujet* ainsi que celle *des SNL composés de plus d'un nœud lexical en position objet* inversent la situation. Ces deux proportions sont plus importantes dans les productions de locuteurs appartenant à l'ordre écrit/oral. Ceci s'explique sans doute par le fait que produire un SNL est assez coûteux cognitivement et, ce, notamment en position sujet ; le fait de produire un texte écrit en premier en facilite donc la production.

9.3. Conclusion du Chapitre 9

Nous pouvons encore une fois souligner l'importance des analyses selon les positions syntaxiques qui révèlent de nouvelles variations. Ainsi, si *la proportion des SNL constitués d'un nœud* est plus importante dans les textes oraux, ceci reste, certes, vrai pour la position objet mais pas pour la position sujet. Les SNL simples en position sujet se retrouvent plus à l'écrit. Il en va de même pour l'ordre : si *la proportion des SNL constitués d'un nœud* est plus importante dans les textes des individus appartenant à l'ordre oral/écrit, ceci reste, certes, vrai pour la position objet mais pas pour la position sujet. Les SNL simples en position sujet ont une proportion plus importante dans les textes des individus ayant produit dans l'ordre écrit/oral.

Nous pouvons conclure en observant d'un peu plus près les statistiques descriptives (Figure 63). Les individus préfèrent produire les SNL dans certaines positions, et ce peu importe leur nombre de nœuds lexicaux. Les individus produisent certes des SNL mais surtout en position objet et autre. La différence de proportion entre les SNL en position objet/autre et les SNL en position sujet est assez conséquente. Cette conclusion va dans le sens des conclusions qui ont été faites lors des chapitres précédents. Les individus de notre étude produisent des SNL mais : (a) ces SNL sont préférentiellement simples ; (b) ces SNL sont préférentiellement produits en position non sujet.

Les Chapitres 8 et 9 se sont penchés sur *la proportion des SNL et des pronoms* ainsi que sur *le degré de lexicalité des SNL*. Les résultats font ressortir qu'il existe un lien assez fort entre formes linguistiques, positions syntaxiques, contexte de production et âge. Avec l'expérience, les enfants produisent de moins en moins de pronoms et de plus en plus de SNL qui sont de plus en plus lourds. Cette remarque générale dépend néanmoins du contexte de production et des positions syntaxiques. Nous pouvons compléter notre vision de la production langagière en mettant en relation les formes linguistiques, les positions syntaxiques avec le statut référentiel. Si la position syntaxique influence le type de forme mobilisé, est-ce que le statut référentiel du référent à encoder est lié également à la position et à la forme ?

Chapitre 10

Le Flux de l'information : Statuts informationnels des syntagmes nominaux

Comme nous l'avons montré dans le Chapitre 1 (Section 1.1, p. 22), pour produire un texte cohérent, le locuteur/scripteur doit jongler entre introduire de la nouvelle information et la maintenir. Une des façons d'y parvenir est l'utilisation des SN. L'individu introduit une thématique, en discute, fait des prédications dessus puis, lorsqu'il pense avoir épuisé le sujet, introduit une nouvelle information, la développe, *etc.* Il ne suffit donc pas d'introduire de nouveaux référents en masse ; il faut également en discuter en les maintenant. Plus les informations sont maintenues, plus le texte est long et développé. Maintenir dans la perspective de développer une thématique est sans doute une acquisition tardive.

Rappelons, que, sous l'expression *ancienne information*, nous intégrons les SNL et les pronoms représentant une information déjà introduite dans le texte ; nous incluons également les réintroductions lexicales. Néanmoins, les réintroductions pronominales, qui sont codées à part (*cf.* Chapitre 11), ne sont pas prises en compte dans la mesure où il s'agit d'un cas non conventionnel allant à l'encontre de la cohérence d'un texte et du traitement naturel des langues. En effet, lorsqu'une information est considérée comme pas assez activée pour être encodée sous forme pronominale, les locuteurs/scripteurs la verbalisent sous forme lexicale.

Sous la catégorie *nouvelle information*¹⁰¹ sont regroupés tous les référents mentionnés explicitement pour la première fois. Ces aspects du codage sont davantage travaillés dans le Chapitre 6 (Section 6.2, p.143).

Dans un premier temps, sont présentés des résultats relatifs à la proportion des SN maintenant et introduisant 134 une nouvelle information toutes positions confondues (Section 10.1, p. 231) et dans un second temps, selon les positions syntaxiques (Section 10.2, p. 240).

(134) *Je me suis embrouillée avec **une meuf** parce qu'**elle** prenait les montages de ma sœur sans lui demander.* (5^{ème} n°28, narratif écrit, O/E)

Dans l'exemple 134, le SNL *une meuf* est introduit puis maintenu avec le pronom *elle*.

10.1. Proportion des syntagmes nominaux introduisant et maintenant une information

10.1.1. Analyses

Cette partie présente *les proportions des SN maintenant et introduisant des informations* selon le nombre total de SN introduisant et maintenant une information en fonction du NIVEAU SCOLAIRE, du TYPE DE TEXTE et de la MODALITÉ.¹⁰² Les différences selon le NIVEAU SCOLAIRE ($F_{(2,503)} = 9,093$, $p < 0,0001$), le TYPE DE TEXTE ($F_{(1,503)} = 33,864$, $p < 0,0001$) et la MODALITÉ ($F_{(1,503)} = 8,982$, $p < 0,0029$) sont significatives. Le facteur ORDRE n'a pas d'effet.

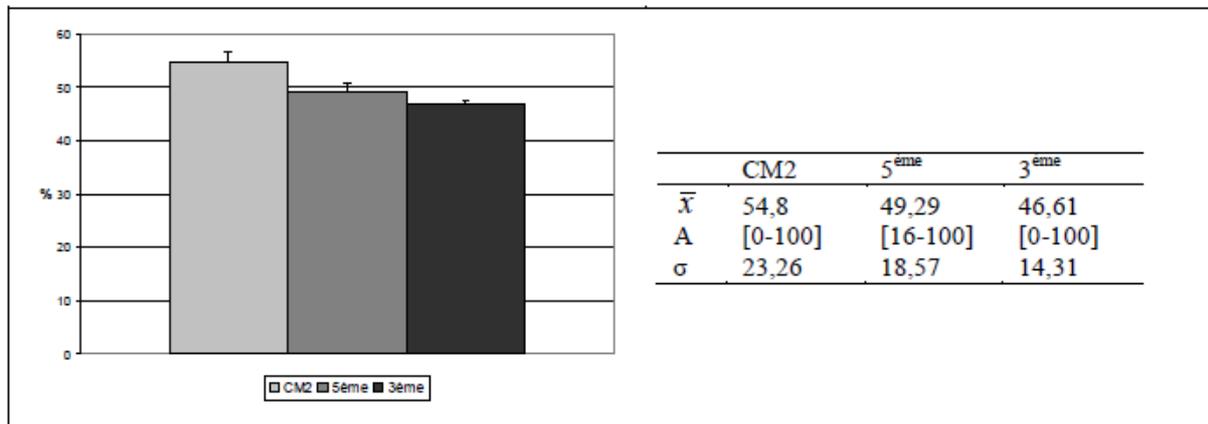
Les figures 74 à 76 présentent *les proportions des SN introduisant une information*.

La proportion des SN introduisant une nouvelle information diminue avec le NIVEAU SCOLAIRE (Figure 74). Les écoliers de CM2 ont une proportion de *SN introduisant une nouvelle information* plus importante que les individus de 5^{ème} ($p = 0,0053$) et de 3^{ème} ($p < 0,0001$). Ces deux dernières populations ont des proportions ne se distinguant pas significativement ($p = 0,1707$). Ce résultat semble assez surprenant. Ce n'est pas que les CM2 introduisent plus de nouvelles informations que les populations collégiennes, mais c'est qu'ils ne maintiennent pas une information introduite pour la développer ensuite. Cette explication, en ayant recours au nombre de *SN introduisant et maintenant une information par texte*, est davantage développée dans la partie discussion (Section 10.1.2, p. 235).

¹⁰¹ Cette catégorie contient que des formes lexicales alors que la catégorie des SN maintenant une information contient des SNL et des pronoms.

¹⁰² Pour les 526 textes contenant ce type de SN sur les 528 au total.

Figure 74 : Proportions des SN introduisant une nouvelle information selon le NIVEAU SCOLAIRE (par rapport au nombre total de SN introduisant et maintenant une information)



De plus, la proportion des SN introduisant une nouvelle information est plus importante dans les textes expositifs et écrits que des textes narratifs et oraux (Figure 75 et Figure 76).

Figure 75 : Proportions des SN introduisant une nouvelle information selon le TYPE DE TEXTE (par rapport au nombre total de SN introduisant et maintenant une information)

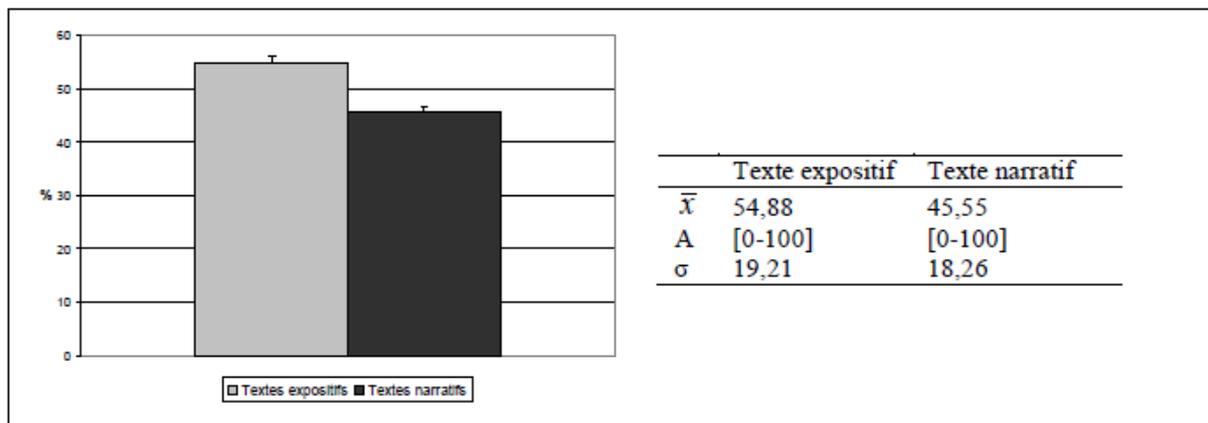
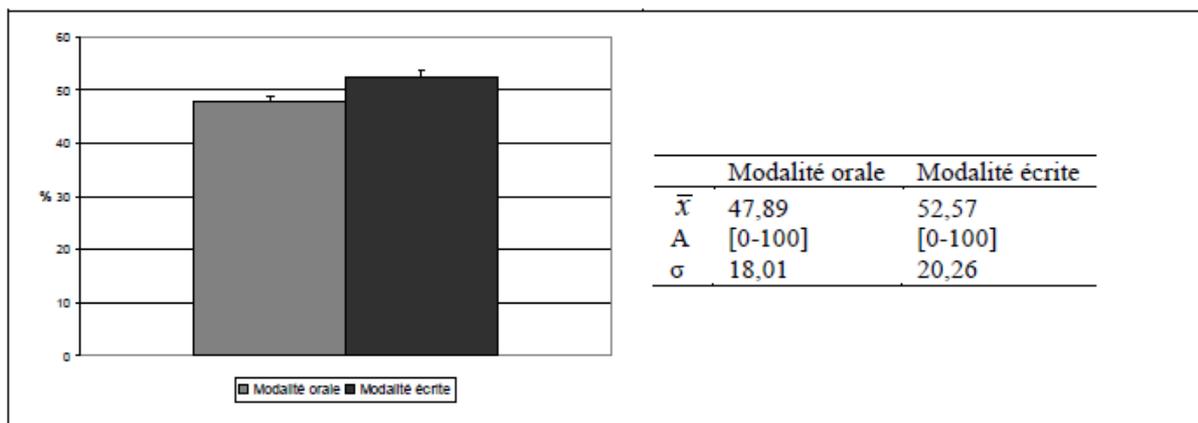


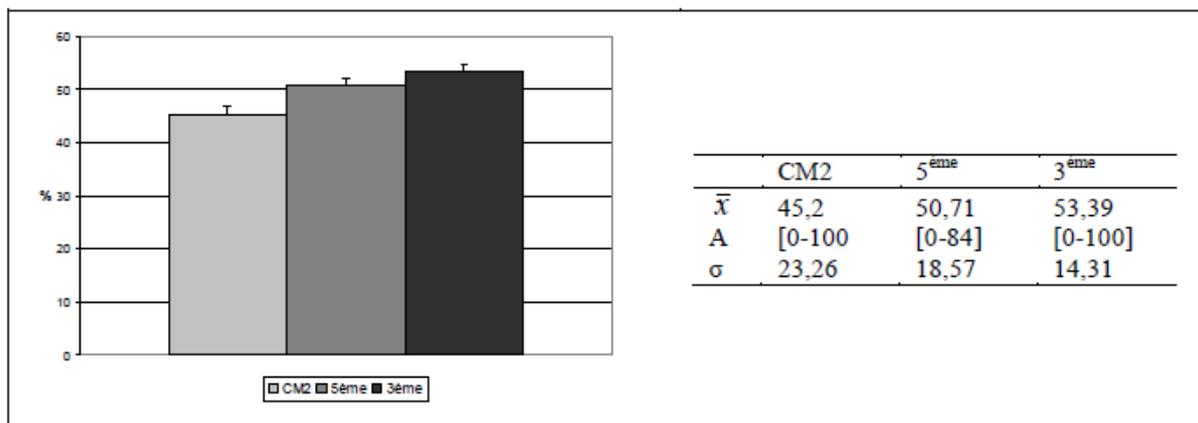
Figure 76 : Proportions des SN introduisant une nouvelle information selon la MODALITÉ (par rapport au nombre total de SN introduisant et maintenant une information)



Les figures 77 à 79 présentent *les proportions des SN maintenant une information*.

La proportion des SN maintenant une information augmente avec le NIVEAU SCOLAIRE. Les écoliers (CM2) ont une proportion de *SN maintenant une information* significativement moins importante que les individus de 5^{ème} ($p = 0,0053$) et ceux de 3^{ème} ($p < 0,0001$). Néanmoins ces deux populations de collège ont des proportions ne se distinguant pas significativement ($p = 0,1707$) (Figure 77).

Figure 77 : Proportions des SN maintenant l'ancienne information selon le NIVEAU SCOLAIRE (par rapport au nombre total de SN introduisant et maintenant une information)



De plus, *la proportion des SN maintenant une information* est plus élevée dans les textes narratifs et oraux que des textes expositifs et écrits (Figure 78 et Figure 79).

Figure 78 : Proportions des SN maintenant l'ancienne information selon le TYPE DE TEXTE (par rapport au nombre total de SN introduisant et maintenant une information)

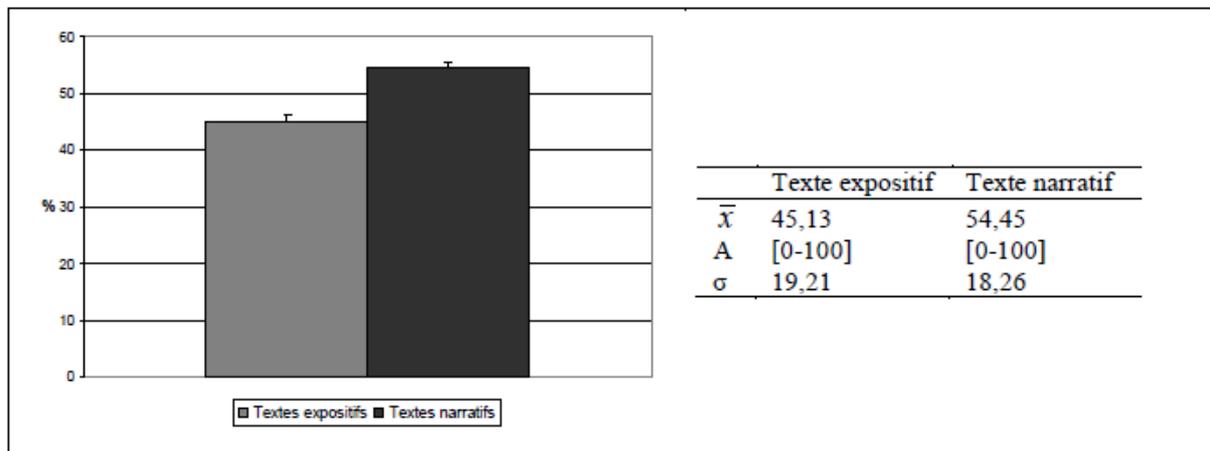
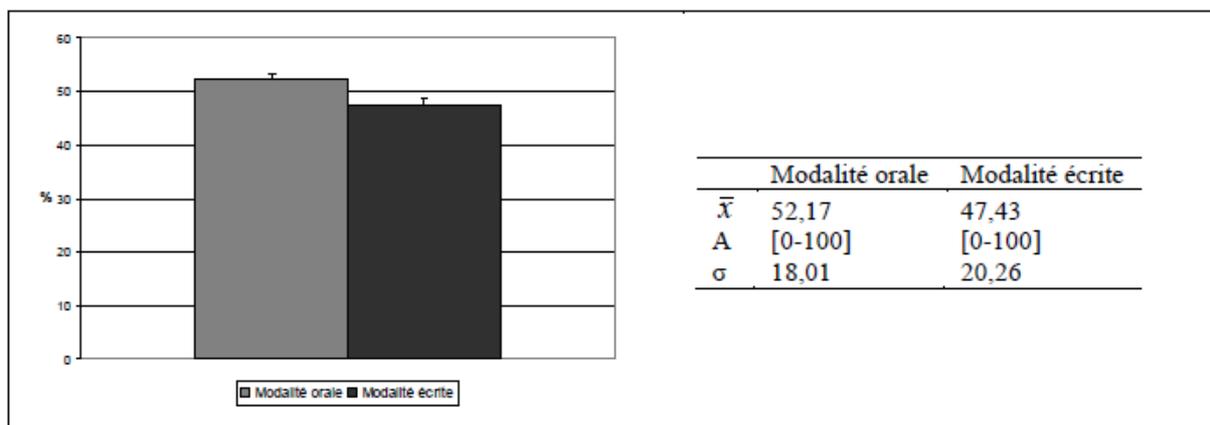
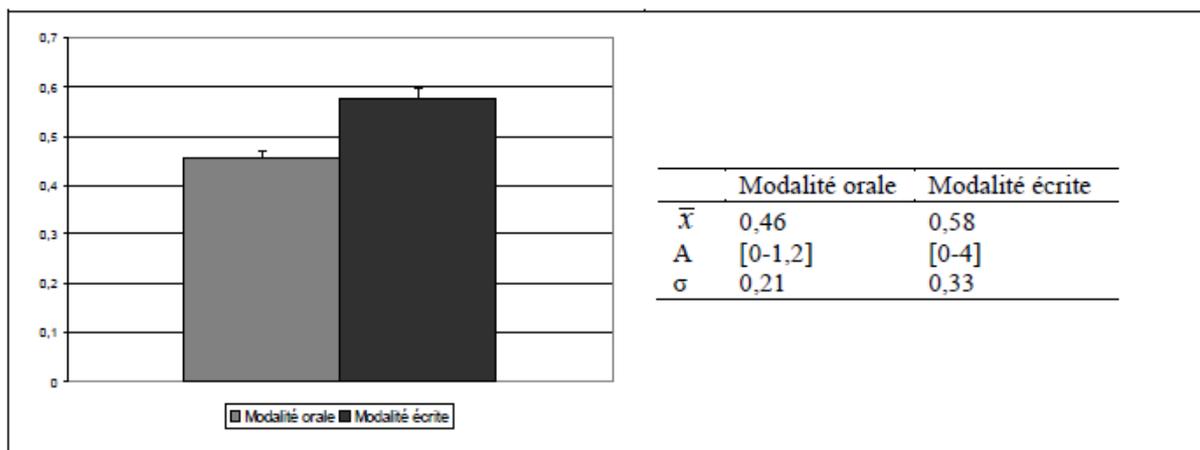


Figure 79 : Proportions des SN maintenant l'ancienne information selon MODALITÉ (par rapport au nombre total de SN introduisant et maintenant une information)



La Figure 80 présente *le nombre moyen de SNL introduisant une information par clause*. Seules les différences selon la MODALITÉ sont significatives ($F_{(1,504)} = 24,469$, $p < 0,0001$) ; les autres facteurs – NIVEAU SCOLAIRE, TYPE DE TEXTE et ORDRE – n'interviennent pas. *Le nombre moyen de SNL introduisant une nouvelle information par clause* est plus élevé dans les textes écrits que dans les textes oraux (Figure 80).

Figure 80 : Nombre moyen de SNL introduisant une nouvelle information par clause selon la MODALITÉ



10.1.2. Discussion de la Section 10.1

Dans ce chapitre, nous voulions voir si les facteurs de notre étude –NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE – faisaient varier *la proportion des SN introduisant une nouvelle information et des SN maintenant une information*.¹⁰³

Une de nos attentes et prédictions était un effet significatif du NIVEAU SCOLAIRE sur ces VD. Nous pensions que l'introduction de nouveaux référents, qui est considérée comme exigeant plus de ressources cognitives que l'ancienne information (Chafe, 1994 ; Khorounjaia *et al.*, 2004), augmentait avec les années de scolarisation, d'autant plus qu'introduire une nouvelle information implique l'utilisation de lexique. Et, comme nous l'avons déjà mentionné, la mobilisation et la production de SNL augmentent les années de scolarisation. Ainsi, la proportion de SN introduisant une information augmenterait avec la longueur des textes et donc avec les années de scolarisation. Cependant, même si l'introduction d'information dans un texte est un aspect essentiel, un autre principe guide le développement d'un texte : le maintien des informations introduites (Charolles, 1978 ; Reinhart, 1980 ; Schneuwly, 1988). Introduire de l'information sans la maintenir n'est pas satisfaisant : le locuteurs/scripteur doit, en effet, également la maintenir pour la développer. Ainsi, nous supposons également que *la proportion de SN maintenant une ancienne information* est plus importante chez les locuteurs plus experts, reflétant qu'ils développent leurs idées.

L'analyse des données confirme un effet significatif du NIVEAU SCOLAIRE sur *la proportion de SN introduisant et maintenant une information*. *La proportion de SN introduisant une nouvelle information* est moins importante dans les textes des individus les plus âgés que dans

¹⁰³ Les attentes, hypothèses et problématiques sont davantage explicitées au Chapitre 3 (Section 3.2.2.3, p.102).

les textes des CM2. En revanche, *la proportion de SN maintenant une information* est plus élevée dans les textes des individus les plus âgés que dans les textes des CM2. Les individus les plus avancés dans le cursus scolaire introduisent de nouvelles informations puis les développent, leurs textes sont plus longs, les idées plus travaillées. Sur la totalité des SN, il y en a davantage qui servent à maintenir qu'à introduire ; ceci se confirme avec l'avancée dans le cursus scolaire. Nous pouvons nuancer notre propos en regardant le nombre moyen de SN introduisant une nouvelle information et le nombre moyen de SN maintenant une information par texte. Ces analyses révèlent que les participants de notre étude maintiennent de plus en plus avec le NIVEAU SCOLAIRE mais qu'ils introduisent également de plus en plus de nouvelle information (Tableau 10).

[Tableau 10 : Nombre moyen de SN introduisant une nouvelle information et maintenant une information par texte]

Tableau 10 : Nombre moyen de SN introduisant une nouvelle information et maintenant une information par texte

	CM2	5 ^{ème}	3 ^{ème}
Nouvelle information ¹⁰⁴	5,03	8,93	11,89
Ancienne information ¹⁰⁵	5,35	10,99	14,67

Ces résultats montrent alors que les textes des 5^{ème} et des 3^{ème} sont de plus en plus riches informationnellement et qu'ils maintiennent également de plus en plus leurs idées en les développant. De plus, nous voyons que, pour les CM2, *le nombre moyen de SN introduisant une nouvelle information* et celui des *SN maintenant une information par texte* sont quasi-similaires. En revanche, pour les 5^{ème}, *le nombre moyen de SN introduisant une nouvelle information* est inférieur à celui des *SN maintenant une information par texte* et cette différence s'accroît chez les 3^{ème}. Les collégiens introduisent de plus en plus de nouveaux référents dans leur production, mais ils assurent également plus la continuité référentielle de leur texte.

Le maintien de l'information n'est pas si évident. Insérer de nouveaux référents dans une production est certes une des règles assurant la cohésion d'un texte (Clark et Haviland, 1974,

¹⁰⁴ Les différences entre les populations sont significatives ($F_{(2,504)} = 55,659$, $p < 0,0001$). Les individus de CM2 produisent significativement moins de *SN introduisant une nouvelle information* que les 5^{ème} ($p = 0,0001$), ces deux groupes en produisent moins que les 3^{ème} ($p = 0,0001$).

¹⁰⁵ Les différences entre les populations sont significatives ($F_{(2,504)} = 46,682$, $p < 0,0001$). Les individus de CM2 produisent significativement moins de *SN maintenant une information* que les 5^{ème} ($p = 0,0001$), ces deux groupes en produisent moins que les 3^{ème} ($p = 0,0001$).

1977 ; Khorounjaia et Tolchinsky, 2004), mais une autre règle est la continuité référentielle, la progression. Si l'analyse qualitative des productions montre que les plus jeunes tendent à énumérer de nouvelles informations sans pour autant beaucoup développer les thématiques (135).

*(135) **Le racisme** est par exemple des gens qui n'aiment pas certaines personnes car elles ne sont pas comme eux. **Le vol** est attendre qu'une personne sorte pour lui voler son goûter ou quelque chose de ce genre. **La bagarre** est quelqu'un qui est énervé par quelqu'un et qui veut se battre. (CM2 n°24, expositif écrit, E/O)*

Dans cet exemple, nous voyons que l'élève introduit trois informations principales sans les maintenir et les discuter davantage. Son texte est court et n'est pas développé. Les collégiens font des textes beaucoup plus longs, développent les idées introduites, et ceci se voit à travers le maintien de l'information (136).

*(136) **Les bagarres** commencent par des regards bizarres entre **les personnes concernées** en général. Puis, **elles** se fixent jusqu'à se dire « qu'est-ce que tu chaab ? ». Ensuite, **ils** se rapprochent et **Ø** s'expliquent en s'insultant. Puis tout le monde vient pour **les** chauffer jusqu'à qu'**ils** en viennent aux mains. Pendant **la bagarre**, les gens autour savotent celui qu'ils n'aiment pas et les personnes intelligentes essayent de les séparer. **Les bagarres** sont aussi dues à la race, la nationalité et l'origine (Arabes et Arméniens se bagarrent beaucoup). Le collègue punit ce comportement mais les parents, eux, disent que les pelos ont eu raison. (3^{ème} n°20, expositif écrit, E/O)*

Une autre de nos attentes et prédictions était l'effet significatif du contexte de production (TYPE DE TEXTE et MODALITÉ) SUR ces deux variables – *la proportion des SN introduisant une nouvelle information*, qui serait plus importante dans les textes expositifs et écrits et *la proportion des SN maintenant une information*, qui serait plus élevée dans les textes narratifs et oraux.

Les résultats confirment un effet significatif du TYPE DE TEXTE : *la proportion des SN maintenant une information* est plus importante dans les textes narratifs alors que celle *des SN introduisant une nouvelle information* est plus élevée dans les textes expositifs. Ce résultat n'est guère surprenant si nous nous fions aux définitions de ces types de textes. Le texte expositif impose aux locuteurs/scripteurs une forte pression informationnelle. Les individus répondent à cette contrainte en ayant une proportion des SN introduisant une nouvelle

information importante (137). Le texte narratif n'impose pas une telle pression informationnelle mais au contraire appelle à un maintien des référents principaux (138).

(137) Dans un collège les sujets tels que la triche, les mises à l' écart (envers certaines personnes), les bastons, etc., sont plutôt fréquents. (3^{ème} n°06, expositif écrit, O/E)

*(138) Je connu **un garçon qui**, dès qu' il parlait ou dès qu' il arrivait au collège, **Ø** souffrait et **Ø** subissait des moqueries incessantes et blessantes. (3^{ème} n°06, narratif écrit, O/E)*

Dans ces extraits, le locuteur/scripteur introduit davantage de SN avec une nouvelle information dans son texte expositif que dans son texte narratif. Dans sa narration, une fois introduit le référent principal est maintenu.

Les analyses confirment également un effet significatif de la MODALITÉ DE PRODUCTION et vont dans le sens de nos attentes. À l' oral, les individus ont une proportion plus importante de *SN maintenant une information* qu' à l' écrit, modalité dans laquelle, à l' inverse, ils ont une proportion plus élevée de *SN introduisant une nouvelle information*. De plus, *le nombre moyen de SNL introduisant une nouvelle information par clause* est plus élevé dans les textes écrits (139) que dans les textes oraux (140). Ce résultat peut être expliqué si nous nous en référons à Chafe (1994) qui propose qu' introduire un nouveau référent est sans doute plus coûteux que d' en maintenir un déjà connu. Les individus doivent activer le nouveau référent et de plus l' introduire sous un SNL pour assurer la bonne compréhension du destinataire ; nous avons vu, par ailleurs, au chapitre précédemment que la production de SNL caractérise davantage les textes écrits que les textes oraux. Ainsi, les locuteurs/scripteurs introduisent davantage de nouveaux éléments dans leur production écrite dans la mesure où la modalité écrite offre aux scripteurs un temps de planification facilitant cette tâche.

Pour illustrer ces différences entre les deux modalités, les exemples suivants, qui sont deux extraits de deux productions expositives, une écrite (139) et une orale (140), d' un même individu, sont proposés.

*(139) **Le recopiage pendant un contrôle peut poser beaucoup de problèmes.** (5^{ème} 17, expositif écrit, O/E)*

(140) On n' a pas à copier. (5^{ème} 17, expositif oral, O/E)

Pour encoder la même idée, ce même individu, en production écrite, préfère utiliser des SNL introduisant une nouvelle information alors qu' à l' oral il tend à pronominaliser son message.

(141) Pendant un **cours de maths** lors d'un **contrôle de géométrie**, le **professeur** passait dans les **rangs**. (5^{ème} n°25, expositif écrit, O/E)

(142) euh et ben y a des gens / euh i(ls) trichaient en classe / euh c'est / quand l(e) prof i(l) passait derrière eux. (5^{ème} n°25, expositif oral, O/E)

Les exemples (141) et (142) sont également deux extraits d'une production expositive, une écrite (141) et une orale (142), d'un même individu de 5^{ème}. À l'écrit, dans une seule clause, plus d'un SNL introduisant une nouvelle information sont comptabilisés (*un cours de maths, lors d'un contrôle de géométrie, le professeur, dans les rangs*) ; à la différence de l'oral où un seul SNL introduisant une nouvelle information (*le prof*) est utilisé par clause.

Le résultat concernant *le nombre moyen de SNL introduisant une nouvelle information par clause* engendre une réflexion liée à la quatrième contrainte de la PAS (Du Bois, 1987, 2003) selon laquelle une seule nouvelle information par clause est énoncée.¹⁰⁶ Nos résultats coïncident certes avec cette idée, cependant cette contrainte, reflétant la transmission de l'information à l'oral, est moins caractéristique de l'écrit. Dans les productions écrites, le *nombre moyen de SNL nouvelle information par clause* est supérieur à celui de la modalité orale.

Ces analyses nous révèlent d'intéressants faits. Mais qu'en est-il si nous observons *la proportion des SN maintenant et introduisant une information* en prenant en considération les positions syntaxiques ? Est-ce que les analyses en prenant en considération la fonction syntaxique peuvent compléter les résultats, comme cela a été le cas pour *la proportion des SNL et des pronoms* au Chapitre 8 ainsi que pour le degré de lexicalité des SNL au Chapitre 9 ? Est-ce que *la proportion des SN maintenant et introduisant une information* dans les différentes positions varie également selon les variables NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE ? Afin de répondre à ces questions, nous allons, dans la prochaine section, présenter les analyses des *proportions des SN maintenant et introduisant une information en position sujet, objet et autre*.

¹⁰⁶ Cette réflexion est à mettre également en lien avec la notion de *the one new idea constraint* de Chafe (1994).

10.2. Proportion des syntagmes nominaux maintenant et introduisant une information selon les positions syntaxiques

Dans cette partie est présentée *la proportion des SN maintenant une information* ainsi que *celle des SN introduisant une information* selon les positions syntaxiques. Ces proportions sont toutes calculées par rapport au nombre total de SN introduisant et maintenant une information.¹⁰⁷ Par exemple, *la proportion de SN maintenant une information en position sujet* selon le nombre total de SN introduisant et maintenant une information est observée.

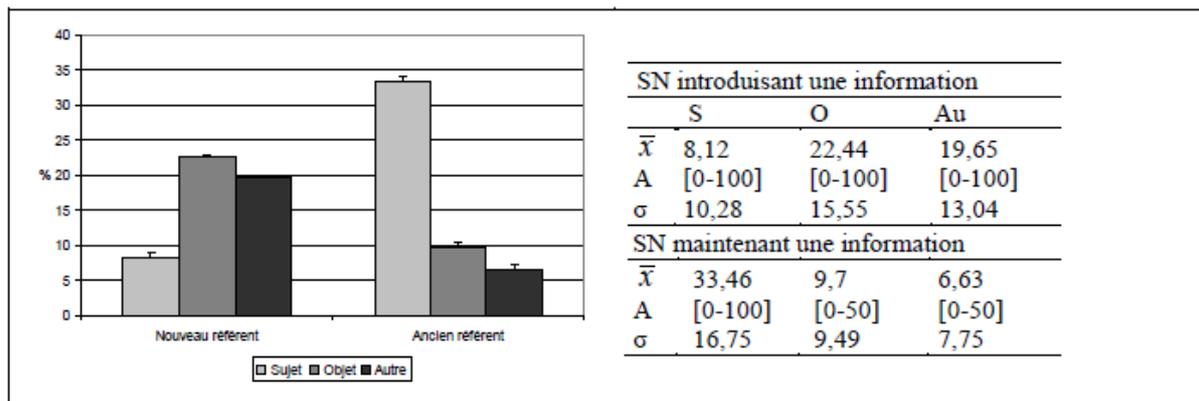
10.2.1. Indicateurs préliminaires

Les informations regroupées en Figure 81 résument *les proportions des SN maintenant et introduisant une information* selon les positions syntaxiques. Ainsi, ces analyses révèlent que :

- 8,12% des SN introduisent une nouvelle information en position sujet ;
- 22,44% des SN introduisent une nouvelle information en position objet ;
- 19,95% des SN introduisent une nouvelle information en position autre ;
- 33,46% des SN maintiennent une information en position sujet ;
- 9,7% des SNL maintiennent une information en position objet ;
- 3,63% des SNL maintiennent une information en position autre.

Figure 81 : Proportions des SN maintenant et introduisant une information selon la position syntaxique (par rapport au nombre total de SN introduisant et maintenant une information)

¹⁰⁷ Pour les 527 textes contenant des SN introduisant et maintenant une information.



Ces résultats suggèrent que l'ancienne information est davantage en position sujet qu'en position objet et autre ; la tendance s'inverse lorsque nous observons les proportions de la nouvelle information qui apparaît davantage en position non sujet qu'en position sujet.

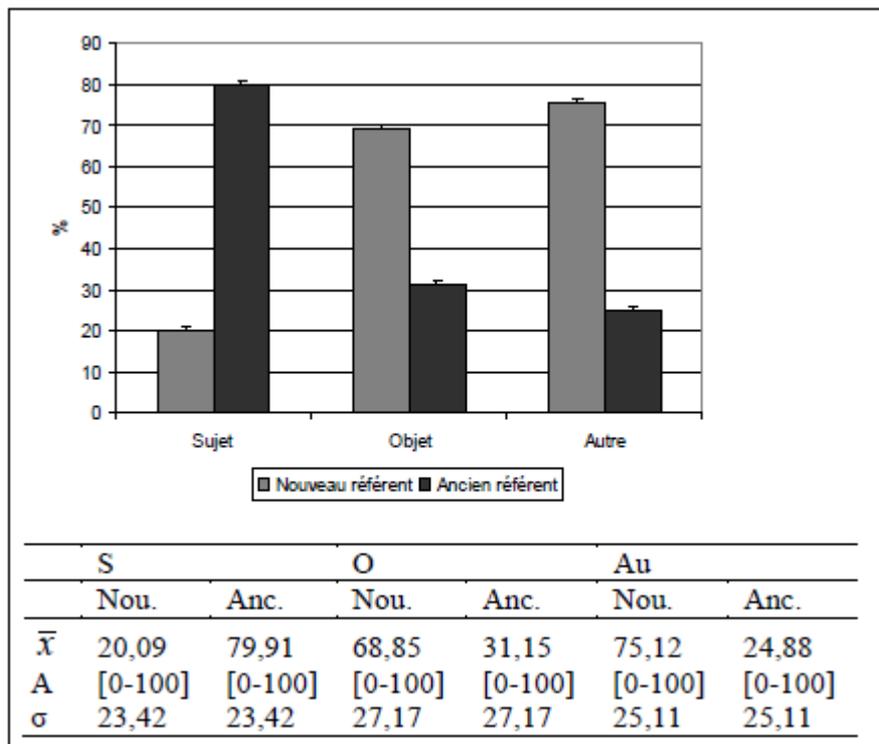
Nous confirmons ces tendances par des tests plus spécifiques. Nous présentons :

- la proportion des SN maintenant et introduisant une information en position sujet selon le nombre total de SN maintenant et introduisant une information en sujet ;
- la proportion des SN maintenant et introduisant une information en position objet selon le nombre total de SN maintenant et introduisant une information en position objet ;
- et enfin la proportion des SN maintenant et introduisant une information en position autre selon le nombre de SN total en position autre.

Les données présentées en Figure 82 récapitulent *les proportions des SN maintenant et introduisant une information* selon les positions syntaxiques. Ainsi :

- 79,91% des SN en position sujet (maintenant ou introduisant une information) maintiennent une information alors que 20,09% d'entre eux introduisent une information (Test T : $t_{(505)} = 28,732$, $p < 0,0001$) ;
- 31,15% des SN en position objet (maintenant ou introduisant une information) maintiennent une information alors que 68,85% d'entre eux introduisent une information (Test T : $t_{(509)} = -15,669$, $p < 0,0001$) ;
- 24,88% des SN en position autre (maintenant ou introduisant une information) maintiennent une information alors que 75,12% d'entre eux introduisent une information (Test T : $t_{(487)} = -22,104$, $p < 0,0001$).

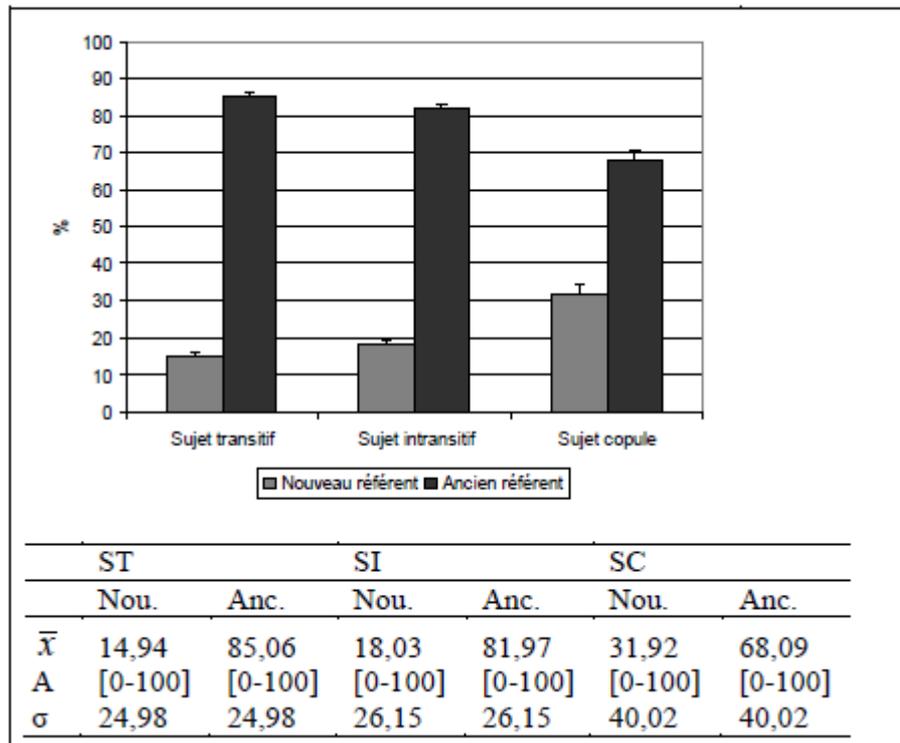
Figure 82 : Proportions des SN maintenant et introduisant selon les positions sujets



Ces données confirment que les positions sujet sont remplies davantage par des SN maintenant une information que par des SN introduisant une nouvelle information. La tendance s'inverse pour les positions non sujet qui sont préférentiellement remplies par des SN introduisant une nouvelle information que par des SN maintenant des référents.

Cette tendance, à avoir une position sujet remplie davantage par une ancienne information, est vraie pour tous les types de sujets (transitifs, intransitifs et copules). La position sujet de verbe transitif (Test T : $t_{(425)} = -28,97$, $p < 0,0001$), la position sujet intransitif (Test T : $t_{(441)} = -25,7$, $p < 0,0001$) et la position sujet de copules (Test T : $t_{(263)} = 7,34$, $p < 0,0001$) maintiennent plus une information qu'elles en introduisent (Figure 83).

Figure 83 : Proportions des SN maintenant et introduisant selon en position sujet



Nous pouvons tout de même observer que la position sujet transitif est la position sujet étant le plus remplie par une ancienne information et la moins remplie par une nouvelle information.

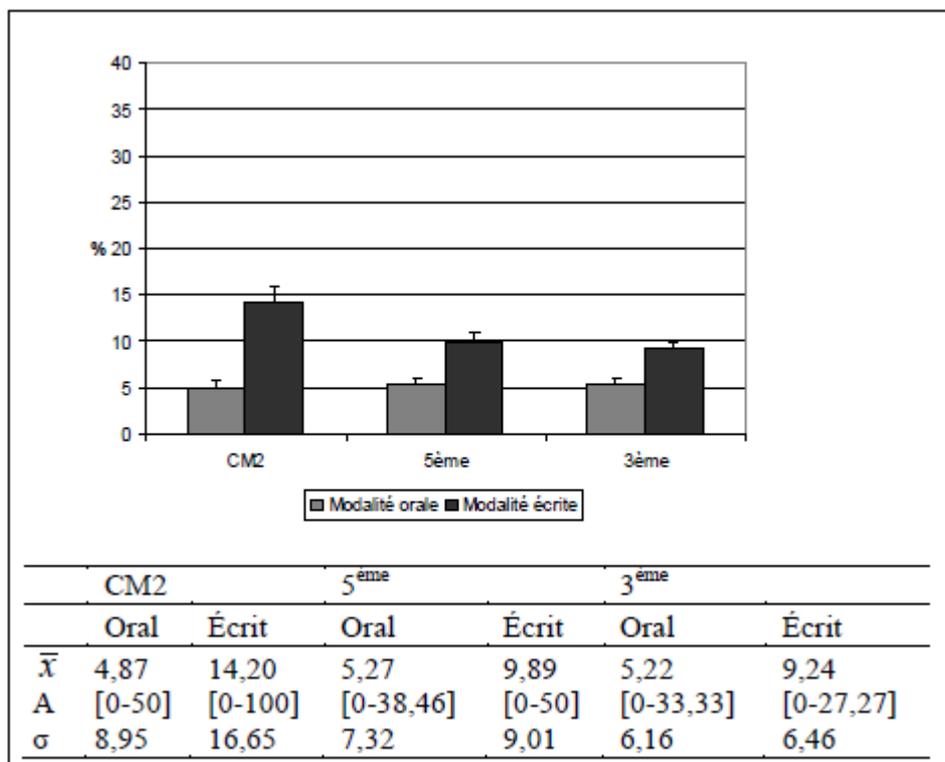
10.2.2. Proportion de SN introduisant une information selon les positions syntaxiques

10.2.2.1. Proportion de SN introduisant une information en position sujet

Les figures 84 et 85 présentent *les proportions des SN introduisant une information en position sujet* en fonction du NIVEAU SCOLAIRE, du TYPE DE TEXTE et de la MODALITÉ.

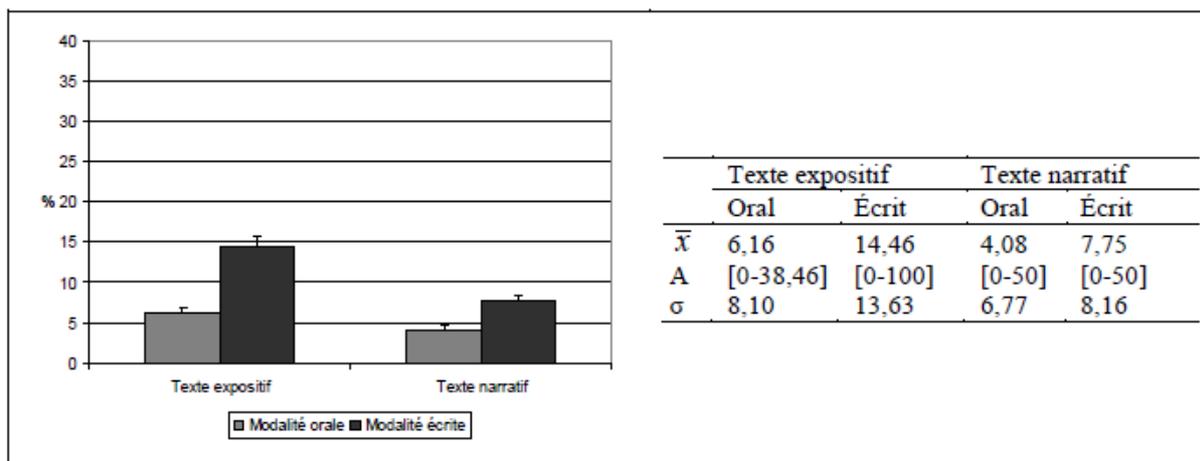
Les différences selon le NIVEAU SCOLAIRE sont significatives ($F_{(2,503)} = 3,025$, $p < 0,0494$). *La proportion des SN introduisant une information en position sujet* diminue avec le NIVEAU SCOLAIRE (Figure 84). Les écoliers de CM2 introduisent significativement plus de nouveaux référents dans cette position que les individus de 5^{ème} ($p = 0,0500$) et de 3^{ème} ($p = 0,0198$) dont les différences de proportion ne sont pas significatives ($p = 0,7262$).

Figure 84 : Proportions des SN introduisant une information en position sujet selon le NIVEAU SCOLAIRE et la MODALITÉ (par rapport au nombre total de SN introduisant et maintenant une information)



Le contexte de production a également un effet significatif (TYPE DE TEXTE ($F_{(1,503)} = 28,624$, $p < 0,0001$) et MODALITÉ ($F_{(1,503)} = 53,545$, $p < 0,0001$)). *L'introduction d'une nouvelle information en position sujet se réalise davantage dans les textes expositifs et écrits que dans les textes narratifs et oraux* (Figure 84 et Figure 85).

Figure 85 : Proportions des SN introduisant une information en position sujet selon le TYPE DE TEXTE et la MODALITÉ (par rapport au nombre total de SN introduisant et maintenant une information)



L'ORDRE n'affecte pas significativement *la proportion des SN introduisant une information en position sujet*.

Les résultats ci-dessus révèlent également deux interactions entre les facteurs NIVEAU SCOLAIRE et MODALITÉ ($F_{(2,503)} = 4,203$, $p < 0,0155$) ainsi que le TYPE DE TEXTE et la MODALITÉ ($F_{(1,503)} = 8,063$, $p < 0,0047$). En production écrite, *la production de SN introduisant une nouvelle information en position sujet* varie davantage selon le NIVEAU SCOLAIRE qu'en production orale. Les enfants les plus jeunes marquent davantage la différenciation entre les deux modalités du langage que les plus âgés. De plus, en production narrative, l'effet de la modalité est atténué. En texte expositif écrit, la proportion des SN introduisant une nouvelle information en position sujet atteint son maximum.

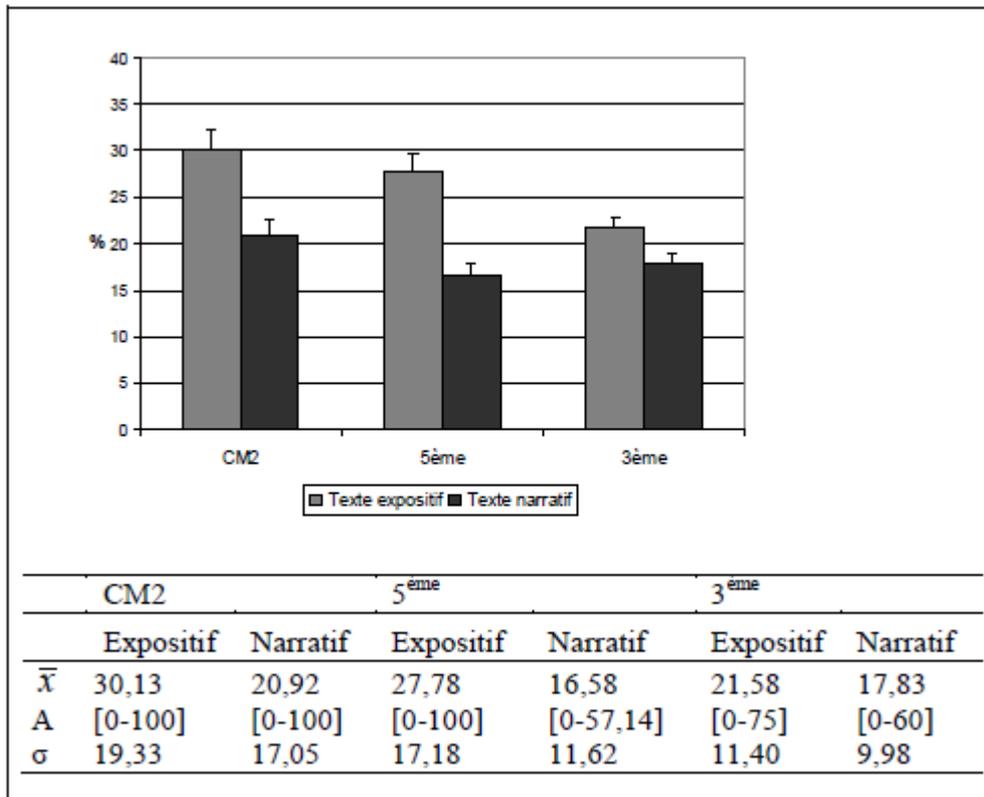
10.2.2.2. Proportion de SN introduisant une information en position objet

Les figures 86 et 87 présentent *les proportions des SN introduisant une information en position objet* en fonction du NIVEAU SCOLAIRE, du TYPE DE TEXTE et de la MODALITÉ.

Les différences selon le NIVEAU SCOLAIRE sont significatives ($F_{(2,503)} = 6,897$, $p < 0,0011$). La proportion des *SN introduisant une nouvelle information en position objet* diminue avec les années de scolarisation (Figure 86) ; les écoliers de CM2 ont une proportion significativement plus importante que les individus de 5^{ème} ($p = 0,0368$) et de 3^{ème} ($p < 0,0002$) dont les différences de proportion ne sont pas significatives ($p = 0,1175$).

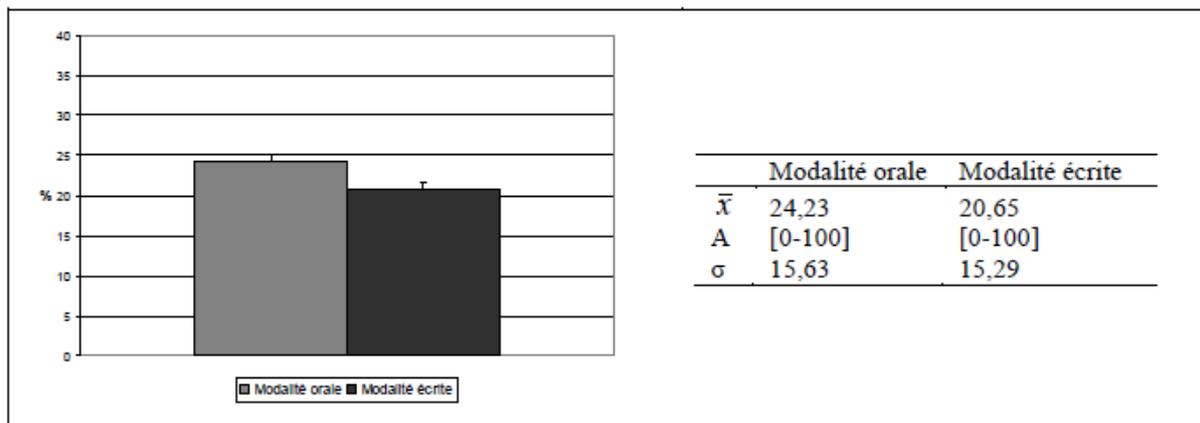
Le TYPE DE TEXTE fait également varier *la proportion des SN introduisant une information en position objet* ($F_{(1,503)} = 38,980$, $p < 0,0001$). Ce type de SN a une proportion plus importante dans les textes expositifs que dans les textes narratifs (Figure 86).

Figure 86 : Proportions des SN introduisant une information en position objet selon le NIVEAU SCOLAIRE et le TYPE DE TEXTE (par rapport au nombre total de SN introduisant et maintenant une information)



La MODALITÉ fait varier la proportion des SN introduisant une information en position objet ($F_{(1,503)} = 7,533, p < 0,0063$). Ce type de SN a une proportion plus importante dans les textes oraux que dans les textes écrits (Figure 87).

Figure 87 : Proportions des SN introduisant une information en position objet selon la MODALITÉ (par rapport au nombre total de SN introduisant et maintenant une information)



L'ORDRE n'a pas d'effet significatif sur cette VD.

Les résultats montrent également une interaction entre le NIVEAU SCOLAIRE et le TYPE DE TEXTE ($F_{(2,503)} = 3,019$, $p < 0,0497$) (Figure 86). La différence entre les deux types de textes diminue avec le niveau scolaire.

10.2.2.3. Proportion de SN introduisant une information en position autre

La Figure 88 présente *les proportions des SN introduisant une information en position autre* en fonction du TYPE DE TEXTE et de la MODALITÉ.

Les différences en fonction du contexte de production sont significatives (TYPE DE TEXTE ($F_{(1,503)} = 7,761$, $p < 0,0055$) et MODALITÉ ($F_{(1,503)} = 4,254$, $p < 0,0397$)) ; les deux autres facteurs – NIVEAU SCOLAIRE et ORDRE – ne font pas varier cette VD. *L'introduction d'un nouveau référent en position autre* caractérise plus les textes narratifs et écrits que dans les textes expositifs et oraux (Figure 88 et Figure 89).

Figure 88 : Proportions des SN introduisant une information en position autre (pronom et SNL) selon le TYPE DE TEXTE (par rapport au nombre total de SN introduisant et maintenant une information)

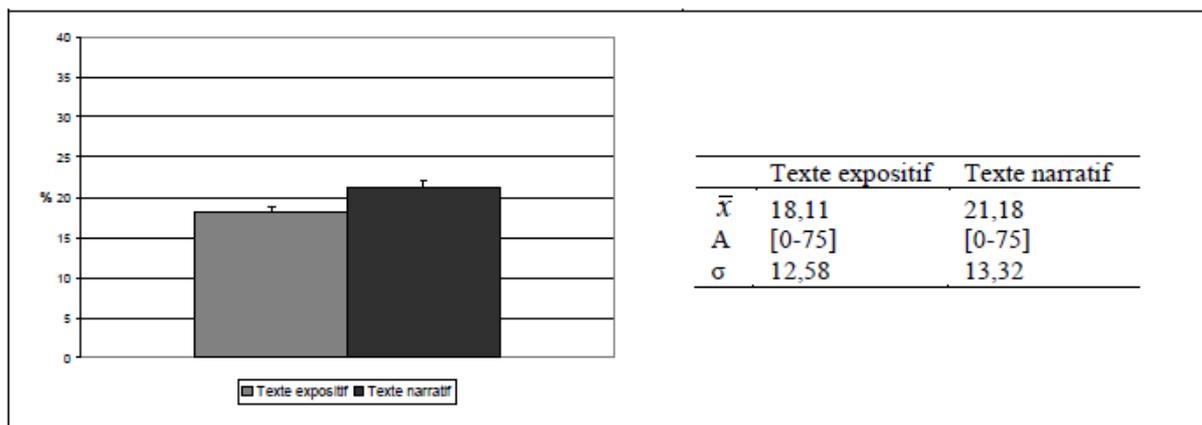
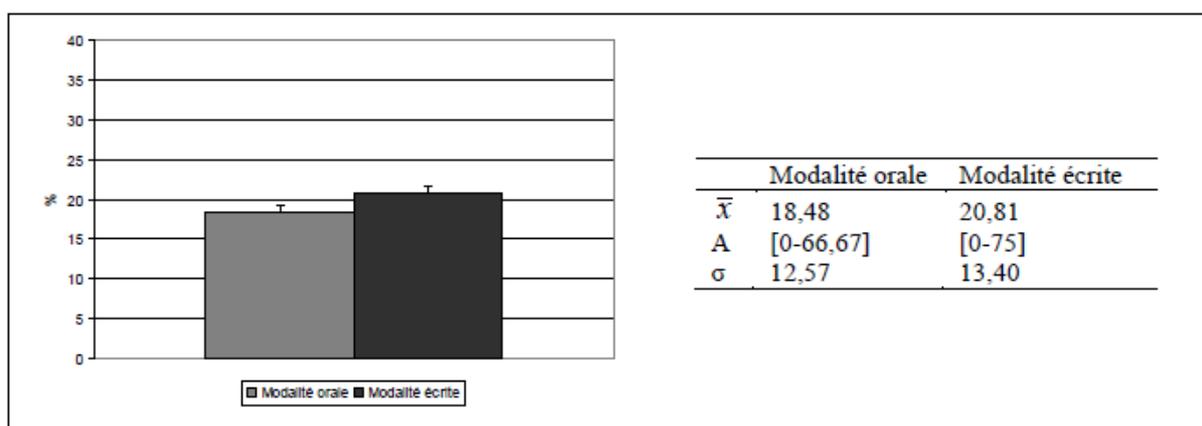


Figure 89 : Proportions des SN introduisant une information en position autre (pronom et SNL) selon la MODALITÉ (par rapport au nombre total de SN introduisant et maintenant une information)



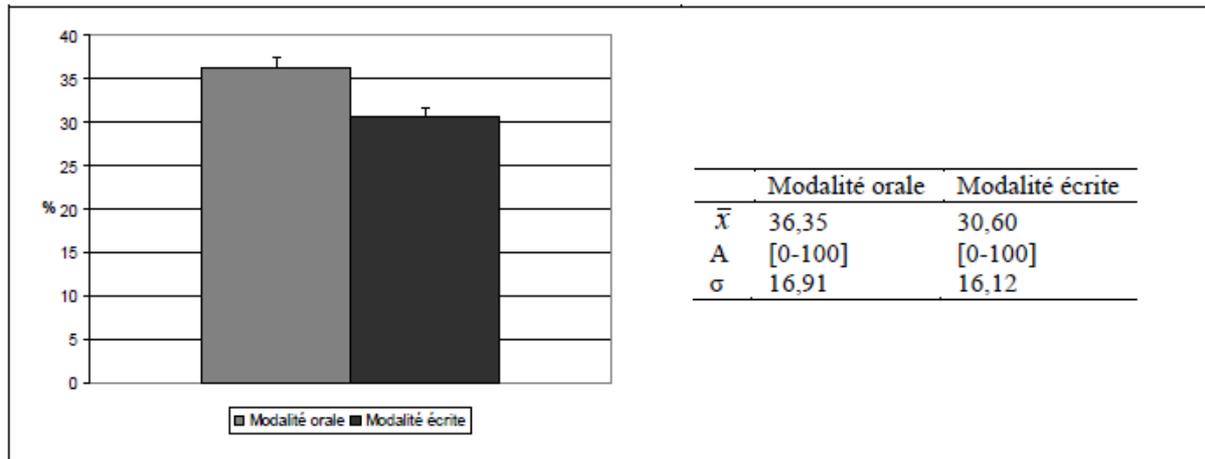
10.2.3. Proportion de SN maintenant une ancienne information selon les positions syntaxiques

10.2.3.1. Proportion de SN maintenant une ancienne information en position sujet

La Figure 90 présente *les proportions des SN maintenant une information en position sujet* en fonction de la MODALITÉ.

Les différences en fonction de la MODALITÉ sont significatives ($F_{(1,503)} = 16,219$, $p < 0,0001$). Les trois autres facteurs – le NIVEAU SCOLAIRE, le TYPE DE TEXTE et l'ORDRE – n'interviennent pas. *La proportion des SN maintenant une ancienne information en position sujet* est plus importante dans les textes oraux que dans les textes écrits (Figure 90).

Figure 90 : Proportions des SN maintenant une information en position sujet selon la MODALITÉ (par rapport au nombre total de SN introduisant et maintenant une information)

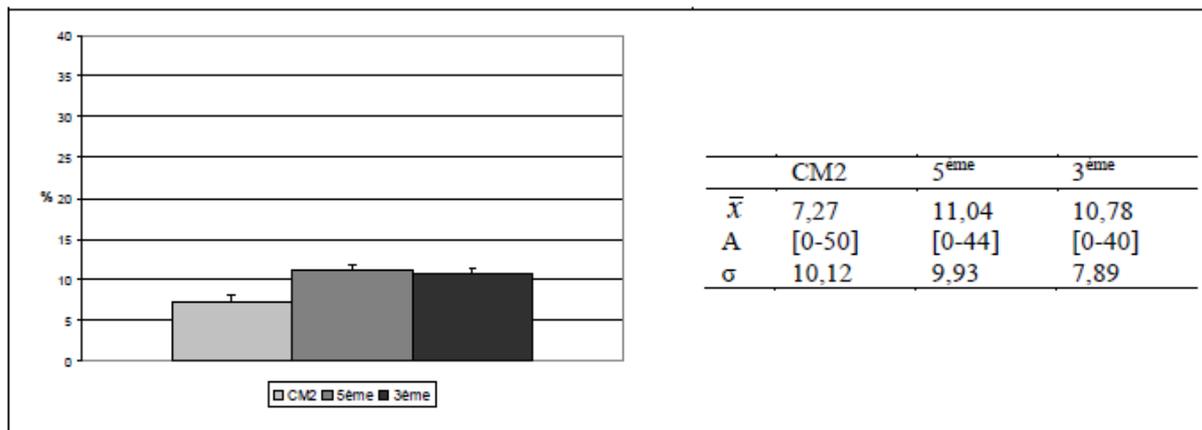


10.2.3.2. Proportion de SN maintenant une ancienne information en position objet

Les figures 91 à 93 présentent *les proportions des SN maintenant une information en position objet* en fonction du NIVEAU SCOLAIRE, du TYPE DE TEXTE et de l'ORDRE DE PRODUCTION.

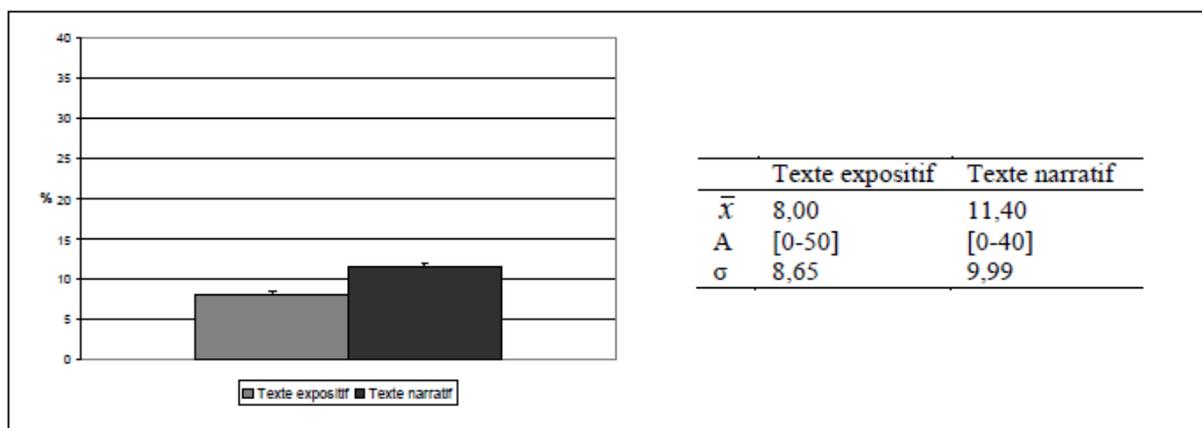
Les différences en fonction du NIVEAU SCOLAIRE sont significatives ($F_{(2,503)} = 9,279$, $p < 0,0001$). *La proportion des SN maintenant une ancienne information en position objet* augmente avec le NIVEAU SCOLAIRE (Figure 91); les individus de 5^{ème} ($p < 0,0001$) et de 3^{ème} ($p < 0,0004$) maintiennent significativement plus l'information que les CM2, néanmoins les deux populations de collège ont des proportions dont la différence n'est pas significative ($p = 0,7891$).

Figure 91 : Proportions des SN maintenant une information en position objet selon le NIVEAU SCOLAIRE (par rapport au nombre total de SN introduisant et maintenant une information)



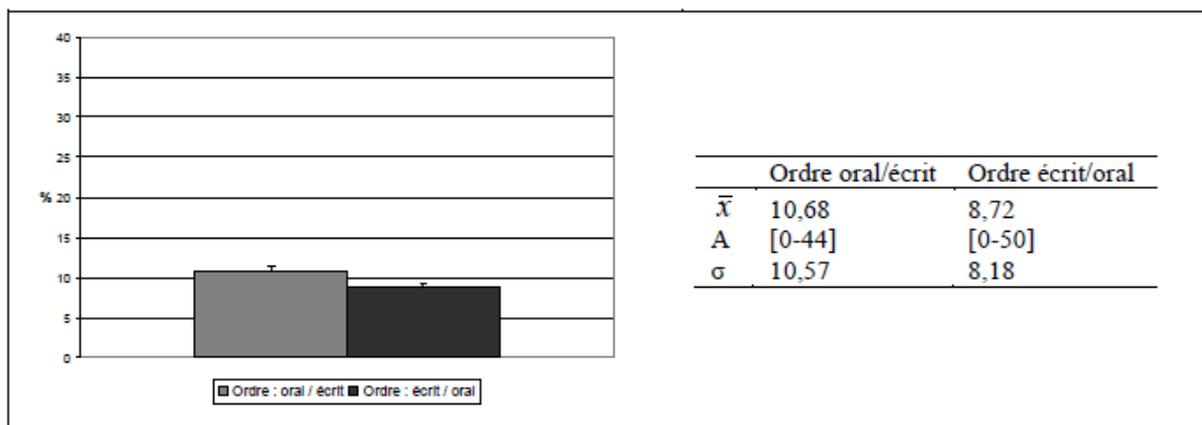
Le TYPE DE TEXTE fait également varier la proportion des SN maintenant une information en position objet, qui a une proportion plus importante dans les textes narratifs que dans les textes expositifs ($F_{(1,503)} = 18,204$, $p < 0,0001$) (Figure 92).

Figure 92 : Proportions des SN maintenant une information en position objet selon le TYPE DE TEXTE (par rapport au nombre total de SN introduisant et maintenant une information)



Pour finir, le maintien de l'information en position objet caractérise davantage les textes ayant été produits dans l'ordre de passation oral/écrit que dans ceux ayant été produits dans l'ordre de passation écrit/oral ($F_{(1,503)} = 6,116$, $p < 0,0137$) (Figure 93).

Figure 93 : Proportions des SN maintenant une information en position objet selon l'ORDRE (par rapport au nombre total de SN introduisant et maintenant une information)



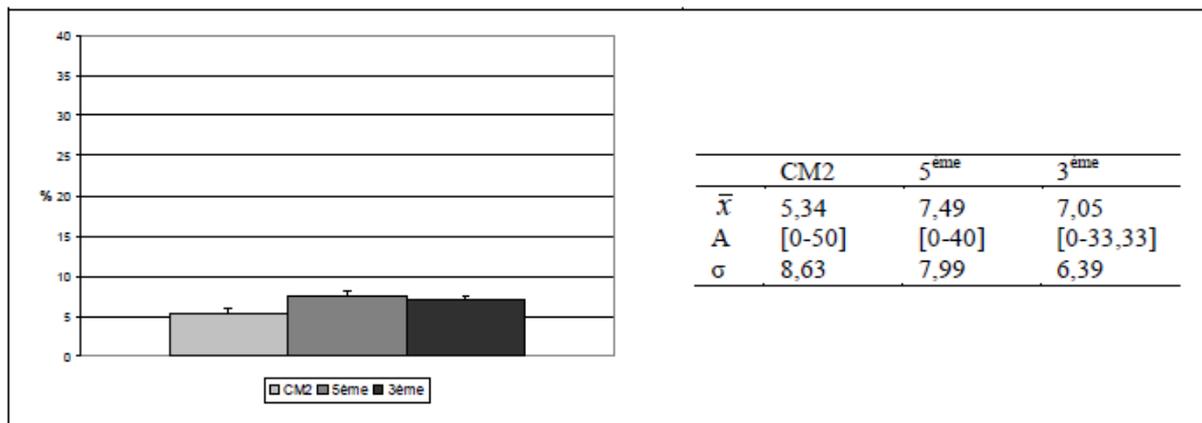
La variable MODALITÉ n'affecte pas significativement la proportion des SN maintenant une information en position objet.

10.2.3.3. Proportion de SN maintenant une ancienne information en position autre

Les figures 94 et 96 présentent *les proportions des SN maintenant une information en position autre* en fonction du NIVEAU SCOLAIRE, du TYPE DE TEXTE et de la MODALITÉ.

Les différences selon le NIVEAU SCOLAIRE sont significatives ($F_{(2,503)} = 4,048$, $p < 0,0180$). *Le maintien de l'information en position autre* augmente avec le NIVEAU SCOLAIRE (Figure 94) ; les collégiens de 5^{ème} et de 3^{ème}, dont les proportions ne se distinguent guère ($p = 0,5822$), maintiennent significativement plus une information en position autre que les écoliers (respectivement, $p = 0,0074$ et $p = 0,0207$).

Figure 94 : Proportions des SN maintenant une ancienne information en position autre selon le NIVEAU SCOLAIRE (par rapport au nombre total de SN introduisant et maintenant une information)



Le contexte de production affecte également *la proportion des SN maintenant une information en position autre* (TYPE DE TEXTE ($F_{(1,503)} = 35,364$, $p < 0,0001$) et MODALITÉ ($F_{(1,503)} = 3,675$, $p < 0,05$)). Ce type de SN est plus fréquent dans les textes narratifs que dans les textes expositifs (Figure 95).

Figure 95 : Proportions des SN maintenant une ancienne information en position autre selon le TYPE DE TEXTE (par rapport au nombre total de SN introduisant et maintenant une information)

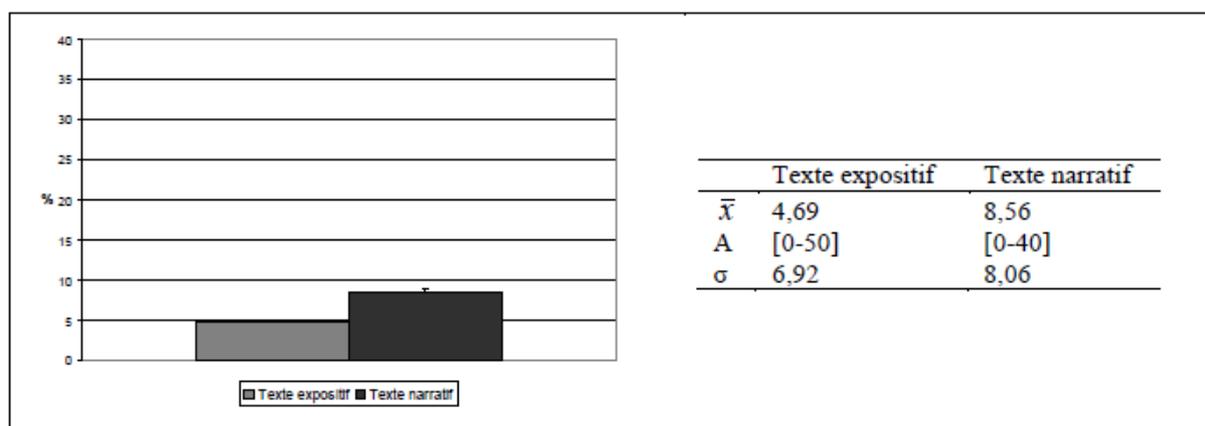
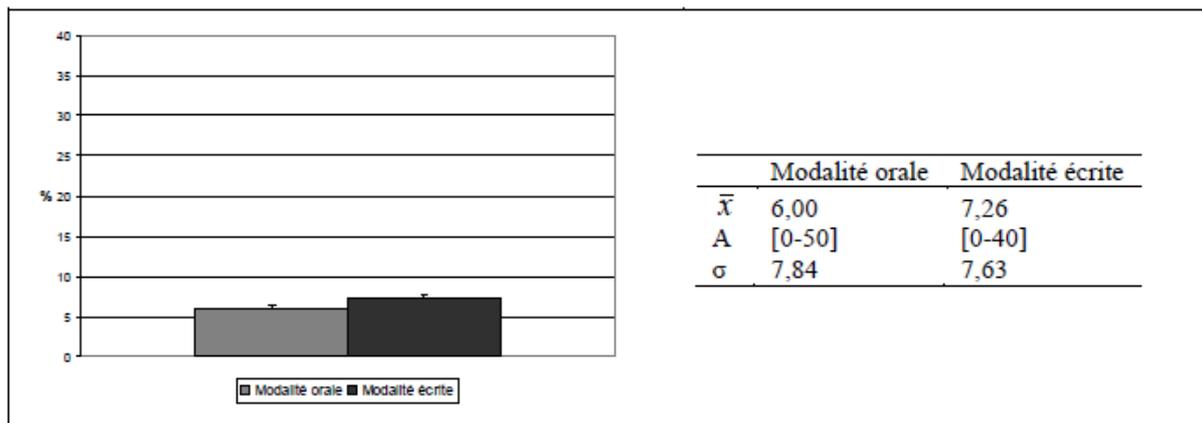


Figure 96 : Proportions des SN maintenant une ancienne information en position autre selon la MODALITÉ (par rapport au nombre total de SN introduisant et maintenant une information)



L'ORDRE n'a pas d'effet significatif cette VD.

10.2.4. Discussion de la Section 10.2

Le but de cette partie était de vérifier notre hypothèse selon laquelle l'observation des deux variables – *la proportion des SN maintenant ou introduisant des informations* – selon les positions syntaxiques (Section 10.2, p. 240) compléterait les résultats que nous trouvons lors des analyses ne prenant pas en considération la position syntaxique (Section 10.1, p. 231).¹⁰⁸ Ceci est confirmé par nos analyses.

Rappelons que *la proportion des SN introduisant une information* diminue avec les années de scolarisation. Les enfants les plus jeunes, comme nous l'avons mentionné déjà précédemment, font des listes de nouveaux référents qu'ils ne discutent guère. En revanche, avec l'expérience scolaire, ils apprennent à développer leurs idées. C'est ainsi que *la proportion des SN maintenant une information* augmente avec les années de scolarisation. De plus, *la proportion des SN introduisant une information* est plus importante dans les textes expositifs et dans les textes écrits que dans les textes narratifs et oraux. Par contraste, *la proportion des SN maintenant une information* augmente avec les années de scolarisation. Cette proportion est plus importante dans les textes narratifs et dans les textes oraux que dans les textes expositifs et écrits.

Les proportions des SN introduisant et maintenant une information en position sujet, objet et autre varient selon le NIVEAU SCOLAIRE.

Les tendances relevées pour *la proportion des SN maintenant et introduisant une information* sont en partie confirmées par les analyses prenant en considération les positions syntaxiques.

¹⁰⁸ Les attentes, hypothèses et problématiques sont davantage explicitées au Chapitre 3 (Section 3.2.2.3, p. 102).

Concernant *la proportion de la nouvelle information* selon les positions syntaxiques, seules les proportions en position sujet et objet varient selon le NIVEAU SCOLAIRE et confirment la tendance observée précédemment en diminuant avec le NIVEAU SCOLAIRE. *La proportion des SN maintenant une information en position objet et autre* confortent, en augmentant avec le NIVEAU SCOLAIRE, ceux obtenus pour la proportion des SN maintenant une information toutes positions confondues.

Les proportions des SN introduisant et maintenant une information en position sujet, objet et autre varient selon le TYPE DE TEXTE.

Si *les proportions de SN introduisant une information en position sujet et objet* sont plus importantes dans les productions expositives, ce qui confirme la tendance observée lors des analyses toutes positions confondues, *la proportion de SN introduisant une information en position autre* est plus élevée dans les textes narratifs. La production de nouvelle information en position autre correspond l'introduction d'informations secondaires (compléments circonstanciels) ayant souvent pour rôle d'enrichir le cadre global (temps, lieu, *etc.*) d'une narration. De plus, si *la proportion des SN maintenant une information* toutes positions confondues est plus importante dans les productions narratives, la conclusion diffère quelque peu selon la position syntaxique. *La proportion des SN maintenant une information en position objet et autre* est plus élevée dans les textes narratifs, alors que *la proportion des SN maintenant une information en position sujet* ne varie pas selon ce facteur.

La MODALITÉ fait varier certaines proportions. *La proportion des SN introduisant une information en position sujet* est plus importante à l'écrit qu'à l'oral, ce qui confirme les résultats de la Section 10.1 (p. 231). Néanmoins, la tendance s'inverse pour *la proportion des SN introduisant une information en position objet* qui est plus élevée en production orale qu'en production écrite. Ces résultats semblent logiques dans la mesure où les principes naturels de transmission de l'information de la langue orale veulent que la nouvelle information soit davantage en position objet (Du Bois, 1987, 2003). Ainsi, mettre une nouvelle information, et donc de surcroît du lexique, en position sujet exige d'aller à l'encontre de ces principes. La modalité écrite, par son rythme de production lent et donc par un temps de planification plus long, offre aux locuteurs/scripteurs les possibilités pour avoir un tel comportement.

Les proportions des SN maintenant une information en position sujet et autre varient selon la MODALITÉ. La proportion en position sujet est plus importante en production orale qu'en

production écrite, ce qui confirme la tendance observée lors des analyses toutes positions confondues (Section 10.1, p. 231). En revanche, la proportion en position autre est plus élevée en production écrite qu'en production orale.

Si l'ORDRE ne fait varier pas *les proportions des SN maintenant et introduisant une information* toutes positions confondues (Section 10.1, p.231), cette variable affecte *la proportion des SN maintenant une information en position objet*. Cette dernière est plus importante dans les textes des individus ayant produit dans l'ordre oral/écrit que dans ceux des individus ayant produit dans l'ordre écrit/oral.

Nous concluons cette partie avec une remarque concernant la distribution des SN maintenant et introduisant une information selon la position syntaxique (Figure 81 et 82) qui rappelle les travaux de Du Bois (1987, 2003) et de Khorounjaia et Tolchinsky (2004). Nos résultats révèlent que l'ancienne information occupe préférentiellement la position sujet. Cette tendance confirme, en partie, l'une des contraintes de la PAS de Du Bois, à savoir que l'ancienne information se trouve préférentiellement en position sujet, notamment en position sujet de verbe transitif. Cette contrainte de la PAS est confirmée : nos individus introduisent préférentiellement les nouveaux référents en position non sujet et maintiennent une information de préférence en position sujet (Figure 97), ce qui va dans le sens des travaux de Salas *et al.* (2008) concernant une population favorisée (classe moyenne supérieure, Projet Spencer).¹⁰⁹

Figure 97 : La position préférée de la nouvelle et de l'ancienne information

Nouvelle information : Objet / Autre > Sujet	Ancienne information : Sujet > Objet / Autre
--------------------------------------------------------	--------------------------------------------------------

La tendance à avoir une position sujet remplie par une ancienne information plutôt que par une nouvelle se confirme en observant les trois types de sujets (transitifs, intransitifs et copules). Nous notons juste que la position sujet transitif est la position sujet ayant la proportion d'anciennes informations la plus élevée.

¹⁰⁹ Dans cette étude, deux langues sont observées, le français et l'espagnol.

10.3. Conclusion du Chapitre 10

Les résultats de ce chapitre confirment en partie nos hypothèses. *Les proportions de SN maintenant et introduisant une information* varient avec le NIVEAU SCOLAIRE. Nous n'avions pas posé d'attentes spécifiques quant à la proportion des SN introduisant et maintenant une information. Les analyses statuent sur cette hypothèse : *la proportion de SN introduisant une nouvelle information* diminue alors que *celle des SN maintenant une information* augmente. Néanmoins, en regardant *le nombre moyen de SN introduisant une nouvelle information* et *celui des SN maintenant une information par texte*, nous observons que les enfants maintiennent certes de plus en plus avec le NIVEAU SCOLAIRE mais ils introduisent également de plus en plus de nouvelles informations (Tableau 10). Ces résultats révèlent qu'avec l'expérience, les enfants enrichissent de plus en plus leurs textes que ce soit par l'introduction de nouvelles informations ou par le maintien de ces informations dans le but de développer des thématiques.

La proportion des SN introduisant une nouvelle information et *celle des SN maintenant une nouvelle information* varient également selon le contexte de production (TYPE DE TEXTE et MODALITÉ). Néanmoins, ces deux variables ne varient pas selon l'ORDRE.

De plus, l'observation *des SN maintenant et introduisant une information* selon les positions syntaxiques (Section 10.2, p.240) rend compte de certains phénomènes relatifs à la position syntaxique. Ainsi, *la proportion des SN introduisant une information en position sujet* est plus élevée en production écrite ; en revanche, *la proportion des SN introduisant une information en position objet* est plus importante à l'oral.

Selon les principes de transmission de l'information de l'oral, introduire une information se fait forcément sous forme lexicale et le choix naturel est de mettre cette nouvelle information en position objet. Ceci est dû à la relation étroite entre position syntaxique et forme linguistique dont nous avons parlé dans les chapitres précédents. Les formes lexicales sont de préférence en position non sujet. Produire cette nouvelle information sous forme lexicale en position sujet exige du locuteur d'aller à l'encontre de deux principes naturels du traitement de l'information : (a) le fait que la nouvelle information est préférentiellement en position non sujet ; et (b) le fait que la forme lexicale ne remplit pas la forme sujet. Les individus trouvent les ressources nécessaires en modalité écrite, qui de par son rythme de production lent,

implique un temps de planification long. Durant ce temps, les scripteurs peuvent, entre autres, en profiter pour défier ces principes.

Dans ce chapitre, nous avons pu voir ce qu'il en était de deux des contraintes de la PAS de Du Bois (1987, 2003) : (a) une seule nouvelle information par clause ; et (b) la nouvelle information est en position non sujet. Les individus de notre étude se soumettent davantage aux contraintes de la PAS dans les contextes narratif et oral. En production expositive et écrite, les enfants la respectent de moins en moins avec les années de scolarisation et l'expérience de l'écrit. En effet, cette modalité du langage implique des contraintes différentes des contraintes naturelles de l'oral auxquelles les enfants apprennent à faire face.

Ces analyses nous ont éclairée sur le statut informationnel des SN dans une optique prenant également en considération les positions syntaxiques. Ces résultats peuvent être complétés par une étude portant sur les formes linguistiques utilisées selon les statuts informationnels. Nous savons que la nouvelle information est lexicale. Néanmoins, nous pourrions trouver dans les productions des enfants : (a) des *réintroductions sous formes pronominales* (b) des cas de *référent manquant*, (c) des cas de *référents multiples*, et (d) des *problèmes d'accord*. C'est pourquoi nous proposons des analyses concernant l'observation de l'utilisation non conventionnelle de pronoms par les enfants. Si les analyses révèlent une faible proportion de ces utilisations, cela montrera alors que les enfants savent comment encoder la nouvelle information et qu'il faut utiliser un SNL lorsqu'il y a ambiguïté ou réintroduction.

Chapitre 11

Utilisation non conventionnelle des pronoms

Lors du codage des productions de notre population, nous avons observé des utilisations non conventionnelles de pronoms. Quatre catégories ont émergé des données : (a) *réintroduction pronominale*, (b) *référent manquant*, (c) *référents multiples*, et (d) *problème d'accord*. La première catégorie, la *réintroduction pronominale*, correspond à l'utilisation d'un pronom pour encoder un référent qui est trop loin dans le texte pour être encore actif en MDT. Le

traitement naturel de l'information de la langue voudrait que la réintroduction d'un tel référent se réalise sous forme lexicale.

(143) *et euh euh euh pour euh dire c(e) qu'il [l'oncle] a vécu ben **au juge** elle [la femme] lui [l'oncle] a dit de faire une lettre pour tout dire mais elle [la femme] lui [l'oncle] a dit d(e) pas mettre son adresse d'où il [l'oncle] est main(te)nant elle [la femme] lui [l'oncle] a dit de mettre son ancienne adresse comme ça **il** [le juge] pourra pas le [l'oncle] poursuivre. (CM2 n°19, narratif oral, E/O)*

Dans cet exemple, l'écolier maintient *juge* par le biais du pronom *il*, mais cette référence se fait trop loin et la réintroduction du référent, *juge*, aussi loin, sous forme pronominale entrave la compréhension. Il aurait fallu un SNL.

La deuxième catégorie, *référent manquant*, concerne l'utilisation de pronoms anaphoriques qui apparaissent sans mention d'un référent au préalable (144). L'exemple suivant est la première clause du texte d'un enfant.

(144) *Et **ils** m'ont accusée. (CM2 n°01, narratif oral, O/E)*

Le participant utilise le pronom *ils* et raconte une histoire concernant les référents de *ils*, mais il ne dit jamais, et ce notamment par le biais d'un SNL, de qui il s'agit.

La troisième catégorie est celle des pronoms anaphoriques à *référents multiples*. Il s'agit d'un pronom pouvant se rapporter à plus d'un référent mentionné au préalable (145).

(145) *À la cantine y avait un garçon là / et il jouait au foot et tout / et y avait un autre garçon dans la même équipe / pa(r)ce que **il** a joué personnel / **il** a commencé à l'insulter. (CM2 n°43, narratif oral, O/E)*

Dans cet exemple, l'enfant utilise le pronom de la troisième personne du singulier pour maintenir *un garçon*. Néanmoins, lors de la deuxième et troisième utilisation, ceci engendre un problème de cohérence pour l'interlocuteur : il est difficile de savoir avec certitude à quel référent renvoie ce pronom, *un garçon* ou *un autre garçon* (notons que même la prosodie ne nous aide pas).

Enfin, nous avons trouvé un quatrième cas d'usage problématique des pronoms anaphoriques, celui des *problèmes d'accord*. Il s'agit d'un accord en genre et en nombre du pronom avec son référent qui s'écarte de la norme de la *langue de scolarisation* et plus spécifiquement de son mode écrit (146)

(146) *Et la bagarre ben euh euh entre **deux personnes** c'est un peu bête / pa(r)ce qu'**i(ls)** s(e) font mal tous les deux.* (CM2 n°17, expositif oral, O/E)

Dans cet exemple constituant un extrait de la production expositive orale, le locuteur, un enfant de CM2, fait une anaphore, il remplace *deux personnes* par *ils* ; l'accord en genre entre le référent et le pronom.

La proportion de ces types d'utilisation (et notamment les deux premières) peut nous indiquer si les participants de notre étude mettent en œuvre une des règles de la transmission de l'information : introduire une nouvelle information sous forme lexicale et utiliser le lexique pour lever toute ambiguïté.

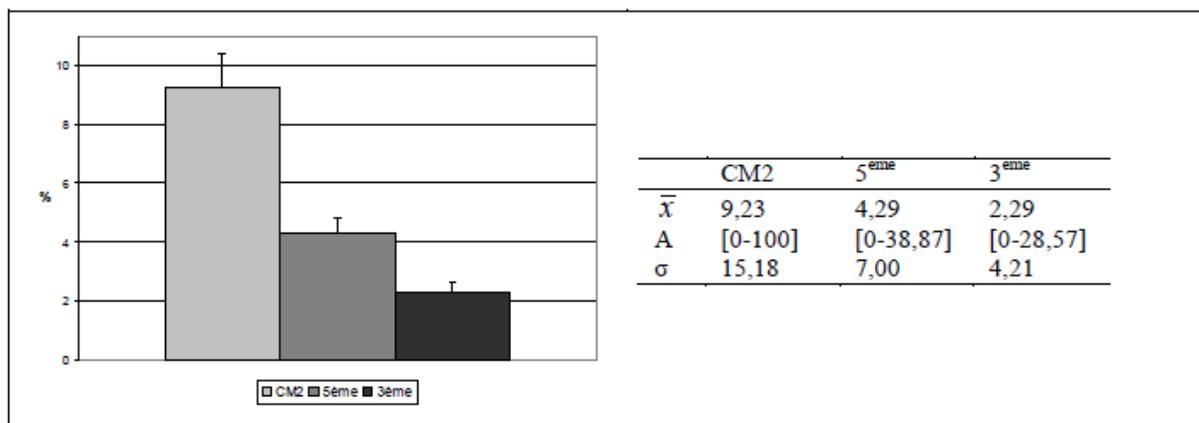
De plus, nous voulons attirer l'attention sur un dernier aspect. En effet, les résultats qui vont suivre sont à interpréter en prenant en considération l'aspect artificiel du monologue et notamment à l'oral. Ainsi, par exemple, la consigne du texte expositif contenait la clause « *Maintenant j'aimerais que tu parles sur le thème des problèmes entre les gens.* » et les investigateurs pouvaient relancer la consigne en disant, par exemple, *qu'est-ce qu'il se passe entre les gens.* Nous avons pu alors remarquer que certains enfants, et notamment les plus jeunes, commençaient alors leur texte par *ils*, faisant référence aux *gens* de la consigne.

11.1. Analyses

Les figures 98 et 99 présentent *la proportion des utilisations non conventionnelles des pronoms* par rapport au nombre total de SN en fonction du NIVEAU SCOLAIRE et du TYPE DE TEXTE.

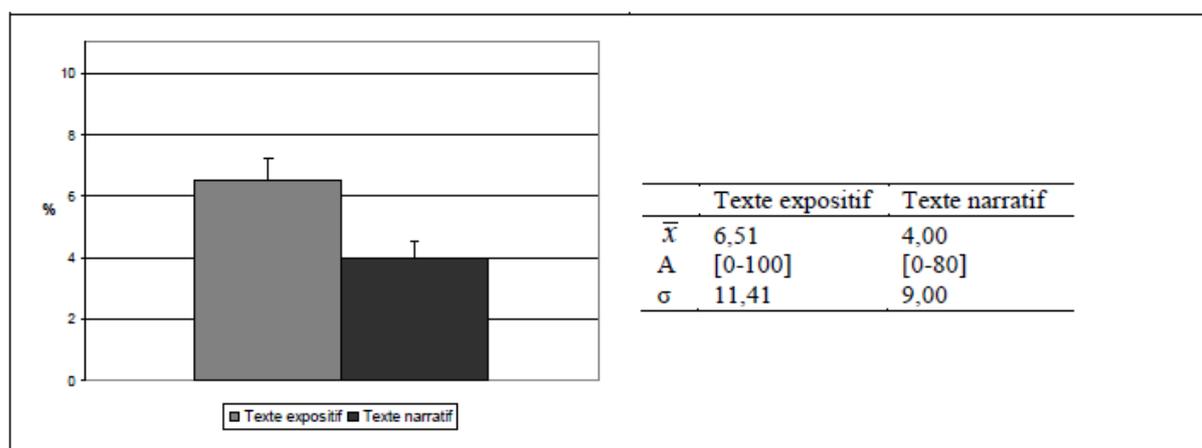
Les différences en fonction du NIVEAU SCOLAIRE sont significatives ($F_{(2,504)} = 22,853$, $p < 0,0001$) : *la proportion des pronoms problématiques* diminue avec les années de scolarisation (Figure 98). Les individus les plus jeunes (CM2) ont une proportion significativement plus élevée que les collégiens de 5^{ème} ($p = 0,0001$) et de 3^{ème} ($p = 0,0001$). La différence de proportion entre ces deux populations d'adolescents n'est pas significative ($p = 0,0603$).

Figure 98 : Proportions de l'utilisation non standard des pronoms selon le NIVEAU SCOLAIRE (par rapport au nombre total de SN)



Le TYPE DE TEXTE a également un effet sur *la proportion des utilisations non conventionnelles des pronoms* ($F_{(1,504)} = 8,535$, $p < 0,0036$) : elle est plus importante dans les textes expositifs que dans les textes narratifs (Figure 99).

Figure 99 : Proportions de l'utilisation non standard des pronoms selon le TYPE DE TEXTE (par rapport au nombre total de SN)



Les deux autres facteurs – MODALITÉ et ORDRE – ne font pas varier significativement *la proportion des pronoms problématiques*.

11.2. Discussion du Chapitre 11

Dans ce chapitre, nous voulions observer la proportion des *utilisations non conventionnelles des pronoms* et voir si elle variait selon les facteurs de notre étude – NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE –.¹¹⁰

¹¹⁰ Les attentes, hypothèses et problématiques sont davantage explicitées au Chapitre 3 (Section 3.2.2.4, p. 103).

Rappelons que nous nous attendions à ce que *les utilisations non conventionnelles des pronoms* diminuent avec le NIVEAU SCOLAIRE. Les enfants respecteraient les principes de transmission de l'information de l'oral : quand il y a introduction d'un nouveau référent, réintroduction et ambiguïté, le pronom est évité. Ceci est confirmé par l'analyse des données. Ainsi, les situations (a) *réintroduction pronominale*, (b) *référent manquant* (147), (c) *référents multiples*, et (d) *problème d'accord* (148) deviennent de plus en plus rares.

(147) Je **lui** dis qu'on peut s'arrêter pour parler mais **lui il** ne veut pas. (CM2 n°15, narratif écrit, E/O)

(148) Au début j'ai eu une dispute avec euh **mon meilleur copain**, ça a dégénéré, et après le jour qui suivait, **ils** m'ont même pas calculé. (CM2 n°46, narratif oral, E/O)

L'individu n°15 introduit le pronom *lui* deux fois de suite puis le pronom *il* mais en aucun cas le référent de ces pronoms n'a été lexicalisé précédemment. Le sujet n°46 utilise le pronom masculin de la troisième personne du pluriel, *ils*, pour reprendre une ancienne information énoncée précédemment, *mon meilleur copain*, qui aurait dû être repris par le pronom masculin de la troisième personne du singulier.

Blanche-Benveniste (1996) énonce que l'une des règles majeures de la coréférence dans les discours est celle « *selon laquelle un syntagme nominal donné dans une première position peut entretenir une coréférence avec un pronom qui le suit* » (1996:109). Il semblerait qu'il s'agisse d'un phénomène de *mise en mémoire discursive* ainsi « *les détails descriptifs apportés par le syntagme nominal sont stockés en mémoire et le pronom qui les suit les relaie sans les répéter* » (1996:109). La mémoire discursive se définirait comme « *le savoir partagé lors des interactions de communication verbale* » (Berrendonner, 1986:259). Les enfants les plus jeunes, donc notamment ceux de CM2, ont des difficultés à gérer ce phénomène de *mise en mémoire discursive* et donc à respecter au mieux les règles de la coréférence dans le discours. La cohérence d'un texte met en jeu des outils syntaxiques, qui, utilisés à bon escient, assurent une bonne cohérence entre la représentation mentale du locuteur/scripteur et le texte, ce qui se développe avec l'âge (Traxler et Gernsbacher, 1995). Ceci permet alors aux destinataires d'accéder à cette représentation mentale. Les enfants feraient preuve d'*égocentrisme cognitif*¹¹¹ : ils auraient des difficultés à concevoir que les destinataires de leur message aient des représentations mentales différentes et non partagées. L'utilisation non conventionnelle des pronoms peut être en partie liée à ce dernier aspect. Le *common ground*

¹¹¹ Notion de Hickmann (2003) selon Piaget (1923).

est de plus en plus pris en considération et respecté. Les enfants prennent conscience que leurs représentations mentales sont différentes de celles des destinataires de leur message et sont alors en mesure d'adapter leurs choix linguistiques.

Nous suggérons également que *la proportion des utilisations non conventionnelles des pronoms* était plus fréquente dans les textes expositifs, contexte dans lequel les participants ont moins d'automatismes concernant le schéma de ce type de texte, que dans les textes narratifs. Les analyses des données vont dans le sens de nos attentes. Ce résultat est compréhensible. En effet, les individus lors de la production d'un texte expositif sont face à un contexte impliquant de fortes exigences cognitives. Rappelons que ce type de texte exige une certaine organisation des connaissances autour d'un contexte général non spécifique. Il appelle un effort d'adéquation entre le niveau du contenu et celui de la rhétorique (Britton, 1994 ; Katzenberger, 2004) ; la mobilisation de *la planification par connaissances transformées* est alors davantage exigée. Nous comprenons donc que ce texte puisse poser plus problème aux enfants qui utilisent davantage *la planification par connaissances rapportées* que *celle par connaissances transformées* (Chanquoy et Alamargot, 2002 ; Fayol, 1997 ; Olive et Piolat, 2005 ; Piolat *et al.*, 1993 ; Piolat et Roussey, 1992). Ainsi, additionner la gestion du *common ground*, du *given-new contract* et donc adapter leurs choix linguistiques en prenant en considération les représentations mentales et éventuelles connaissances de leur interlocuteur semble être trop lourd en plus des difficultés de ce texte.

Pour conclure, deux observations générales peuvent être faites : (a) *les proportions des utilisations problématiques des pronoms* sont de façon générale assez faibles ; et (b) avec les années de scolarisation, les proportions diminuent (9,23% pour les CM2 et 2,29% pour les 3^{ème}). Ces deux conclusions générales laissent penser que les enfants de notre étude gèrent tout de même assez bien certaines règles de la langue. Notamment, ils savent que pour produire une nouvelle information, pour réintroduire un référent et en cas d'ambiguïté lexicale, ils ont à utiliser une forme lexicale. La nouvelle information est forcément lexicale, mais qu'en est-il de l'ancienne information, statut informationnel pour lequel la langue permet d'utiliser aussi bien les pronoms que les SNL ? Nous avons vu, de plus, que maintenir une information reflète le développement du texte. Développer son texte semble être une capacité tardive, il peut alors être intéressant de se focaliser essentiellement sur cet aspect. Le chapitre qui suit se consacre alors à cette question.

Chapitre 12

Syntagmes nominaux maintenant une information, quel type de référence ?

Dans ce chapitre, nous traitons de la forme syntaxique des SN maintenant une information. Nous le divisons en deux parties. La première se consacre à savoir si l'ancienne information est plus fréquemment maintenue sous forme pronominale ou sous forme lexicale. La seconde partie se penche plus précisément sur le maintien sous forme lexicale (substitution lexicale) qui semble être une acquisition tardive, comme nous le verrons dans les résultats.

Dans la première partie (Section 12.1, p. 270) est présentée *la proportion des SN maintenant une information* et nous nous penchons sur la forme que prend ce maintien : pronom ou SNL. Dans l'exemple qui suit, l'information *les affaires de ma copine* est maintenue par un pronom. (149) *Les affaires de ma copine* avaient disparues et personne ne savait où **elles** étaient passées. (3ème n° 05, narratif écrit, E/O)

Un individu expert maintient aisément une information déjà connue par un SNL (150).

(150) « Ce qui m'a frappé, déclare **Nicolas Sarkozy**, c'est l'extrême violence de ces manifestations, une violence rarement vue en France. Je pense que nous avons vécu là un moment de vérité," ajoute **le ministre de l'Intérieur** qui affirme que la haine est d'abord territoriale, une bande s'approprie un territoire et y fait régner la loi de la force." A propos des destructions de bâtiments publics comme les écoles, **Nicolas Sarkozy** assure que les jeunes s'en prennent à l'école comme un symbole d'un service public : c'est une façon de dire 'ce territoire est le mien', de contester toutes les formes d'organisation publique au profit des lois de la bande." **Il** estime enfin que la crise a permis de percevoir le décalage entre le pays réel et le pays virtuel." On raconte aux Français une histoire qui ne correspond pas à ce qu'ils pensent. Ils rejettent la pensée unique, le prêchi-prêcha social-démocrate qui explique la délinquance par la pauvreté, qui, de façon excessive, privilégie les zones, les territoires plutôt que l'individu," lance **le président de l'UMP**. »¹¹²

¹¹² Extrait d'un article du Nouvel Observateur du mercredi 7 Décembre 2005.

Le journaliste maintient le référent *Nicolas Sarkozy*, introduit dès la première ligne, qu'une seule fois sous forme pronominale. Les trois autres mentions sont sous forme lexicale (*Nicolas Sarkozy, le ministre de l'Intérieur, le président de l'UMP*). Le scripteur fait preuve d'une grande capacité en utilisant son éventail de formes linguistiques tout en variant les utilisations ; il réalise des substitutions lexicales et produit des SNL comportant deux têtes lexicales.

Ainsi, sont présentées des analyses de variance pour *la proportion de l'ancienne information sous forme pronominale et sous forme lexicale* puis nous effectuons des analyses en prenant en considération les positions syntaxiques.

Dans la seconde partie (Section 12.2, p. 277) sont présentés les résultats relatifs aux structures des substitutions lexicales.¹¹³ Dans l'exemple (150), le scripteur expert fait preuve de flexibilité dans la mesure où il varie les formes faisant référence à un même référent. Ainsi, l'auteur introduit le référent sous la mention *Nicolas Sarkozy*, poursuit avec *le ministre de l'Intérieur*, utilise de nouveau *Nicolas Sarkozy* puis termine en encodant cette même entité sous la forme d'un SNL encore différent soit avec le SNL *le président de l'UMP*.

Dans notre étude, ont été distinguées les anciennes informations reprenant un référent déjà mentionné sous la forme exacte de la précédente mention c'est-à-dire sous forme de répétition (151) des anciennes informations reprenant un référent sous une forme différente de sa précédente mention (152).

(151) **Mon père** les a injuriés. Une demi-heure plus tard, **mon père** est parti. (CM2 n°19, narratif écrit, E/O)

(152) **Un garçon** l'embêtait. Donc je suis allé voir **ce garçon**, **l'enfant** disait que c'était mon frère. (5^{ème} n° 08, narratif écrit, E/O)

Sous la catégorie ancienne information lexicale avec modifications, nous avons différencié les modifications portant sur le déterminant du SNL (152) (**un** garçon ~ **ce** garçon), sur les modificateurs des SNL (153) (*un foot* Ø ~ *du foot de rue*) ou sur le nom du SNL (152) (**ce garçon** ~ **l'enfant**).

(153) Hier j'ai fait **un foot** avec sept pelos. Au début ça joue normal mais après quelques buts ça bloque ou ça met des coups vu que c'était **du foot de rue** il n'y avait pas d'arbitre. (3^{ème} n°38, narratif écrit, E/O)

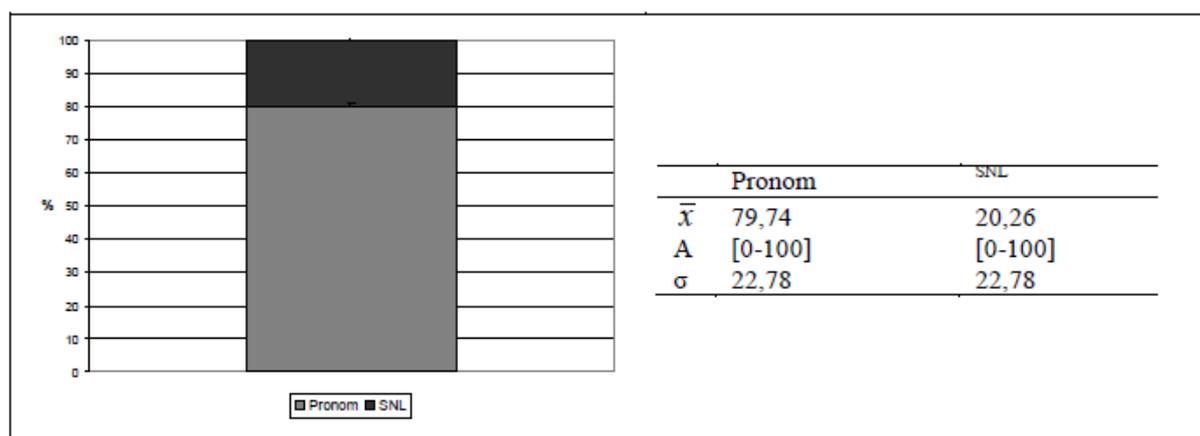
¹¹³ Nous incluons les réintroductions sous formes lexicales.

12.1. Maintien de l'information : SNL ou pronoms ?

12.1.1. Analyses

Les proportions *des SN maintenant une information sous forme pronominale et lexicale* (Figure 100) sont calculées par rapport au nombre total de SN maintenant une information.¹¹⁴ Un référent, une fois introduit, est préférentiellement maintenu par le biais d'une forme pronominale que par une forme lexicale (Test T : $t_{(503)} = 29,31$, $p < 0,0001$).

Figure 100 : Proportions des SN ancienne information sous forme pronominale et lexicale



Les figures 101 à 106 présentent *les proportions de l'ancienne information sous forme pronominale et lexicale* en fonction des facteurs de notre étude. Les différences selon le NIVEAU SCOLAIRE ($F_{(2,480)} = 3,185$, $p < 0,0423$), la MODALITÉ ($F_{(1,480)} = 37,023$, $p < 0,0001$) et l'ORDRE ($F_{(1,480)} = 3,563$, $p < 0,05$) sont significatives. Le facteur TYPE DE TEXTE n'affecte pas *la proportion de l'ancienne information sous forme pronominale et lexicale*.

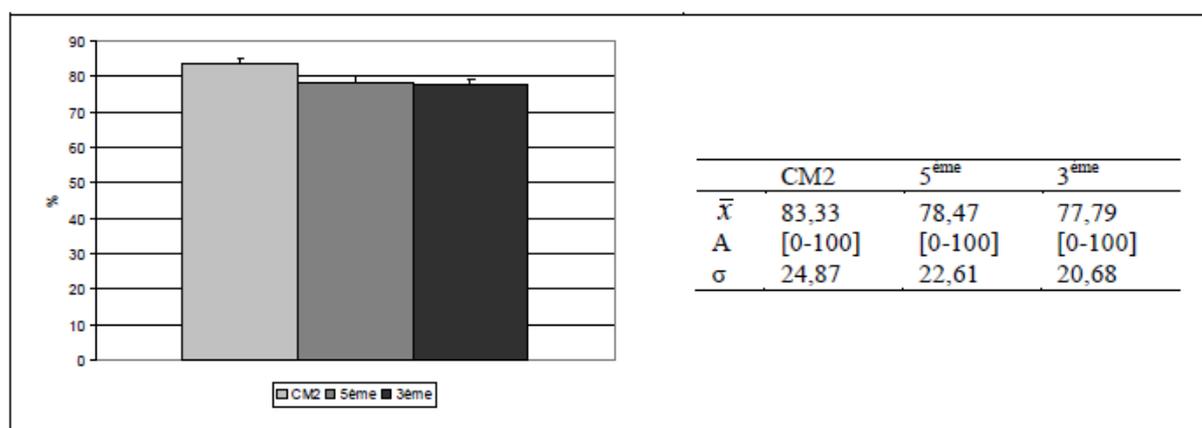
Les figures 101 à 103 présentent *les proportions de l'ancienne information sous forme pronominale* selon le NIVEAU SCOLAIRE, la MODALITÉ et l'ORDRE.

La proportion des SN maintenant une information sous forme pronominale diminue avec le NIVEAU SCOLAIRE. Les écoliers de CM2 ont une proportion de *SN maintenant une information sous forme pronominale* plus importante que les individus de 5^{ème} ($p = 0,0427$) et de

¹¹⁴ Ceci concerne 504 textes sur 528.

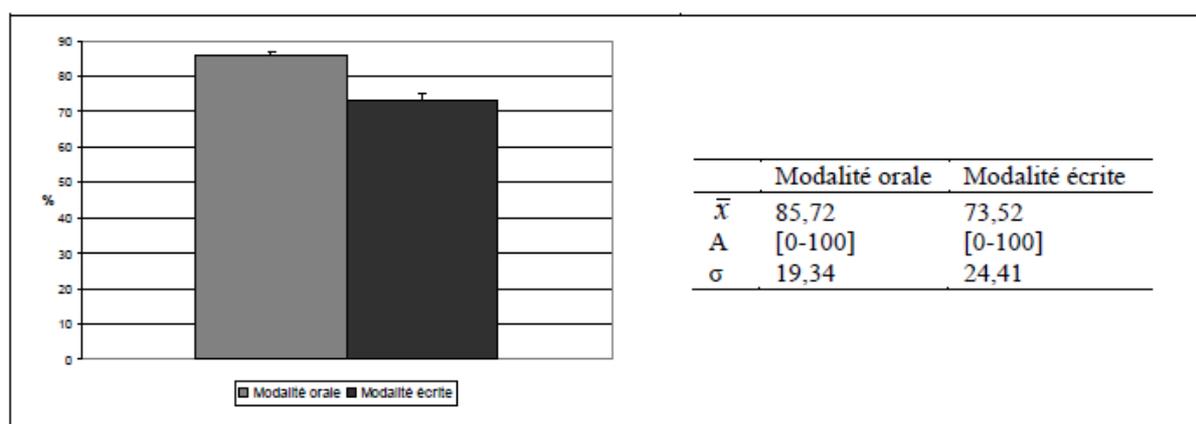
3^{ème} ($p < 0,0190$). Ces deux dernières populations ont des proportions ne se distinguant pas significativement ($p = 0,7677$) (Figure 101).

Figure 101 : Proportions des SN pronominaux maintenant une information selon le NIVEAU SCOLAIRE (par rapport au nombre total de SN maintenant une information)



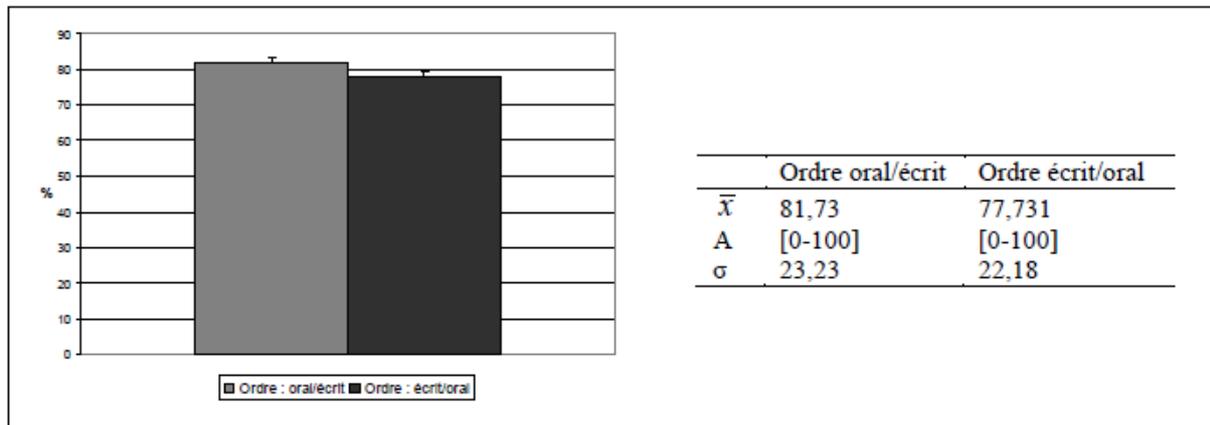
La proportion de l'ancienne information sous forme pronominale et lexicale varie selon la MODALITÉ : l'utilisation de pronoms pour maintenir un référent constitue davantage une caractéristique de la modalité orale que de la modalité écrite (Figure 102).

Figure 102 : Proportions des SN pronominaux maintenant une information selon la MODALITÉ (par rapport au nombre total de SN maintenant une information)



De plus, l'ORDRE intervient : les individus utilisent davantage les pronoms pour maintenir une information lorsqu'ils appartiennent à l'ordre de production oral/écrit (Figure 103).

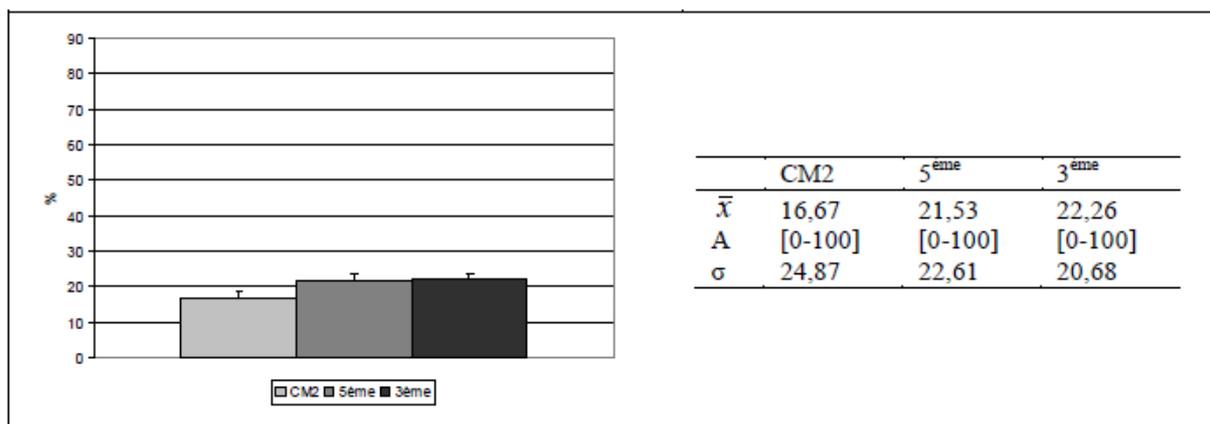
Figure 103 : Proportions des SN pronominaux maintenant une information selon l'ORDRE DE PRODUCTION (par rapport au nombre total de SN maintenant une information)



Les figures 104 à 106 présentent les proportions de l'ancienne information sous forme pronominale selon le NIVEAU SCOLAIRE, la MODALITÉ et l'ORDRE.

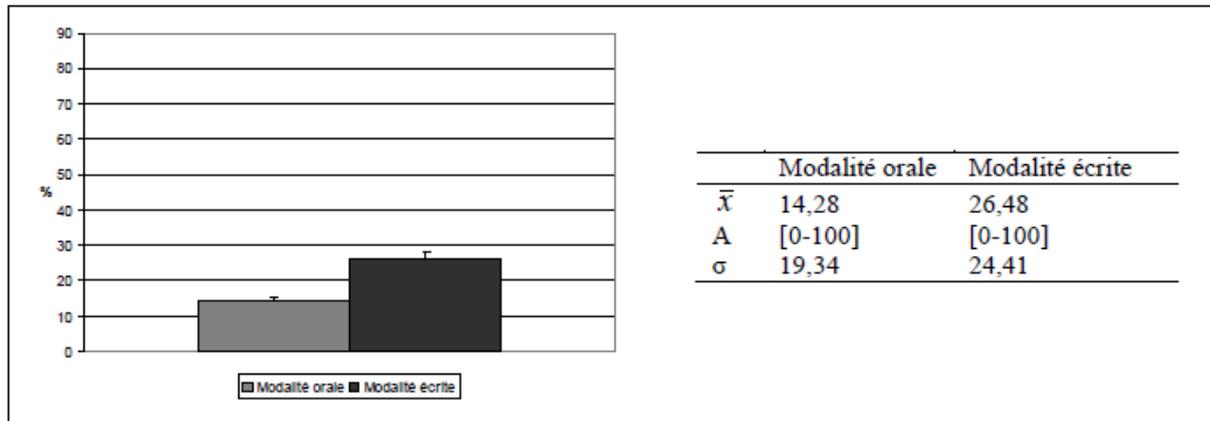
Les proportions des SNL maintenant une information augmente avec le NIVEAU SCOLAIRE. Les écoliers de CM2 ont une proportion de SN maintenant une information sous forme pronominale plus faible que les individus de 5^{ème} ($p = 0,0427$) et de 3^{ème} ($p < 0,0190$). Ces deux dernières populations ont des proportions ne se distinguant pas significativement ($p = 0,7677$) (Figure 104).

Figure 104 : Proportions des SNL maintenant une information selon le NIVEAU SCOLAIRE (par rapport au nombre total de SN maintenant une information)



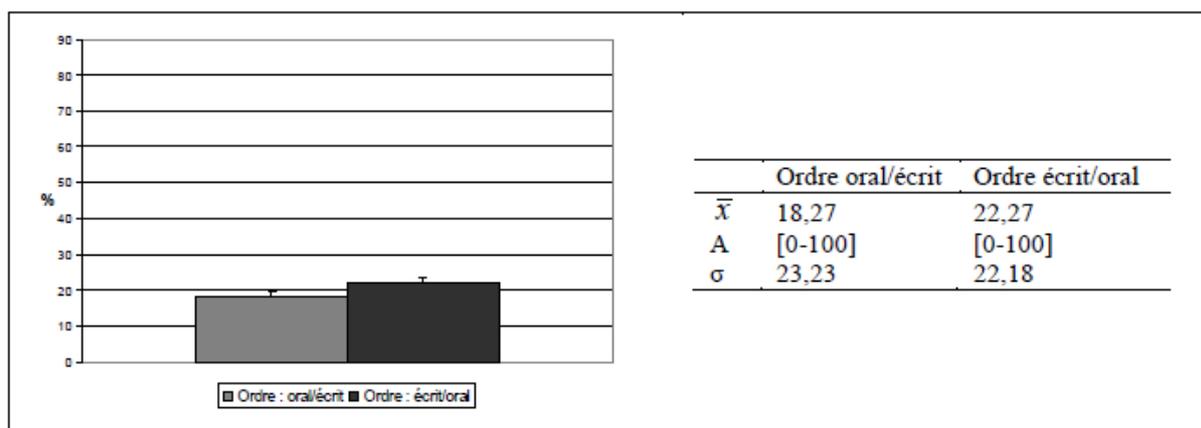
La proportion de SNL pour maintenir un référent constitue davantage une caractéristique de la modalité écrite que de la modalité orale (Figure 105).

Figure 105 : Proportions des SNL maintenant une information selon la MODALITÉ (par rapport au nombre total de SN maintenant une information)



De plus, la proportion de SNL pour maintenir un référent est plus importante dans les textes des individus qui appartiennent à l'ordre de production écrit/oral que dans les textes des individus qui appartiennent au groupe de passation oral/écrit (Figure 106).

Figure 106 : Proportions des SNL maintenant une information selon l'ORDRE DE PRODUCTION (par rapport au nombre total de SN maintenant une information)



12.1.2. Discussion de la Section 12.1

Le but de cette partie était d'observer *la proportion des SN maintenant une information* en nous penchant sur la forme que prend ce maintien : pronom ou SNL. Un second intérêt était de voir si *la proportion des SN maintenant une information sous forme pronominale* et *la proportion des SN maintenant une information sous forme lexicale* variaient selon les facteurs de notre étude – NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE –.¹¹⁵

Les résultats révèlent que la forme la plus sollicitée pour encoder un référent déjà introduit dans un texte est le pronom. Ce résultat est trivial dans la mesure où il confirme de nombreuses théories et données allant dans ce sens (Ariel, 1990, 1996 ; Chafe, 1994 ; Fox, 1987 ; Jones, 1996 ; Khorounjaia *et al.*, 2004). Les pronoms sont utilisés préférentiellement pour encoder un référent de haute accessibilité ; en revanche, les SNL encodent préférentiellement des référents de basse accessibilité.

Une de nos prédictions était un effet significatif du NIVEAU SCOLAIRE : avec les années de scolarisation, les individus de notre étude maintiendraient de plus en plus une information sous forme lexicale. Ceci est confirmé par nos résultats. Les individus développent leur répertoire, ils utilisent alors une nouvelle forme, *le SNL*, pour une ancienne fonction, *maintenir une information*. Rappelons que les écoliers de CM2 (154) maintiennent significativement moins l'ancienne information sous forme lexicale que les deux populations de collègue (155) ; ces deux derniers groupes ne se distinguent guère.

(154) *C'était avec **un copain**, j'avais ramené **un cahier**, il a voulu **le** prendre. (CM2 n°42, narratif écrit, E/O)*

(155) *Elle n'avait pas vu qu'elle avait perdu **ses sous**. Moi, j'ai ramassé **l'argent** et je lui ai rendu **ses sous**. (5ème n°47, narratif écrit, E/O)*

Dans ces exemples, l'écolier introduit les nouvelles informations *un copain* et *un cahier*, puis les maintient par les pronoms *il* et *le*. Le collégien introduit les référents *ses sous* puis le maintient deux fois sous forme pronominale, *l'argent* et *ses sous*.

Le choix naturel et le plus économique, qui est lié aux principes de transmission de l'information de l'oral, est l'encodage sous forme pronominal (Ariel, 1990 et 1996 ; Fox, 1987 ; Khorounjaia et Tolchinsky, 2004). Avec l'expérience textuelle et le perfectionnement de l'écrit, les enfants répondent de moins en moins à ce précepte. Nous avons vu, aux

¹¹⁵ Les attentes, hypothèses et problématiques sont davantage explicitées au Chapitre 3 (Section 3.2.2.5, p. 105).

chapitres 8 et 9, que produire des SNL exige un coût cognitif important et notamment lorsqu'il faut les mettre en position sujet. Ainsi, produire un SNL pour maintenir une information exige des individus (a) non seulement un effort pour l'accès lexical et (b) d'aller à l'encontre d'un principe bien connu et intégré : l'encodage de l'ancienne information sous forme pronominale.

Une autre de nos attentes était un effet significatif de la MODALITÉ ce qui est confirmé par nos résultats. Maintenir un référent par le biais d'un pronom constitue plus une caractéristique des textes produits à l'oral (156) alors que maintenir un référent grâce à un SNL correspond davantage aux textes produits à l'écrit (157). Le fait de maintenir une information sous forme lexicale exige d'aller à l'encontre des principes naturels de la transmission de l'information, et notamment ceux de l'oral. Ceci implique alors un effort cognitif important qui est facilité en production écrite par son temps de planification important. De plus, la production de SNL est beaucoup plus aisée à l'écrit qu'à l'oral ; en effet, la production de SNL est coûteuse et les individus trouvent davantage de ressources à l'écrit. Les exemples ci-dessous représentatifs de cette situation sont proposés.

(156) *J'avais changé un peu radicalement **de style vestimentaire**. On s'est revu voilà il a vu que mentalement j'avais pas changé.* (3^{ème} n°27, narratif oral, O/E)

(157) *Il y a quelques années, il m'est arrivé une dispute avec mon oncle à cause de **mon changement vestimentaire**. [...] Il avait fait une erreur en s'arrêtant sur **mon nouveau style vestimentaire**.* (3^{ème} n°27, narratif écrit, O/E)

Le collégien introduit à l'écrit une nouvelle information *mon changement vestimentaire* qu'il maintient par le biais du SNL *mon nouveau style vestimentaire*. Ce même individu introduit à l'oral également un nouveau SNL, *style vestimentaire*, mais ne le maintient pas.

Pour finir, l'ORDRE DE PRODUCTION affecte également *la proportion de l'ancienne information sous forme pronominale et lexicale*, ce qui va dans le sens de nos hypothèses. Le maintien de l'information par le biais d'un pronom est plus fréquent dans les textes d'individus ayant produit dans l'ordre oral/écrit. En revanche, la proportion du maintien grâce à une forme lexicale est plus importante dans les textes des individus ayant produit dans l'ordre écrit/oral (158) et (159).

(158) Alors un jour, elle a lu **mes S_M_S personnels**. Alors pendant une période j'avais peur que mes parents prennent mon téléphone pour lire **mes messages**. (3^{ème} n°14, narratif écrit, E/O)

(159) Ben entre temps elle a vu **mes messages** et tout et donc elle a lu **des messages euh personnels**. (3^{ème} n°14, narratif oral, E/O)

L'utilisation de formes lexicales pour maintenir un référent (*mes SMS personnels ~ mes messages*), caractéristique de l'écrit, se trouve transférée à l'oral (*mes messages ~ des messages*).

Le fait de produire un texte écrit puis un texte oral favorise le maintien sous forme lexicale étant donné que l'écrit facilite la proportion de SNL ; à l'écrit, les individus ont le temps d'aller au-delà des principes de transmission de l'information de l'oral. Une fois cette stratégie (maintenir un SNL par un autre SNL) activée et appliquée à l'écrit, elle est transférée à l'oral. Néanmoins, si le sujet commence par faire un texte oral, il tend à respecter les tendances naturelles du traitement de l'information en maintenant l'ancienne information sous forme pronominale dans la mesure où la modalité orale ne permet pas de les transgresser. Cette attitude se répercute en production écrite.

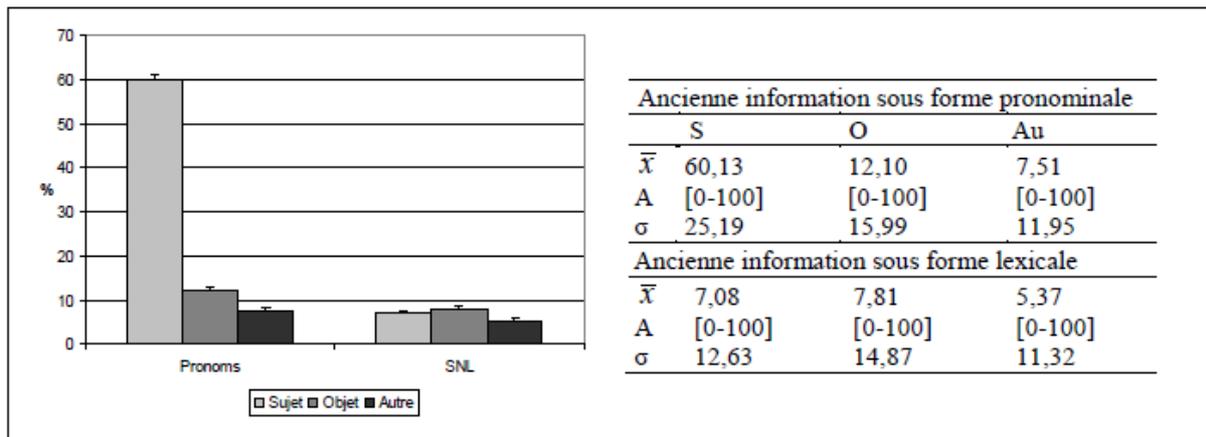
Ces analyses nous révèlent des faits intéressants et confirment bon nombre de nos hypothèses. Mais qu'en est-il si nous observons *la proportion des SN maintenant une information sous forme pronominale et lexicale* en prenant en considération les positions syntaxiques ? Est-ce que les analyses en prenant en considération la fonction syntaxique peuvent compléter les résultats, comme cela l'a été dans les chapitres précédents ? C'est dans cette mesure que nous allons, dans la prochaine section, présenter les analyses des *proportions des SN maintenant une information sous forme pronominale et lexicale en position sujet, objet et autre*.

12.2. Proportion de l'ancienne information : formes et position syntaxiques

Dans cette partie est présentée *la proportion des SN maintenant une information sous forme lexicale et pronominale* selon les positions syntaxiques. Ces proportions sont toutes calculées par rapport au nombre total de SN maintenant une information.

Les informations regroupées en Figure 81 résument *les proportions des SN maintenant une information sous forme lexicale et sous forme pronominale* selon les positions syntaxiques.

Figure 107 : Proportions de l'ancienne information selon sa forme et sa position syntaxique (par rapport au nombre total de SN maintenant une information)



Ainsi :

- 60,13% des SN maintenant une information sont sous forme pronominale en position sujet ;
- 12,10% des SN maintenant une information sont sous forme pronominale en position objet ;
- 7,51% des SN maintenant une information sont sous forme pronominale en position autre ;
- 7,08% des SN maintenant une information sont sous forme pronominale en position sujet ;
- 7,81% des SN maintenant une information sont sous forme pronominale en position objet ;
- 5,37% des SN maintenant une information sont sous forme pronominale en position autre.

Ces résultats suggèrent que l'ancienne information est davantage en position sujet qu'en position objet et autre et est encodée plus sous forme pronominale que sous forme lexicale (ce qui confirme les résultats précédents). La différence entre la proportion d'ancienne

information sous forme lexicale et celle d'ancienne information sous forme pronominale est très importante pour la position sujet. Ceci n'est pas le cas pour les deux autres positions.

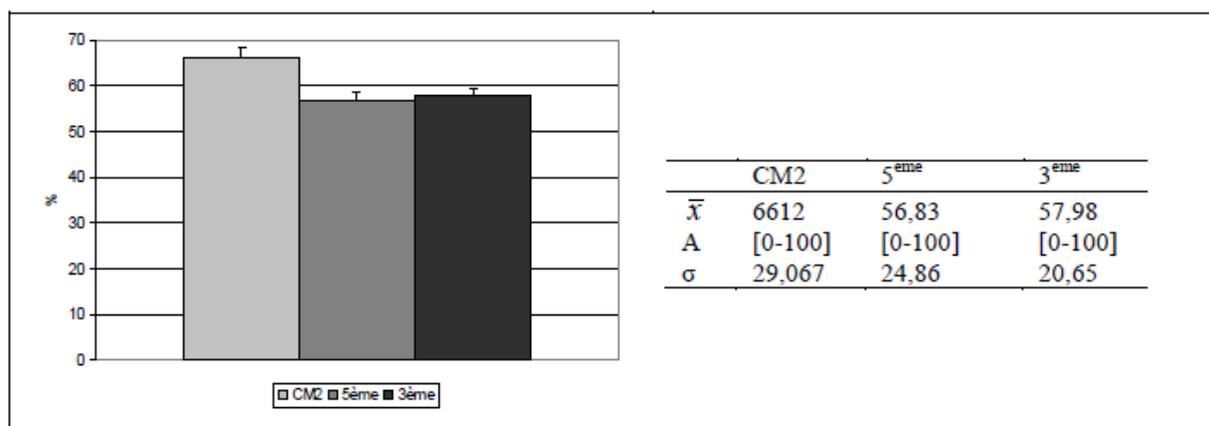
12.2.1. Maintien de l'information sous forme pronominale selon les positions syntaxiques

12.2.1.1. Proportion de l'ancienne information sous forme pronominale en position sujet

Les figures 108 à 110 présentent *les proportions des SN maintenant une information sous forme pronominale en position sujet* en fonction du NIVEAU SCOLAIRE, du TYPE DE TEXTE et de la MODALITÉ.

Les différences en fonction du NIVEAU SCOLAIRE sont significatives ($F_{(2,480)} = 7,473$, $p < 0,0006$). *La proportion des SN maintenant une information sous forme pronominale en position sujet* diminue avec le NIVEAU SCOLAIRE (Figure 108). Les écoliers de CM2 ont une proportion plus élevée que les individus de 5^{ème} ($p = 0,0006$) et de 3^{ème} ($p = 0,0021$). Les proportions des deux populations de collège sont relativement proches et la différence entre les deux est non significative ($p = 0,6579$).

Figure 108 : Proportions des SN maintenant une information sous forme pronominale en position sujet selon le NIVEAU SCOLAIRE (par rapport au nombre total de SN maintenant une information)



Le contexte de production a également un impact significatif sur *la proportion des SN maintenant une information sous forme pronominale en position sujet* (TYPE DE TEXTE

($F_{(1,480)} = 9,833$, $p < 0,0018$) et MODALITÉ ($F_{(1,480)} = 20,574$, $p < 0,0001$). Ce type de maintien en position sujet est également plus fréquent dans les textes expositifs et oraux que dans les textes narratifs et écrits (Figure 109 et Figure 110).

Figure 109 : Proportions des SN maintenant une information sous forme pronominale en position sujet selon le TYPE DE TEXTE (par rapport au nombre total de SN maintenant une information)

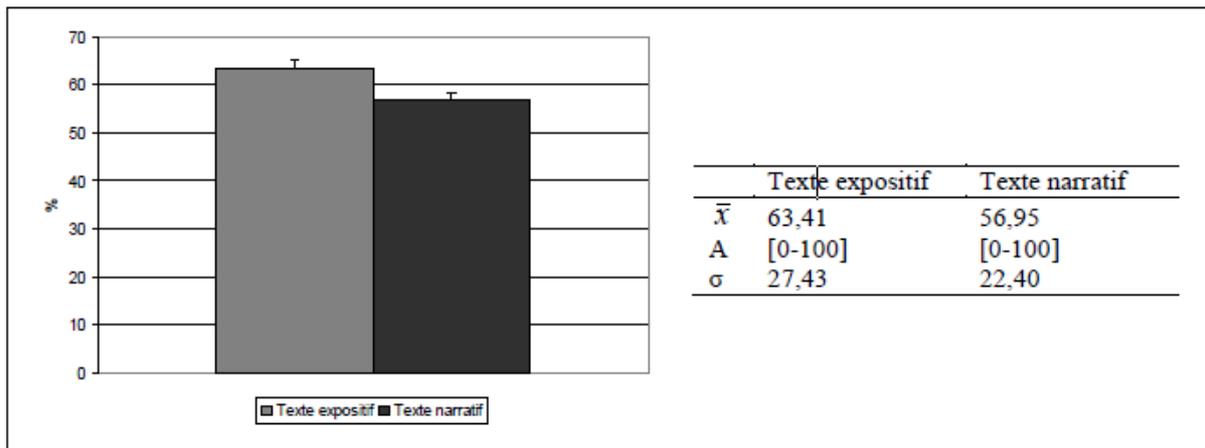
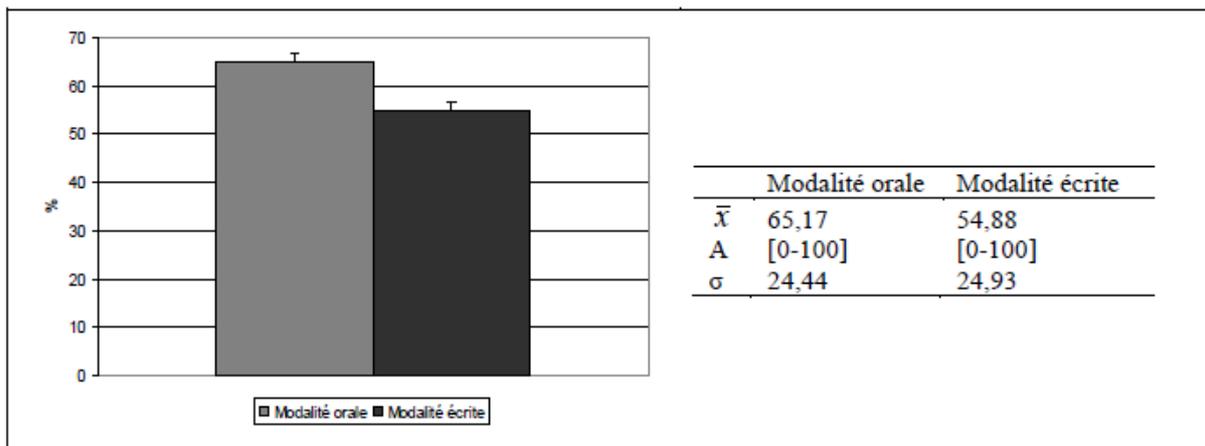


Figure 110 : Proportions des SN maintenant une information sous forme pronominale en position sujet selon la MODALITÉ (par rapport au nombre total de SN maintenant une information)



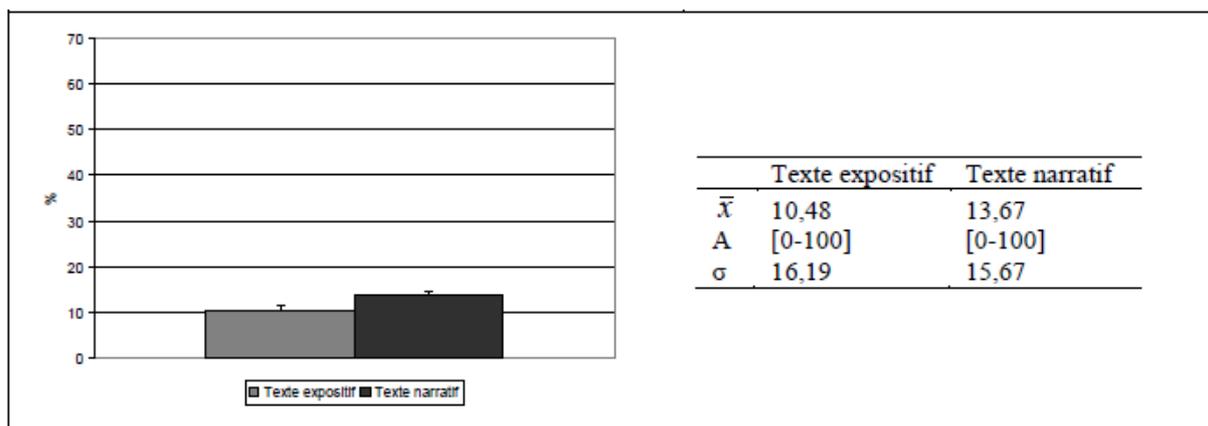
L'ORDRE n'affecte pas significativement *la proportion des SN maintenant une information sous forme pronominale en position sujet*.

12.2.1.2. Proportion de l'ancienne information sous forme pronominale en position objet

La Figure 111 présente *les proportions des SN maintenant une information sous forme pronominale en position objet* en fonction du TYPE DE TEXTE.

Les différences en fonction du TYPE DE TEXTE sont significatives ($F_{(1,480)} = 5,189, p < 0,0232$). Les autres variables n'affectent pas cette VD. *La proportion des SN maintenant une information sous forme pronominale en position objet* a une proportion plus importante dans les textes narratifs que dans les textes expositifs (Figure 92).

Figure 111 : Proportions des SN maintenant une information en position objet selon le TYPE DE TEXTE (par rapport au nombre total de SN maintenant une information)

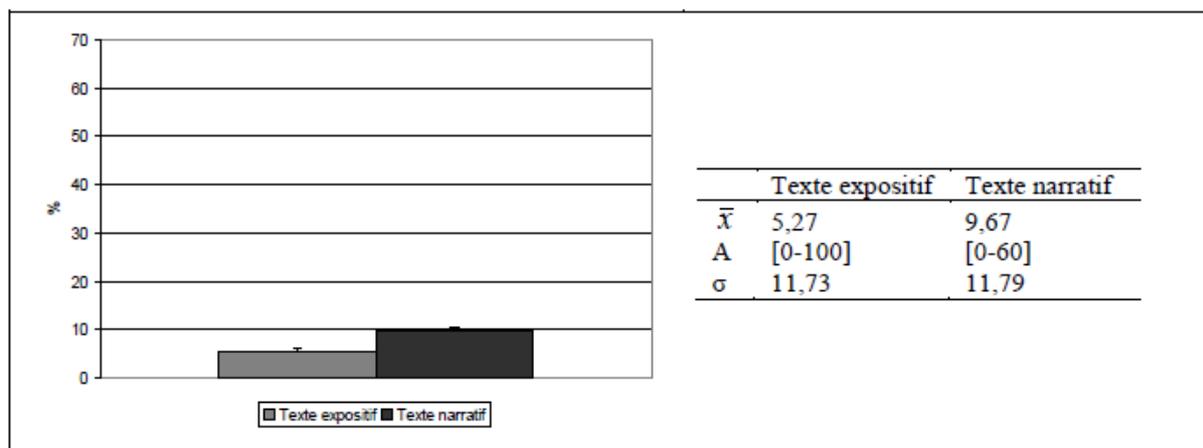


12.2.1.3. Proportion de l'ancienne information sous forme pronominale en position autre

La figure 112 présente *les proportions des SN maintenant une information sous forme pronominale en position autre* selon le TYPE DE TEXTE.

Seules les différences en fonction du TYPE DE TEXTE sont significatives ($F_{(1,480)} = 5,189, p < 0,0232$). Les autres variables n'affectent pas *la proportion des SN maintenant une information sous forme pronominale en position autre* qui est plus importante dans les textes narratifs que dans les textes expositifs (Figure 112).

Figure 112 : Proportions des SN maintenant une information en position autre selon le TYPE DE TEXTE (par rapport au nombre total de SN maintenant une information)



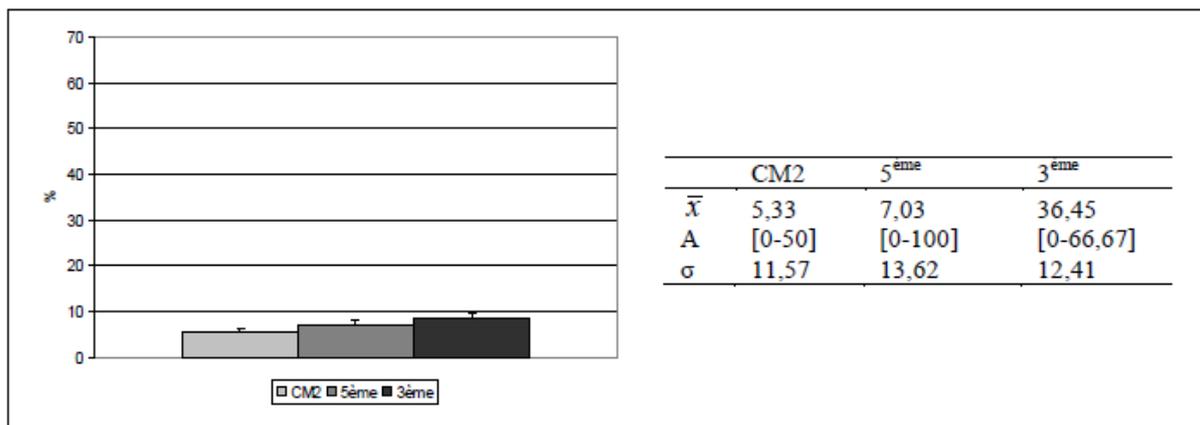
12.2.2. Maintien de l'information sous forme lexicale selon les positions syntaxiques

12.2.2.1. Proportion de l'ancienne information sous forme lexicale en position sujet

Les figures 113 à 115 présentent *les proportions des SN maintenant une information sous forme lexicale en position sujet* en fonction du NIVEAU SCOLAIRE, de la MODALITÉ et de l'ORDRE.

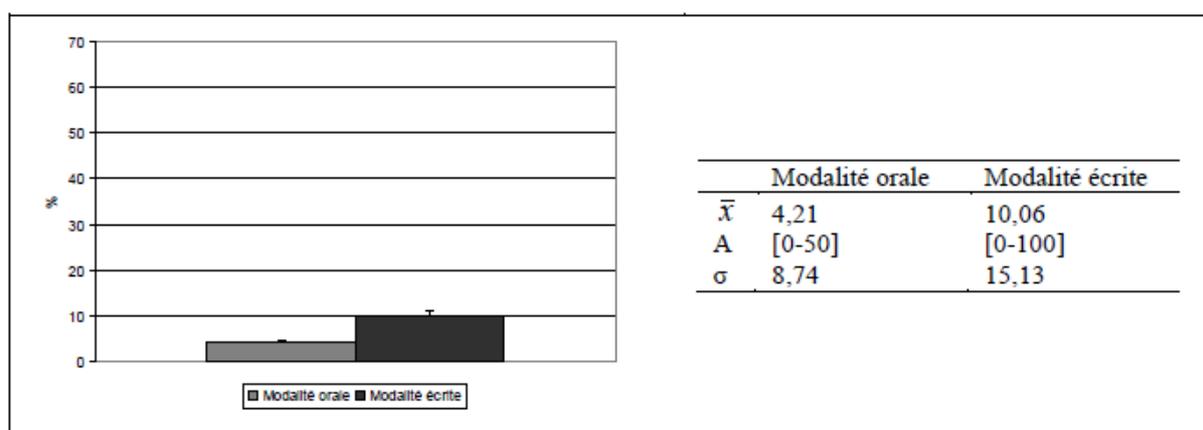
Les différences en fonction du NIVEAU SCOLAIRE ($F_{(2,480)} = 3,148, p < 0,0438$) sont significatives. *La proportion des SN maintenant une information sous forme lexicale en position sujet* augmente avec le NIVEAU SCOLAIRE (Figure 113). Seules les populations extrêmes se distinguent significativement ; les écoliers de CM2 ont une proportion moins élevée les individus de 3^{ème} ($p = 0,0120$).

Figure 113 : Proportions des SN maintenant une information sous forme lexicale en position sujet selon le NIVEAU SCOLAIRE (par rapport au nombre total de SN maintenant une information)



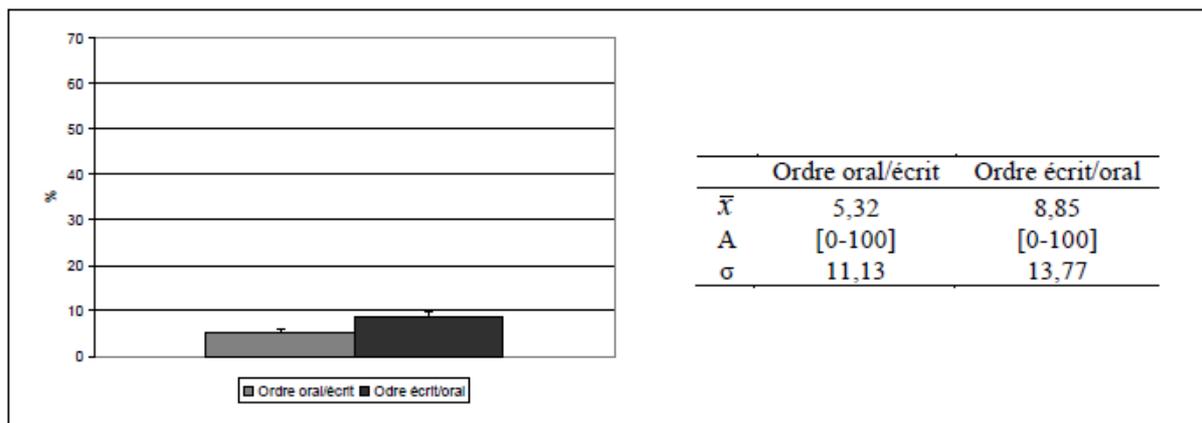
La MODALITÉ fait également varier *la proportion des SN maintenant une information sous forme lexicale en position sujet* ($F_{(1,480)} = 26,461$, $p < 0,0001$) : elle est plus importante dans les textes écrits que dans les textes oraux (Figure 114).

Figure 114 : Proportions des SN maintenant une information sous forme lexicale en position sujet selon la MODALITÉ (par rapport au nombre total de SN maintenant une information)



L'ORDRE a aussi un effet sur *la proportion des SN maintenant une information sous forme lexicale en position sujet* ($F_{(1,480)} = 10,039$, $p < 0,0016$). Ce type de maintien en position sujet est également plus fréquent dans les textes des individus ayant produit dans l'ordre de production écrit/oral (Figure 115).

Figure 115 : Proportions des SN maintenant une information sous forme lexicale en position sujet selon l'ORDRE (par rapport au nombre total de SN maintenant une information)



Le TYPE DE TEXTE n'affecte significativement pas la proportion des SN maintenant une information sous forme lexicale en position sujet.

12.2.2.2. Proportion de l'ancienne information sous forme lexicale en position objet

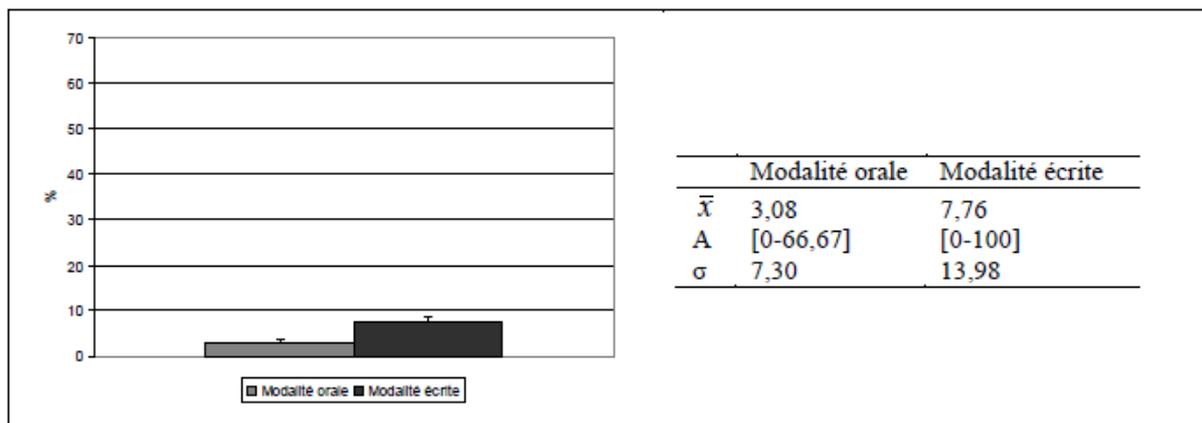
L'analyse de variance montre que, pour *la proportion des SN maintenant une information sous forme lexicale en position objet*, les différences en fonction du NIVEAU SCOLAIRE, du TYPE DE TEXTE, de la MODALITÉ et de l'ORDRE ne sont pas significatives.

12.2.2.3. Proportion de l'ancienne information sous forme lexicale en position autre

La Figure 89 présente *les proportions des SN maintenant une information sous forme lexicale en position autre* en fonction du NIVEAU SCOLAIRE.

Seules les différences en fonction de la MODALITÉ ($F_{(1,480)} = 4,254, p < 0,0397$) sont significatives ; les trois autres facteurs – NIVEAU SCOLAIRE, TYPE DE TEXTE et ORDRE – ne font pas varier cette VD. *La proportion des SN maintenant une information sous forme lexicale en position autre* est plus importante dans les textes écrits que dans les textes oraux (Figure 89).

Figure 116 : Proportions des SN maintenant une information sous forme lexicale en position autre selon la MODALITÉ (par rapport au nombre total de SN maintenant une information)



12.2.3. Discussion de la Section 12.2

Nous faisons l’hypothèse que l’observation des deux variables – *la proportion des SN maintenant une information sous forme lexicale* et *sous forme pronominale* – selon les positions syntaxiques (Section 12.2, p. 277) complèterait les résultats que nous trouvons lors des analyses ne prenant pas en considération la position syntaxique (Section 12.1, p.270).¹¹⁶ Ceci est confirmé par nos analyses.

Les statistiques descriptives (Figure 107) révèlent que les SN maintenant une information sont préférentiellement en position sujet et sous forme pronominale (60,13%), ce qui va dans le sens des analyses précédentes (Section 12.1, p. 270). Le nouvel aspect à souligner est le fait qu’en position objet et autre les individus maintiennent pratiquement autant sous forme pronominale (respectivement 12,10% et 7,51%) que sous forme lexicale (respectivement 7,81% et 5,37%). En revanche, en position sujet, la différence entre les deux types de maintiens est beaucoup plus importante (60,13% sous forme pronominale et 7,08% sous forme lexicale).

Rappelons que *les proportions des maintiens sous forme lexicale* et *des maintiens sous forme pronominale* toutes positions confondues varient par rapport aux quatre variables indépendantes : NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE. *La proportion des maintiens sous forme lexicale* toutes positions confondues augmente avec le NIVEAU SCOLAIRE, et est plus élevée dans les textes écrits et plus importante dans les textes produits par des individus appartenant à l’ordre de passation écrit/oral. En revanche, *la proportion des maintiens sous forme pronominale* toutes positions confondues diminue avec le NIVEAU SCOLAIRE, et est plus

¹¹⁶ Les attentes, hypothèses et problématiques sont davantage explicitées au Chapitre 3 (Section 3.2.2.5, p. 105).

élevée dans les textes oraux et plus importante dans les textes produits par des individus appartenant à l'ordre de passation oral/écrit.

L'analyse des données montre que le NIVEAU SCOLAIRE affecte seulement *la proportion des maintiens sous forme lexicale et sous forme pronominale en position sujet*. *La proportion des maintiens sous forme lexicale en position sujet* augmente avec les années de scolarisation. En revanche, *la proportion des maintiens sous forme pronominale en position sujet* diminue avec les années de scolarisation : le maintien d'un référent se fait préférentiellement sous forme pronominale en position sujet. La substitution lexicale en position sujet implique d'aller à l'encontre d'un des principes de transmission de l'information, les formes lexicales sont préférentiellement en position non sujet. L'enfant acquiert de l'expérience, qui lui permet de maintenir un référent (a) sous forme lexicale et (b) en position sujet.

La proportion des maintiens sous forme lexicale et sous forme pronominale selon leurs positions syntaxiques varient également significativement selon le TYPE DE TEXTE.

Les analyses prenant en considération les positions syntaxiques dévoilent que *les proportions des maintiens sous forme lexicale en position sujet, objet et autre* ne varient pas selon ce facteur. Néanmoins, si *la proportion des maintiens sous forme pronominale en position sujet* est plus importante dans les textes expositifs, *les proportions des maintiens sous forme pronominale en position objet et autre* sont plus élevées dans les textes narratifs.

Les proportions des maintiens sous forme lexicale et des maintiens sous forme pronominale selon leurs positions syntaxiques varient également selon la MODALITÉ DE PRODUCTION. Les analyses distinguant les positions syntaxiques mettent en avant que *les proportions des maintiens sous forme lexicale en position sujet et autre* sont plus importantes dans les textes écrits. En revanche, la proportion *la proportion des maintiens sous forme pronominale en position sujet* est plus élevée dans les productions orales.

L'ORDRE DE PRODUCTION affecte seulement une variable : *la proportion des maintiens sous forme lexicale en position sujet* qui est plus élevée dans les productions produites par des individus appartenant à l'ordre écrit/oral.

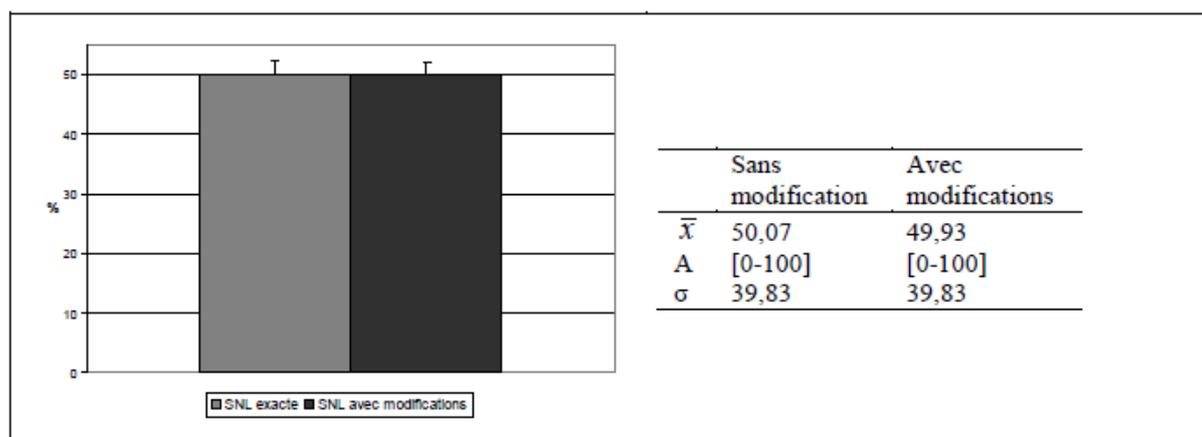
La substitution lexicale, surtout en position sujet, s'acquiert avec le temps et l'expérience, il s'agit d'une acquisition assez tardive. Nous pouvons alors nous intéresser plus spécifiquement au maintien lexical en nous penchant sur l'architecture des SNL permettant ce maintien.

12.3. Le maintien sous forme lexicale

12.3.1. Analyses

La Figure 117 présente les proportions des SNL maintenant une ancienne information de même forme que la mention précédente et des SNL maintenant une ancienne information mais impliquant des modifications par rapport à la mention précédente selon le nombre total des SNL maintenant une information.¹¹⁷ La différence entre ces deux types de maintien n'est pas significative (Test T : $t_{(331)} = 0,03$, $p = 0,9735$).

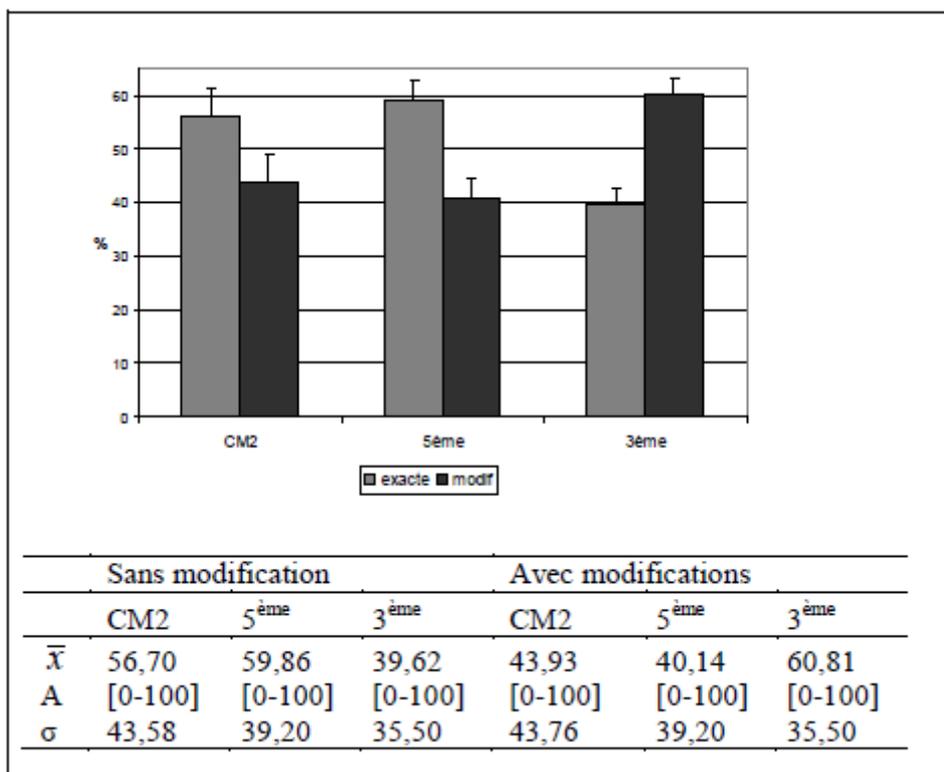
Figure 117 : Proportions des SNL maintenant une information impliquant une répétition ou des modifications linguistiques



Or, en introduisant le facteur NIVEAU SCOLAIRE, le maintien lexical par le biais d'une répétition ou par le biais de modifications linguistiques varie. La Figure 118 montre que si les écoliers de CM2 et les individus de 5^{ème} produisent davantage de SNL ancienne information sans modification, les collégiens de 3^{ème}, eux, maintiennent davantage une information sous forme lexicale avec modifications.

Figure 118 : Proportions des SNL maintenant une information impliquant une répétition ou des modifications linguistiques selon le NIVEAU SCOLAIRE

¹¹⁷ Pour les 332 textes contenant un maintien lexical sur les 528 au total.



Ainsi, il peut être intéressant de voir comment *les proportions des SNL maintenant une ancienne information de même forme que la mention précédente et des SNL maintenant une ancienne information mais impliquant des modifications par rapport à la mention précédente* varient selon tous nos facteurs : NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE.

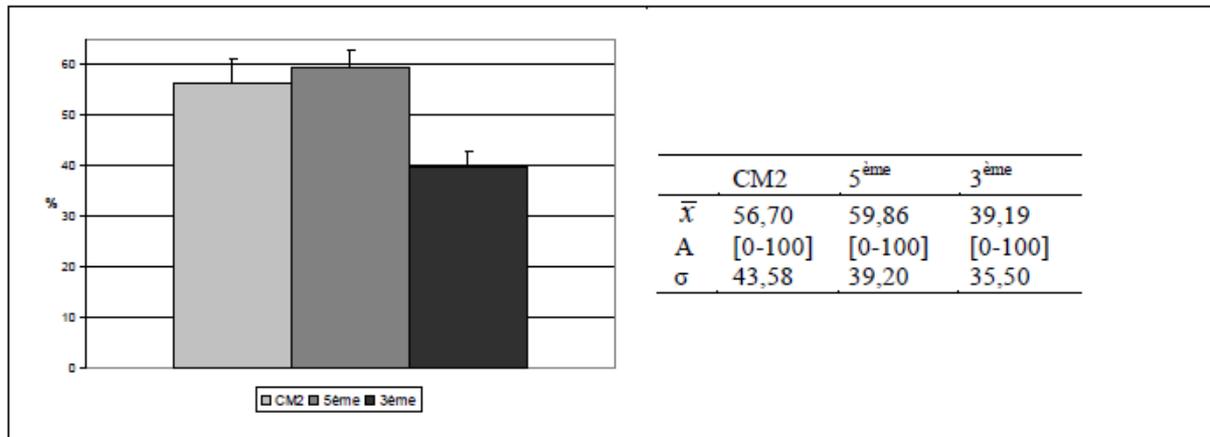
Les figures 119 à 122 présentent *les proportions des SNL ancienne information sans et avec modifications* selon le nombre total de SNL ancienne information en fonction du NIVEAU SCOLAIRE et la MODALITÉ.¹¹⁸ Les différences en fonction du NIVEAU SCOLAIRE ($F_{(2,308)} = 9,35$, $p < 0,0001$) et la MODALITÉ ($F_{(1,308)} = 10,33$, $p < 0,0013$). Les facteurs TYPE DE TEXTE et ORDRE n'ont pas d'effet significatif sur ces VD.

Les figures 119 et 120 présentent *la proportion des SNL maintenant une ancienne information de même forme que la mention précédente*.

Les individus scolarisés en CM2 ($p = 0,0031$) et en 5^{ème} ($p < 0,0001$) utilisent plus la répétition pour maintenir une information sous forme lexicale que les individus de 3^{ème} (Figure 119). La différence de proportions entre les CM2 et les 5^{ème} est minime ($p = 0,5870$).

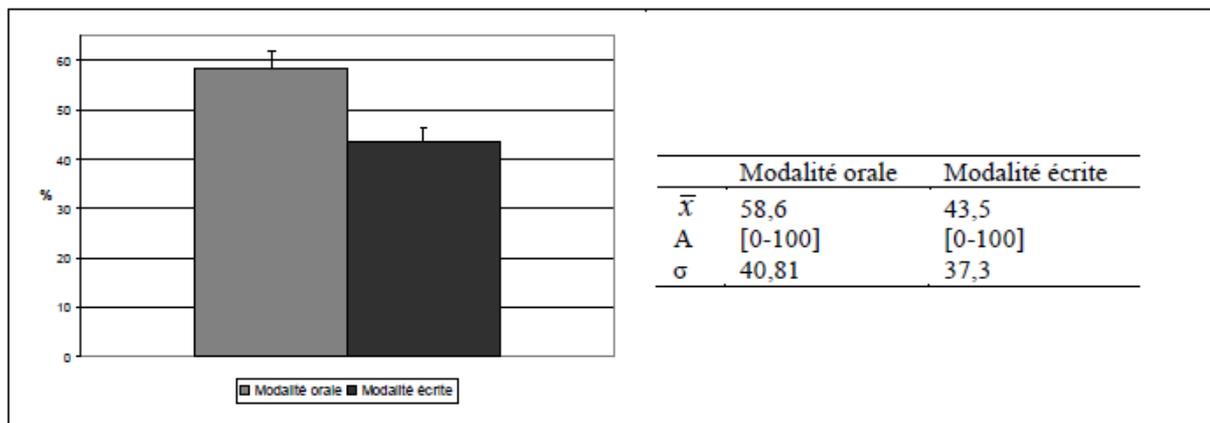
¹¹⁸ Pour les 332 textes contenant des SNL maintenant une information sur 528.

Figure 119 : Proportions des SNL maintenant une information n'impliquant pas de modification linguistique selon le NIVEAU SCOLAIRE (par rapport au nombre total de SNL maintenant une information)



La proportion des SNL ancienne information sans modification est plus élevée dans les productions orales que les productions écrites (Figure 120).

Figure 120 : Proportions des SNL maintenant une information n'impliquant pas de modification linguistique selon la MODALITÉ (par rapport au nombre total de SNL maintenant une information)

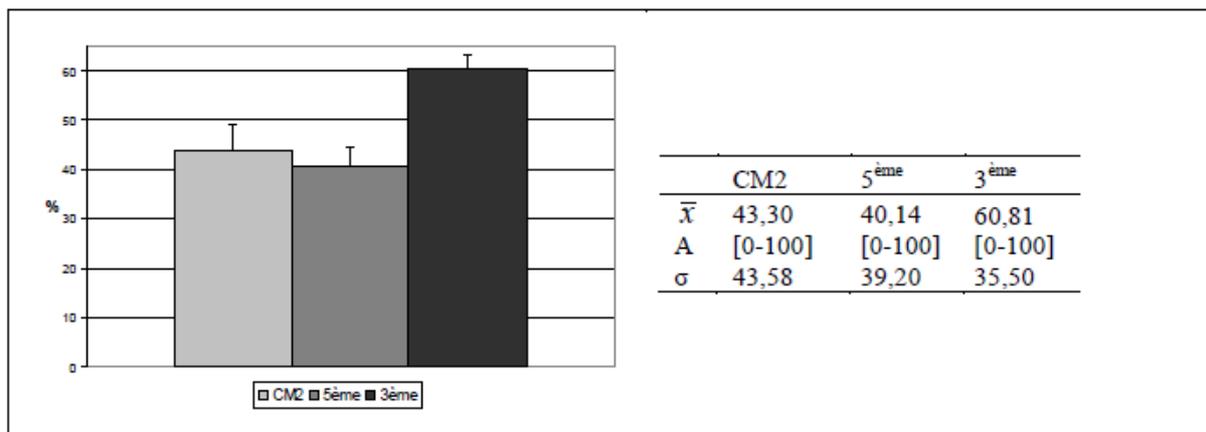


Les figures 121 et 122 présentent les proportions des SNL ancienne information avec modifications.

La proportion des SNL ancienne information avec modifications augmente avec les années de scolarisation. Les individus de 3^{ème}, lorsqu'ils maintiennent une information sous forme lexicale, le font davantage en effectuant des modifications linguistiques qu'en faisant des

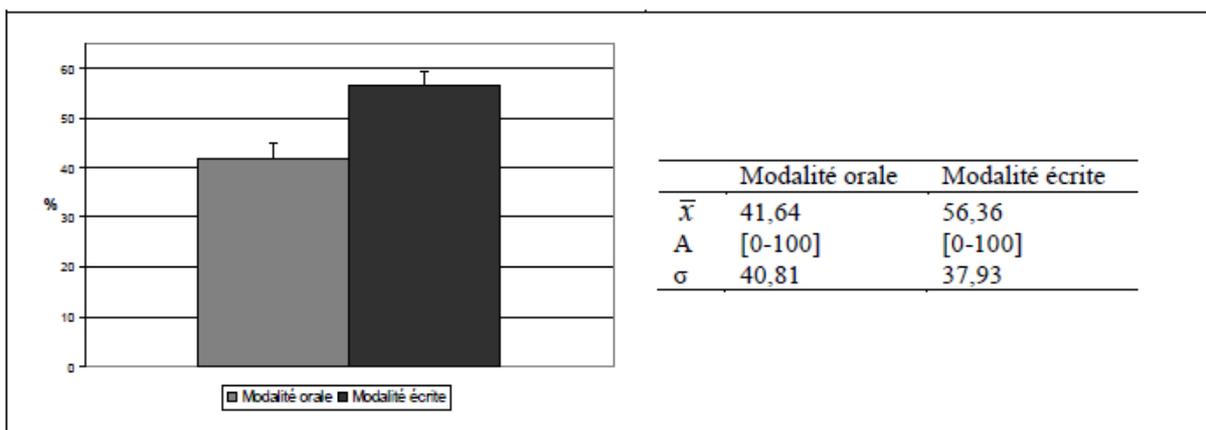
répétitions ; ils maintiennent significativement plus en faisant des modifications que les individus de 5^{ème} ($p < 0,0001$) et de CM2 ($p = 0,0031$) dont les deux proportions ne se distinguent pas significativement ($p = 0,5870$) (Figure 121).

Figure 121 : Proportions des SNL maintenant une information impliquant des modifications linguistiques selon le NIVEAU SCOLAIRE (par rapport au nombre total de SNL maintenant une information)



La proportion des SNL ancienne information avec modifications est plus importante dans les productions écrites que les productions orales (Figure 122).

Figure 122 : Proportions des SNL maintenant une information impliquant des modifications linguistiques selon la MODALITÉ (par rapport au nombre total de SNL maintenant une information)



12.3.2. Discussion de la Section 12.3

Dans cette partie, nous voulions observer *la proportion du maintien sous forme lexicale impliquant des changements linguistiques* ainsi que *celle n'impliquant pas de changement* et, notamment, voir si ces VD variaient selon les facteurs de notre étude – NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE –.¹¹⁹

Nous supposons que *le maintien sous forme lexicale impliquant des changements linguistiques* augmente avec le NIVEAU SCOLAIRE et que celui *via la répétition* diminue. Ceci est confirmé par nos analyses. En effet, *le maintien d'une ancienne information sans modification via sa précédente mention* diminue de façon générale avec les années de scolarisation (160). La différence de proportion entre les CM2 et les 5^{ème} ne se révèle pas significative. Néanmoins, il apparaît que les 3^{ème} en produisent significativement moins que les deux populations les plus jeunes. La raison pour laquelle les 3^{ème} maintiennent moins une information en utilisant la répétition que les deux autres populations, est qu'ils la maintiennent en faisant des modifications (161). Les collégiens de 3^{ème} utilisent davantage cet encodage pour maintenir une information connue (161) que les 5^{ème} et les CM2 dont la différence de proportion est non significative.

(160) *Et après j(e) l'ai dit à **ma mère** / **ma mère** elle est venue. (CM2 n°08, narratif oral, E/O)*

L'enfant introduit le nouveau référent *ma mère* et le maintien dans la clause suivante sous forme lexicale sans faire de modification linguistique.

(161) *On faisait **un exercice** individuellement quand ma copine m'a demandé si j'avais utilisé la première méthode pour répondre **au problème**. (3^{ème}n°39, narratif écrit, O/E)*

La collégienne introduit une nouvelle information, *un exercice*, puis la maintient sous une forme lexicale impliquant des changements linguistiques, *au problème*.

De plus, une étude qualitative des *SNL maintenant une information en impliquant des changements linguistiques* révèle que les modifications des CM2 sont minimales par rapport à celles effectuées par des 3^{ème}. En effet, les individus les plus jeunes réalisent des modifications sur les déterminants (162). En revanche, les participants les plus âgés font des modifications exigeant plus d'efforts cognitifs telles que la substitution lexicale (163). Les 3^{ème} font preuve

¹¹⁹ Les attentes, hypothèses et problématiques sont davantage explicitées au Chapitre 3 (Section 3.2.2.5, p. 105).

de flexibilité et d'ouverture de leur répertoire linguistique alors que les CM2 sont restreints à des changements morphosyntaxiques.

(162) *Une fille s'était mise avec moi dans une chambre avec d'autres filles aussi **cette fille** s'énervait contre sans arrêt pour rien. (CM2 n°08, narratif écrit, E/O)*

(163) *Ça s'appelle **une antisèche**, ça prouve qu'on n'a pas appris nos leçons. Si le ou la professeur voit **le papier** dans la trousse, il va peut être mettre une heure de colle ou une sanction plus grave. (3^{ème} n°44, expositif écrit, E/O)*

Une seconde remarque qualitative concernant le rôle de ces maintiens lexicaux peut être ajoutée. Le maintien lexical, que ce soit avec ou sans modification, peut avoir divers rôles. Si les plus jeunes maintiennent sous forme lexicale en créant de la redondance (164), les plus âgés se servent du maintien lexical pour créer des effets de style ou encore pour désambiguïser des situations (165). Les participants les plus experts ont le souci de faciliter au mieux la compréhension du destinataire.

(164) ***Ma mère** est arrivée et il a menti à **ma mère** et **ma mère** l'a cru. (CM2 n°08, narratif écrit, E/O)*

(165) *Il y a quelques jours, j'ai acheté **un nouveau téléphone portable** sur internet. Ce n'est pas du tout pratique car il y a trop de problèmes. En fait, j'en ai rencontrés plusieurs : lors de la livraison du colis, il n'eu pas **l'objet** que j'avais commandé. (3^{ème} n°24, narratif écrit, O/E)*

Le CM2 utilise la répétition pour maintenir le référent *ma mère* crée de la redondance.

En revanche, le maintien de l'information sous forme lexicale réalisé par un individu de 3^{ème} n'a pas cet effet (165). Dans cet exemple, le référent *un nouveau téléphone portable* est maintenu par une substitution lexicale, *l'objet*, ce qui produit un certain effet de style et révèle une certaine maîtrise de la langue. Puis, de nouveau l'individu apporte des modifications linguistiques lors de la mention suivante, *le téléphone*, et ce maintien lexical a pour but de réintroduire cet élément qui aurait pu être désactivé ; il prend en considération que la distance entre les deux mentions peut impliquer un problème de compréhension si le référent est maintenu avec un pronom.

Une autre de nos attentes était que selon le contexte de production (TYPE DE TEXTE et MODALITÉ), le type de maintien était différent : les individus maintiendraient plus fréquemment une information en variant la forme linguistique dans les textes expositifs et écrits. Les analyses révèlent que seule la MODALITÉ DE PRODUCTION a un effet significatif. Si la proportion de

l'ancienne information maintenue par le biais d'une répétition est plus importante dans les textes oraux, l'ancienne information maintenue grâce à la production d'un SNL impliquant des changements linguistiques caractérise plus les productions écrites.

Ces résultats semblent compréhensifs si nous nous en référons aux caractéristiques des modalités de production. La modalité écrite, notamment par son temps de planification important, donne la possibilité aux locuteurs/scripteurs d'utiliser un type d'encodage se réalisant moins aisément en production orale. Les locuteurs/scripteurs apprennent à aller à l'encontre de certains principes de transmission de l'information : ils produisent des SNL dans le but de maintenir l'information

12.4. Conclusion du Chapitre 12

Ce chapitre montre que les participants de notre étude maintiennent une information préférentiellement sous forme pronominale que sous forme lexicale. Cette conclusion suit la tendance révélée par des travaux précédents tels que ceux de Fox (1987) ou encore Khorounjaia et Tolchinsky (2004). Ces résultats révèlent également que le maintien sous forme pronominale et lexicale varie selon trois facteurs de ce travail de recherche – NIVEAU SCOLAIRE, MODALITÉ et ORDRE DE PRODUCTION–. La proportion des pronoms maintenant une information diminue avec les années de scolarisation, et est plus importante dans les textes oraux et dans les productions réalisées par des individus appartenant à l'ordre oral/écrit. En revanche, la proportion des SNL maintenant une information augmente avec l'âge, et est plus importante dans les textes écrits et dans les productions réalisées par des participants appartenant à l'ordre écrit/oral. Le TYPE DE TEXTE n'a aucun alors que nous en attendions un.

Le maintien sous forme lexicale est non naturel dans le sens où il exige d'aller à l'encontre des principes de transmission de l'information de l'oral. Le cheminement vers la capacité de maintenir une information par le biais de substitutions est bien engagé. Cela nous amène à faire un lien avec les types de planification introduits (Bereiter et Scardamalia, 1987 ; Kellogg, 2008) : nos individus n'auraient atteint la capacité de faire appel à *la planification par connaissances révisées* (Kellogg, 2008). Les 3^{ème} sont capables de faire appel à *la planification par connaissances rapportées* et *la planification par connaissances transformées* mais les plus jeunes doivent se limiter à *la planification par connaissances rapportées*. Les résultats suggèrent que les 3^{ème} ont le souci de varier les formes et d'être

flexible. En effet, la substitution lexicale implique de plus en plus de modifications linguistiques entre la première mention et les suivantes.

Conclusion et perspectives

i. Discussion générale

i.i. Le développement du langage au-delà de 10 ans : Bilan

i.i.i. Transmission de l'information

Ce travail de thèse montre que le développement du langage perdure au-delà de cinq ans. Les enfants, avec les années de scolarisation et l'expérience avec des contextes divers, développent leurs capacités linguistiques et font preuve de plus en plus de flexibilité et d'adaptabilité rhétorique (Berman, 1998, 2006 ; Berman et Ravid, 2009 ; Ravid et Tolchinsky, 2000, 2002). Ils produisent des formes linguistiques de plus en plus variées : leur répertoire linguistique augmente et se diversifie. En effet, ils produisent plus fréquemment des SNL qui ont, par ailleurs, eux-mêmes une architecture plus complexe (soit, plus d'un nœud lexical).

Pour qu'un texte soit cohérent, rappelons que plusieurs principes et règles sont à respecter (Charolles, 1978 ; Clark et Haviland, 1974, 1977 ; Reinhart, 1980 ; Schneuwly, 1988). Un texte doit comporter de la nouvelle information et de l'ancienne (*given-new contract*, Clark et Haviland, 1974, 1977). Produire un texte cohérent exige alors une symbiose et un équilibre entre l'introduction de nouvelles informations et la continuité référentielle. Les productions de nos individus vont dans ce sens. Elles sont, en effet, de plus en plus riches : les locuteurs/scripteurs introduisent, certes, de nouvelles informations, mais ils maintiennent et

développent également davantage les référents introduits. Ils ont alors des référents à récurrence stricte (Bellert, 1970 ; Charolles, 1978 ; Reinhart, 1980) qu'ils maintiennent et sur lesquels ils font des prédications (Charolles, 1978 ; Clark et Haviland, 1974, 1977 ; Schneuwly, 1988). Certaines *méta-règles* de Charolles, le *given-new contract* de Clark et Haviland (1974, 1977) ou encore *le travail de conservation/progression* de Schneuwly (1988) sont alors respectés.

Les choix linguistiques des participants de notre étude montrent qu'ils respectent un principe essentiel de transmission de l'information, et notamment dans les productions narratives et orales. Nos résultats confirment l'idée que les référents, selon leur statut informationnel, sont préférentiellement encodés sous une forme spécifique. En effet, pour introduire, réintroduire ou afin d'éviter toute ambiguïté, les formes lexicales sont mobilisées, ce qui confirme des études antérieures (Ariel, 1990, 1996 ; Allen *et al.*, 2008 ; Khorounjaia et Tolchinsky, 2004). Selon leur statut référentiel, les référents sont également placés différemment dans la clause (entre autres, Allen *et al.*, 2008 ; Ariel, 1990 et 1996 ; Khorounjaia et Tolchinsky, 2004 ; Levelt, 1989 ; Sanford et Garrod, 1981a, 1981b). Une information nouvelle est préférentiellement en position non sujet, ce qui va dans le sens de travaux antérieurs (Blanche-Benveniste, 1990, 1995 ; Chafe, 1994 ; Jeanjean, 1980 ; Lambrecht, 1987). Un autre principe de transmission de l'information est le maintien sous forme pronominale (Ariel, 1990, 1996 ; Allen *et al.*, 2008 ; Khorounjaia et Tolchinsky, 2004).

Néanmoins, le maintien sous forme lexicale, qui apparaît comme un indice de maturité linguistique, se réalise plus fréquemment, en impliquant notamment des formes diverses. La substitution lexicale¹²⁰ dépend de la capacité de l'enfant à produire un SNL, ce qui représente un effort cognitif plus important que l'utilisation d'un pronom, pour un référent déjà connu. Ce choix linguistique exige également de l'enfant de comprendre que, selon le contexte de production, il peut être utile, pour créer un effet de style, d'aller à l'encontre d'un des principes de transmission de l'information : maintenir une information avec un pronom. Les individus de 3^{ème} de notre population, avec l'expérience de l'écrit et de différents types de textes, arrivent à aller à l'encontre de ce principe de transmission de l'information de l'oral.

¹²⁰ Quand il ne s'agit pas de répétition.

i.i.ii. Production langagière et contexte

Nous avons vu que la littéracie linguistique (Ravid et Tolchinsky, 2000 et 2002) implique un trait spécifique, *le contrôle des variations linguistiques*, une condition, *la familiarité et maîtrise de l'action d'écrire et du langage écrit* et un processus, *le métalangage*.

Le contrôle des variations linguistiques et l'encodage des messages par le biais de structures linguistiques adaptées à divers contextes apparaissent comme un des objectifs principaux des systèmes d'éducation (Viguié et Jisa, 1999¹²¹). Les locuteurs/scripteurs deviennent experts (« *linguistically literate* », Ravid et Tolchinsky, 2000:1 ; Ravid et Tolchinsky, 2002:420) en ayant la capacité d'utiliser diverses structures adaptées aux contextes et aux buts communicationnels. Ils possèdent alors un répertoire de structures important et font preuve de flexibilité quant à leur utilisation. Ils deviennent conscients que, selon le contexte, ils ont à encoder un message différemment.

Par *familiarité et maîtrise de l'action d'écrire et du langage écrit*, nous entendons la maîtrise de deux aspects concernant la modalité écrite : (a) les individus doivent appréhender la production écrite comme un type de discours spécifique ; (b) les individus doivent apprendre à considérer l'écrit comme un système notationnel spécifique.

Le processus spécifique *du métalangage* réfère au développement des connaissances plus explicites et analytiques du langage par le biais de l'expérience.

Les résultats révèlent que les locuteurs de notre étude : (a) contrôlent les variations du contexte ; et (b) appréhendent la modalité écrite comme un discours spécifique. Selon le type de texte et la modalité de production, les choix linguistiques sont différents ; le message est adapté au contexte et au but communicationnel de la situation d'énonciation, comme le prédit le modèle de Levelt (1989).

En effet, les textes expositifs se caractérisent comme comportant plus de SNL qui sont plus variés et plus de SN introduisant une nouvelle information. Les choix linguistiques des individus de notre étude révèlent que les contraintes de ce type de texte, notamment la pression informationnelle et le registre académique, sont respectées. De plus, ce type de texte

¹²¹ Les ouvrages qui suivent sont notamment cités : La maîtrise de la langue au collège, Apprendre à lire et à écrire : dix ans de recherche sur la lecture et la production des textes dans la revue française de pédagogie, Les cycles à l'école primaire : Une école pour enfant des outils pour les maîtres, La maîtrise de la langue à l'école : Une école pour enfant des outils pour les maîtres, Lire et écrire à l'école primaire, Programmes à l'école primaire : Une école pour enfant des outils pour les maîtres, Le système éducatif de la France, ou l'ouvrage de Auduc et Bayard-Pierlot (1995).

presse tant les locuteurs/scripteurs, qu'il contient plus d'utilisations non conventionnelles de pronoms. En effet, la macrostructure non linéaire et pas clairement identifié de ce type de texte est acquise tardivement. Les ressources cognitives sont allouées à la gestion de ce type de texte et donc, des utilisations non conventionnelles apparaissent. En revanche, le type de texte narratif est acquis relativement tôt ; en production narrative, les utilisations non conventionnelles de pronoms ont une proportion moins importante. Les textes narratifs contiennent davantage de pronoms, de SNL *simples* et d'ancienne information. En effet, ce type de texte n'exige pas un registre aussi académique que le type de texte expositif, ainsi les formes plus *légères* sont préférées. De plus, la narration n'implique pas une pression informationnelle importante, une fois un personnage introduit, il est maintenu par un pronom. Les caractéristiques de ces types de textes sont résumées dans le Tableau 11.

[Tableau 11 : Caractéristiques des textes expositifs et narratifs]

Tableau 11 : Caractéristiques des textes expositifs et narratifs

Textes expositifs	Textes narratifs
Plus de SNL (et plus de SNL en position sujet)	Plus de pronoms
Plus de SNL complexes/lourds	Plus de SNL simples/légers
Plus de nouvelles informations	Plus d'anciennes informations
Plus d'utilisations non conventionnelles de pronoms	Moins d'utilisations non conventionnelles de pronoms

Les textes écrits se caractérisent comme contenant davantage de SNL (complexes de surcroît) que de pronoms qui sont plus utilisés en production orale. L'écrit comporte également plus de nouvelles informations que l'oral qui contient davantage d'anciennes informations. L'ancienne information est, par ailleurs, davantage encodée sous forme lexicale dans les textes écrits et sous forme pronomiale en production orale. Ce maintien lexical se réalise en variant davantage les formes en production écrite qu'en production orale. Les caractéristiques des modalités du langage sont regroupées ci-dessous (Tableau 12).

[Tableau 12 : Caractéristiques des textes écrits et oraux]

Tableau 12 : Caractéristiques des textes écrits et oraux

Textes écrits	Textes oraux
----------------------	---------------------

Plus de SNL (et plus de SNL en position sujet)	Plus de pronoms
Plus de SNL complexes	Plus de SNL simples
Plus de nouvelles informations	Plus d'anciennes informations
Plus d'anciennes informations sous forme lexicale	Plus d'anciennes informations sous forme pronominale

Les résultats montrent que les enfants, même les plus jeunes, savent distinguer l'écrit de l'oral et s'adapter à ces modalités du langage.¹²² Les individus de notre étude savent profiter du rythme de production lent de la production écrite : des formes dites plus coûteuses sont mobilisées, sans doute tirent-ils avantage d'un temps de planification plus long. L'utilisation de telles formes révèle également que les enfants répondent aux contraintes rédactionnelles de l'écrit.

i.i.iii. Attentes et résultats

Le Tableau 13 ci-dessous résume pour chaque variable principale sa variation quant aux facteurs, NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ et ORDRE.

[Tableau 13 : Récapitulation des résultats]

Tableau 13 : Récapitulation des résultats

	NIVEAU SCOLAIRE	TYPE DE TEXTE	MODALITÉ	ORDRE
SNL (Chapitre 8)	Augmente	Plus en expositif	Plus à l'écrit	Plus écrit/oral
Pronoms (Chapitre 8)	Diminue	Plus en narratif	Plus à l'oral	Plus oral/écrit
SNL 1 nœud (Chapitre 9)	Diminue	Plus en narratif	Plus à l'oral	Plus oral/écrit
SNL +1 nœuds (Chapitre 9)	Augmente	Plus en expositif	Plus à l'écrit	Plus écrit/oral
Ancienne information (Chapitre 10)	Cf. Discussion p. 231	Plus en narratif	Plus à l'oral	NS

¹²² Les enfants de cette étude gèrent de mieux en mieux le système orthographique du français en faisant notamment moins de fautes d'orthographe et en se révisant de plus en plus (Maggio, 2007 ; Maggio, Jisa, Lété et Fayol, 2008 ; Maggio, Lété et Fayol, 2008 ; Cautain, Chenu et Jisa, 2008).

Nouvelle information (Chapitre 10)	Cf. Discussion p. 231	Plus en expositif	Plus à l'écrit	NS
Utilisations des pronoms non conventionnelles (Chapitre 11)	Diminue	Plus en expositif	NS	NS
Ancienne information sous forme lexicale (Chapitre 12)	Augmente	NS	Plus à l'écrit	Plus écrit/oral
Ancienne information sous forme pronominale (Chapitre 12)	Diminue	NS	Plus à l'oral	Plus oral/écrit
Substitution lexicale : répétition (Chapitre 12)	Diminue	NS	Plus à l'oral	NS
Substitution lexicale : changements linguistiques (Chapitre 12)	Augmente	NS	Plus à l'écrit	NS

Ces résultats répondent, d'une manière générale, positivement à nos attentes de départ, regroupées dans le Tableau 4, à la Section 3.3 (p. 108). Si nous n'avions pas posé d'attentes strictes quant à *la proportion de l'ancienne information et celle de la nouvelle information*, les analyses font émerger des tendances. Les résultats montrent que *la proportion de l'ancienne information* augmente avec le NIVEAU SCOLAIRE et que celle de *la nouvelle information* diminue. Néanmoins, rappelons que ce résultat a été source de discussion (cf. p. 235). Il en a été déduit que les productions des individus les plus âgés (3^{ème}) sont plus longs grâce à la capacité d'introduire et aussi de maintenir. Leurs textes sont plus riches en nouvelles informations mais ces nouvelles informations sont également plus maintenues et développées. Avec l'expérience, les individus enrichissent leurs textes aussi bien en multipliant les nouveaux référents qu'en les maintenant davantage, dans le but de discuter dessus.

De plus, nous notons que certains facteurs dont nous attendions des effets n'affectent pas certaines variables. Seuls les facteurs NIVEAU SCOLAIRE et MODALITÉ ont toujours un impact. Avec les années de scolarisation et donc avec l'expérience à l'écrit et à divers types de textes, la production langagière en situation monologique des enfants de notre étude évolue. Chaque

variable principale observée dans ce travail de recherche (Tableau 13) varie selon le niveau de scolarisation des individus. La modalité de production est un des aspects du contexte de production influençant toujours les choix linguistiques des individus.

L'ORDRE et le TYPE DE TEXTE sont les deux facteurs qui interviennent plus sporadiquement. L'ORDRE n'affecte pas *la proportion de l'ancienne et de la nouvelle information* ou encore l'architecture des SNL maintenant une information (répétition *versus* changements linguistiques). Le TYPE DE TEXTE ne fait varier aucune des variables concernant la forme du maintien d'un référent (à savoir : *la proportion de l'ancienne information sous forme lexicale et sous forme pronominale*¹²³ ainsi que *la substitution avec ou sans changement linguistiques*).

Les résultats de cette recherche laissent alors penser que nos individus évoluent quant à la planification de leur message : ils mobilisent des formes linguistiques différentes avec l'expérience et adaptent ces formes au contexte. Nous avons vu que, pour organiser les contenus activés, les locuteurs font appel à deux types de planification : *la planification par connaissances rapportées* et *la planification par connaissances transformées* (entre autres, Bereiter et Scardamalia, 1987 et 1998 ; Bereiter, Burtis et Scardamalia, 1988).¹²⁴ À partir de l'analyse des formes, à considérer comme des traces de planification, nous pensons que, selon le NIVEAU SCOLAIRE et le CONTEXTE DE PRODUCTION, le type de planification sollicité est différent. Les analyses suggèrent que les plus jeunes de nos locuteurs/scripteurs ont plus tendance à faire appel à *la planification par connaissances rapportées* qu'à *la planification par connaissances transformées*, type de planification davantage sollicité avec les années de scolarisation. Les enfants mobilisent de plus en plus *la planification par connaissances transformées*, ils font preuve de flexibilité et d'adaptabilité.¹²⁵

i.ii. La Preferred Argument Structure face au développement et aux contextes de production

La PAS suppose quatre contraintes reflétant la transmission de l'information à l'oral. La première contrainte A entend qu'une clause est constituée au maximum d'un élément lexical et que selon la contrainte B, cet élément lexical n'apparaît pas en position sujet d'un verbe

¹²³ Excepté pour les analyses des proportions de SN pronominaux ancienne information selon les positions syntaxiques.

¹²⁴ Notions présentées dans la Section 2.6 du Chapitre 2 (p. 79).

¹²⁵ De plus, avec les années de scolarisation, ils font de plus en plus de révisions dans leur textes ; néanmoins, ces révisions restent ciblées aux erreurs orthographiques (Cautain *et al.*, 2008).

transitif. Selon la contrainte C, une clause est constituée d'au maximum un seul nouvel élément. Enfin, selon la contrainte D, l'ancienne information se trouve préférentiellement en position sujet (et notamment de verbe transitif).

Comme nous l'avons mentionné, la PAS possède divers niveaux : deux des contraintes concernent la forme (contraintes A et B) et deux autres la pragmatique (contraintes C et D). Une contrainte de chaque est également soumise à un aspect syntaxique (contraintes B et D).¹²⁶

Les résultats révèlent que le NIVEAU SCOLAIRE,¹²⁷ le TYPE DE TEXTE, la MODALITÉ et l'ORDRE ont un effet significatif sur *le nombre moyen d'éléments lexicaux par clause (A)* ou encore *la proportion des sujets lexicaux (B)*. Ceci révèle que les enfants, avec les années de scolarisation et l'expérience de divers types de textes et contextes, vont à l'encontre des deux contraintes syntaxiques de la PAS. Ainsi, les enfants adaptent, de plus en plus, leurs choix linguistiques selon le contexte de production : en production expositive et écrite, contexte les moins *naturels*, ils produisent davantage de SNL par clause et de SNL en position sujet.

En revanche, le niveau discursif de la PAS (contraintes C et D) est moins affecté par nos facteurs (NIVEAU SCOLAIRE, TYPE DE TEXTE, MODALITÉ, ORDRE). Ainsi, *le nombre moyen de nouvelle information (C)* et *la proportion de SN maintenant une information en position sujet (D)* varie qu'en fonction de la MODALITÉ DE PRODUCTION.

Le Tableau 14 résume les résultats en lien avec les quatre contraintes de la PAS.

[Tableau 14 : Récapitulation des résultats concernant les contraintes de la Preferred Argument Structure]

Tableau 14 : Récapitulation des résultats concernant les contraintes de la Preferred Argument Structure

	NIVEAU SCOLAIRE	TYPE TEXTE	DE	MODALITÉ	ORDRE
Respect de la Preferred Argument Structure	Diminue	Plus narratif	en	Plus à l'oral	Plus oral/écrit
A. Un seul élément lexical par clause Observation du <i>nombre moyen d'items lexicaux par clause</i> (Section 8.1.2, p. 180) Davantage vrai dans les textes narratifs, oraux	Augmente	Plus expositif	en	Plus à l'écrit	Plus écrit/oral

¹²⁶ Se reporter à la Section 1.2.4, Erreur : source de la référence non trouvée (p. 38).

¹²⁷ Excepté pour *la proportion des sujets lexicaux*.

et produits dans l'ordre oral/écrit

B. Position sujet non lexicale

Observation de la <i>proportion des SNL en position sujet</i> ¹²⁸ (Section 8.2, p. 189)	NS	Plus expositif	en Plus à l'écrit	Plus écrit/oral
Davantage vrai dans les textes narratifs, oraux et produits dans l'ordre oral/écrit				

C. Un seul nouvel élément par clause

Nombre moyen de nouvelle information par clause (Section 10.1, p. 231)	NS	NS	Plus à l'écrit	NS
Davantage vrai dans les textes oraux				

D. Ancienne information en position sujet

Proportion de l'ancienne information en position sujet ¹²⁹ (Section 10.2, p. 240)	NS	NS	Plus à l'oral	NS
Davantage vrai pour les textes oraux				

Les enfants de notre étude respectent les contraintes de la PAS reflétant les principes de transmission de l'information de la langue orale, et ce , notamment dans les textes narratifs et oraux. En plus des contraintes de la PAS, rappelons qu'ils maintiennent, en général, une information préférentiellement sous forme pronominale (et ceci en position sujet) et la nouvelle information est encodée sous forme lexicale (en position non sujet). Trois conclusions principales peuvent alors être faite.

Premièrement, les quatre contraintes de la PAS ne sont pas logées à la même enseigne. Les contraintes d'ordre pragmatique varient beaucoup moins que les syntaxiques. Nous pouvons alors penser que les contraintes pragmatiques, dont la syntaxe dépend (Du Bois, 1987, 2003), sont les plus naturelles et difficilement amendables. L'acquisition et le maniement de *la langue de scolarisation* implique d'aller à l'encontre des contraintes syntaxiques : les productions des enfants contiennent de plus en plus de SNL et de SNL en position sujet.

Deuxièmement, le respect de ces contraintes et des principes de la transmission de l'information dépend du contexte de production. Ainsi, en production orale et narrative, ce patron est davantage retrouvé qu'en production écrite et expositive.

Troisièmement, ceci montre l'impact de l'expérience de l'écrit et des divers types de textes : en production écrite, les enfants vont à l'encontre de ces contraintes et ces principes. Ainsi, à

¹²⁸ Parmi les positions sujet, la position sujet transitif est la moins lexicale (Section 8.1.2, p. 192).

¹²⁹ Parmi les positions sujet, la position sujet transitif est celle portant le plus l'ancienne information (Section 10.2, p. 243).

l'écrit, de nouveaux principes se mettent en place et certains principes de l'oral en sont ébranlés.

i.iii. Formes, positions et statut informationnel

Selon la position des unités observées, les facteurs impliquant des résultats significatifs ne sont pas les mêmes : les positions syntaxiques ne sont pas traitées de la même manière.

Dans notre Partie I (Théorie), nous avons vu que, selon divers travaux (Bock et Warren, 1985 ; Bock et Levelt, 1994 ; Keenan et Comrie, 1977 ; Leuwers, 2002 ; Levelt, 1999), l'attribution des rôles syntaxiques se ferait selon le caractère plus ou moins activé d'un élément : plus un référent serait activé, plus cet élément se verrait attribuer une fonction syntaxique dite de haut niveau (position sujet). De plus, certaines positions syntaxiques seraient prédisposées à être remplies par certaines formes. Ainsi, le lexique serait préférentiellement en position non sujet et les pronoms se verraient attribuer la position sujet (*la contrainte du sujet léger*, Chafe, 1994). Ces grandes tendances sont confirmées par nos résultats (Tableau 15).

[Tableau 15 : Formes et statuts des référents remplissant les positions sujet et non-sujet]

Tableau 15 : Formes et statuts des référents remplissant les positions sujet et non-sujet

Position SUJET remplie préférentiellement de :	Positions NON-SUJET (objet et autre) remplie préférentiellement de :
Pronoms	SNL
Si présence de SNL = SNL simples/légers	SNL simples et complexes/lourds
Anciens référents	Nouveaux référents
Ancienne information sous forme pronominale	Si présence d'ancienne information, aussi bien sous forme pronominale que lexicale

Nos analyses montrent alors que formes, syntaxe et statut référentiel sont liés quant les choix linguistiques sont effectués. Nous ne pouvons les analyser séparément.

Les participants de notre étude préfèrent remplir la position sujet avec une forme pronominale qu'avec une forme lexicale, plutôt destinée à remplir les positions non-sujet. Les individus de notre étude respectent les principes selon lesquels : (a) l'ancienne information est

préférentiellement en position sujet qu'en positions non-sujet et (b) la nouvelle information est préférentiellement en position non-sujet.

De plus, en nous penchant un peu plus sur la forme de l'ancienne information, nous notons que la position sujet est plus remplie par de l'ancienne information sous forme pronominale. En revanche, les positions non-sujet privilégient l'ancienne information sous forme lexicale. Enfin, si la position sujet est remplie par un SNL, ce qui est plutôt rare, ce SNL est simple. En revanche, les SNL sont plus lourds en position non-sujet.

Nous pouvons alors conclure que l'étude du syntagme nominal dans une perspective fonctionnaliste intégrant syntaxe et pragmatique semble essentielle à la compréhension du développement de la syntaxe discursive de la langue (Ravid et Berman, sous presse).

ii. Perspectives

Des améliorations peuvent être apportées à ce travail de thèse.

Premièrement, nous pouvons travailler davantage sur la notion de nouvelle information afin de subdiviser cette catégorie en nouvelle information et information accessible. Il faudrait mettre au point des critères extrêmement stricts pour distinguer ces deux types informationnels et vérifier qu'ils soient adaptables à différents types de textes. L'ajout d'un nouveau niveau dans le continuum nouvelle/ancienne information pourrait alors nuancer les résultats relatifs à la proportion de la nouvelle information tels que celui montrant que la proportion de la nouvelle information diminue avec le NIVEAU SCOLAIRE.

Deuxièmement, il pourrait également être intéressant de compléter cette étude avec une étude suivant le même protocole mais en prenant en considération les capacités de la MDT des locuteurs/scripteurs. En effet, il apparaît clairement que « *la capacité de la MDT est limitée et qu'elle augmente significativement avec l'âge* » et que « *la capacité de la MDT est supposée exercer de fortes contraintes quant à l'efficacité avec laquelle tout individu est en mesure de mener à bien une activité cognitive complexe* » (Monnier, 2008:279). Sur la base de ces présupposés, nous pouvons imaginer que les différences de capacité cognitive entre les locuteurs/scripteurs peuvent avoir un impact important sur le produit fini. Seigneuric et Megherbi (2008) ont déjà montré l'impact de la capacité de la MDT en compréhension lors d'une étude sur le traitement des pronoms (contexte phrastique). Ainsi, les enfants à capacité

faible de MDT sont moins performants que les enfants à capacité élevée de MDT. Les chercheurs en déduisent alors que « *le traitement des pronoms est contraint par la capacité de MDT des enfants* » (Seigneuric et Megherbi, 2008:281). Pour distinguer les enfants à capacité élevée de MDT de ceux à capacité faible, plusieurs outils sont utilisés. Seigneuric et Megherbi (2008) comme Seigneuric, De Guibert, Megherbi, Potier et Picard (2008), Seigneuric et Ehrlich (2005) et Seigneuric, Ehrlich, Oakill et Yuill (2000), proposent de se baser sur une adaptation du *Listening Span Text* de Siegel et Ryan (1989). Cet outil de mesure consiste à placer l'enfant « *dans une situation de double tâche : il doit trouver le dernier mot de phrases incomplètes lues par l'expérimentateur et maintenir en mémoire ces mots afin de les rappeler à la fin d'une série de phrases.* » (Seigneuric et Megherbi, 2008:288). Olive et Piolat (2005), suite à la récapitulation de quelques travaux concernant la production textuelle et *la mémoire de travail*, en concluent également que les processus rédactionnels requièrent des ressources de la MDT. « *Aussi, la quantité de ressources disponibles, entendue comme la capacité de la MDT, devrait influencer la production de textes* » (Olive et Piolat, 2005:384). Déjà Benton, Kraft, Glover et Plake dans les années 80 (1984) observe que « *de bons rédacteurs et des mauvais rédacteurs adultes pouvaient être différenciés en fonction de leurs scores à des épreuves d'empan de MDT* » (Olive et Piolat, 2005:385). L'impact de la capacité de la MDT sur la composition d'un texte, sur l'automatisation des processus de rédaction mais également sur les processus de révision a déjà inspiré de nombreux travaux (entre autres, Olive, 2004 ; Piolat et Fruttero, 1999 ; Piolat, Roussey, Olive et Amada, 2004).

Ce travail de recherche peut être également compléter par divers type de travaux.

Ainsi, dans un premier temps, la CSP pourrait devenir une variable indépendante. En effet, dans ce travail de recherche, pour chaque résultat, lorsque nous disposions d'analyses similaires, nous avons comparé nos individus de CSP défavorisés avec des francophones natifs de CSP plus favorisée examinés dans d'autres études. Nous pouvons alors dire que, dans l'état actuel de nos connaissances, nos participants ne se révèlent pas comme ayant un comportement extravagant : ils suivent, en règle générale, les mêmes tendances que leurs pairs de classe moyenne supérieure. Néanmoins, nous en saurions davantage si la CSP devient un facteur.

Dans un second temps, nous pouvons également introduire une population adulte défavorisée, ce qui pourrait constituer un nouveau projet. L'éducation nationale semble jouer son rôle de précepteur de la maîtrise de la langue orale et écrite, mais nous supposons qu'une fois

déscolarisés, les individus défavorisés n'ont plus autant accès aux préceptes de la littéracie linguistique et fossilisent leur acquisition. Des différences entre les deux CSP pourraient apparaître entre des populations adultes favorisée et défavorisée. Nous pourrions alors réfléchir à une solution au problème d'approche des adultes défavorisés. En effet, lors que nous avons essayé de compléter notre population avec un groupe contrôle d'adultes défavorisés, nous nous sommes heurtée à divers problèmes : (a) les adultes ne semblaient guère intéressés par ce projet ; (b) ils se sentaient mal à l'aise et avaient peur du jugement ; (c) il nous était difficile de les mettre à l'aise tout en leur demandant leur niveau scolaire, de vie, leur emploi actuel, leur pays de naissance et s'ils avaient éventuellement une pathologie langagière. Il faudrait alors mettre au point une stratégie bien spécifique afin d'approcher ce type de population, si ce n'est de collaborer avec des sociologues et des professionnels de ce terrain.

Dans un troisième temps, il serait intéressant de compléter cette analyse des produits finis (analyse *off line*) par une étude des processus en temps réel ou par des expérimentations visant à tester le niveau d'explicitation des règles discursives de la langue. Aussi, par exemple, nous pouvons nous demander si la pause avant un SN introduisant une information est plus longue que celle précédant un SN maintenant une information. De nouveaux outils ont fait leur apparition et ont permis l'utilisation de techniques particulièrement sophistiquées telles que le recueil des mouvements oculaires lors de la production ou le recueil d'indicateurs comportementaux comme les pauses ou les débits à l'aide de tablettes graphiques à l'écrit (Coirier *et al.*, 1996). De plus, les analyses de la localisation et de la durée des pauses, combinées avec des descriptions textuelles, ne peuvent être qu'un avantage dans la tentative de la compréhension des processus rédactionnels (Schilperoord, 1996).

Enfin, dans un quatrième temps, cette étude porte sur des enfants et adolescents sains, ne présentant aucune pathologie langagière, issus de milieux défavorisés. Notre travail de recherche peut servir de tremplin pour engager des études ayant comme objectif par exemple de comparer les spécificités et le développement linguistique de nos individus sains à une population présentant des troubles langagiers telle que les dyslexiques, les enfants atteints du syndrome de Williams ou encore ceux ayant des lésions cérébrales. Depuis une quinzaine d'année, les études sur le développement langagier de cas atypiques se multiplient. Nous avons commencé à nous documenter sur ce sujet avec notamment les travaux de Bates (2004), Bernicot, Lacroix et Reilly, (2003), Nichols, Jones, Roman, Wulfeck, Delis, Reilly and

Bellugi (2004), Reilly et Bernicot (2003), Reilly, Losh, Bellugi et Wulfeck (2004), Reilly, Weckerly et Bates (2003), Reilly et Wulfeck (2004), Thal, Reilly, Seibert, Jeffries et Fenson (2004), Wulfeck, Bates, Krupa-Kwiatkowski et Saltzman (2004). L'objectif de ces études semble être double : « *fondamental, pour accroître nos connaissances sur le développement et le fonctionnement du langage, et appliqué, pour améliorer la prise en charge des enfants atteints de troubles de l'acquisition du langage* » (Reilly et Bernicot, 2003:205). Ainsi, « *étudier les cas où l'acquisition du langage « se passe mal » peut nous aider à déterminer les facteurs impliqués quand « tout va bien »* » (Reilly et Bernicot, 2003:206). Ainsi, comprenons-nous que des études portant sur la morphosyntaxe (Reilly *et al.*, 2003) ou encore sur la syntaxe ont été réalisées avec, pour certaines, la particularité de prendre en considération aussi bien des aspects structuraux et pragmatiques (Bernicot *et al.*, 2003). Néanmoins, nous-pouvons nous demander ce qu'il en est des aspects syntaxiques tels que la production de SNL, des aspects informationnels (avec le continuum nouvelle information/ancienne information) ou encore des aspects discursifs avec par exemple la PAS. Il serait alors intéressant de voir dans quelles mesures les contraintes de la PAS sont respectées ou non par ce type de population. Si les contraintes pragmatiques ne sont guère amendables par les individus sains, les contraintes syntaxiques sont largement dépassées dans certains contextes. Qu'en est-il des populations pathologiques ?

Bibliographie

- Adam, J.M. (1985). *Le texte narratif, précis d'analyse textuelle*, Poitiers: Fernand Nathan.
- Adam, J.M. (1992). *Les textes : types et prototypes. Récit, description, argumentation, explication et dialogue*, Nathan université, Série Linguistique, Paris: Nathan.
- Adam, J.M. et Revaz, F. (1996). *L'analyse des récits*, Paris: Éditions du Seuil.
- Aisenman, R. et Assayag, N. (1999). Fact and fiction in evaluating narrative evaluation, *Developing Literacy Across Genres, Modalities, And Languages*, 1, 38-57.
- Aisenman, R. et Berman, R. (2000). Rethinking lexical analysis, *Developing Literacy Across Genres, Modalities, And Languages*, 3, 187-196.
- Aksu-Koç, A. et Küntay, A. (2001). Reformulating causal relations while retelling narratives: evidence from Turkish, *Colloque Acquisition et construction du sens dans une perspective interlangue* Paris.
- Alamargot, D. et Chanquoy, L. (2002). Les modèles de rédaction de textes, In Fayol, M. (eds.), *Production du langage*, 45-65, Paris: Lavoisier.

- Alamargot, D., Chanquoy, L. et Chuy, M. (2005). L'élaboration du contenu du texte : de la mémoire à long terme à l'environnement de la tâche, *Psychologie Française*, 50, 287-304.
- Alamargot, D., Lambert, E. et Chanquoy, L. (2005). La production écrite et ses relations avec la mémoire, *Approche Neuropsychologique des Acquisitions de l'Enfant*, 17, 41-46.
- Allal, L., Chanquoy, L. et Lardy, P. (2004). *Revision: Cognitive and Instructional Processes*, Dordrecht: Kluwer Academic Publishers.
- Allen, S. (1997). *A discourse-pragmatic explanation for the subject-object asymmetry in early null arguments: The Principle of Informativeness revisited*, Paper presented at the GALA '97 Conference on Language Acquisition, University of Edinburgh.
- Allen, S. (2000). A discourse-pragmatic explanation for argument representation in child Inuktitut, *Linguistics*, 38, 483-521.
- Allen, S. (2007). Interacting pragmatic influences on children's argument realization, In Bowerman, M., Brown, P. (eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability*, 191-210, Mahwah, NJ: Erlbaum
- Allen, S. et Schröder, H. (2003). Preferred Argument Structure in early Inuktitut spontaneous speech data, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 11-60, Amsterdam: John Benjamins.
- Allen, S., Skarabela, B. et Hughes, M. (2008). Using corpora to examine discourse effects in syntax, *Trends in language acquisition research, Corpora in language acquisition research*, 6.
- Anderson, J. (1983). *The architecture of cognition*, Cambridge, MA: Harvard University Press.
- Anderson, J. (1993). *Rules of mind*, Hillsdale, N. J.: Lawrence Erlbaum Associates.
- Anglin, J. (1993). Vocabulary development: A morphological analysis, *Monographs of the society for Research in Child Development*, 58, 1-165.
- Aparici, M., Tolchinsky, L. et Rosado, E. (2000). On defining longer Units in narrative and expository Spanish texts, *Developing Literacy Across Genres, Modalities, And Languages*, 3, 95-122.
- Argerich, N. et Tolchinsky, L. (2000). On a definition of lexical items in written and spoken texts, *Developing Literacy Across Genres, Modalities and Languages*, 3, 197-204.
- Ariel, M. (1990). *Accessing Noun Phrases Antecedents*, London: Routledge.
- Ariel, M. (1996). Referring expressions and the +/- coreference distinction, In Fretheim, T., Gundel, J. (eds.), *Reference and referent accessibility*, 13-37, Amsterdam/Philadelphia : John Benjamins.
- Arnold, J. (2003). Multiple constraint on reference form: Null, pronominal, and full reference in Mapudungun, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 225-272, Amsterdam: John Benjamins.
- Arrivé, M., Gadet, F. et Galmiche, M. (1986). *La grammaire d'aujourd'hui*, Flammarion.
- Ashby, W. et Bentivoglio, P. (1993). Preferred argument structure in spoken French and Spanish, *Language Variation and Change*, 5, 61-76.

- Ashby, W. et Bentivoglio, P. (2003). A comparative diachronic analysis of French and Spanish , In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred argument structure: Grammar as architecture for function*, 61-81, Amsterdam: John Benjamins.
- Atkinson, R. et Shiffrin, R. (1968). Human memory: A proposed system and its control processes, In Spence, K. (eds.), *The psychology of learning and motivation: Advances in research and theory*, Vol. 2, 89-185, New York: Academic Press.
- Atkinson, R. et Shiffrin, R. (1971). The control of short-term memory, *Scientific American*, 225, 82-90.
- Auduc, J.L. et Bayard-Pierlot, J. (1995). *Le système éducatif français. Les guides de l'école, 4ème édition*, Académie de Créteil: CRPD (Centre Régional de Documentation Pédagogique).
- Auer, P. (2009). On-line syntax: Thoughts on the temporality of spoken language, *Language Sciences*, 31, 1-13.
- Baddeley, A. et Hitch, G. (1974). Working memory, In Bower, G. (ed.), *The psychology of learning and motivation: Advances in research and theory*, 47-89, New York: Academic Press.
- Baddeley, A. (1986). *Working memory*, Oxford Oxfordshire, New York: Clarendon Press ; Oxford University Press.
- Baddeley, A. (1990/1992). *La mémoire humaine, Théorie et pratique*, Grenoble: Presse Universitaire de Grenoble.
- Baddeley, A. (1992). Working Memory, *Science*, 255, 556-559.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory?, *Trends in Cognitive Sciences*, 4, 417-423.
- Baddeley, A. (2001). Is working memory still working?, *American Psychologist*, 56, 849-864.
- Bamberg, M. (1986). A functional approach to the acquisition of anaphoric relationships, *Linguistics*, 24, 227-284.
- Bamberg, M. (1987). *The acquisition of narratives: learning to use language*, Berlin: Mouton de Gruyter.
- Bartlett, E. (1984). Anaphoric reference in written narratives of good and poor elementary school writers, *Journal of Verbal Learning and Verbal Behaviour*, 23, 540-552.
- Bassano, D., Maillolchon, I., Klampfer, S. et Dressler, W. (2001). L'acquisition de la morphologie verbale à travers les langues, *Enfance*, 53, 81-99.
- Bates, E. (2004). Explaining and interpreting deficits in language development across clinical groups: Where do we go from here? , *Brain and Language*, 88, 248-253.
- Bates, E., McNew, S., MacWhinney, B., Devescovi, A. et Smith, S. (1982). Functional constraints on sentence processing: a cross-linguistic study, *Cognition*, 11, 245-299.
- Bates, E. et McWhinney, B. (1989). *The crosslinguistic study of sentence processing*, Cambridge: Cambridge University Press.
- Begnet, Z. (2007). *Étude des pauses en production écrite*, Mémoire de M2, Université de Lyon.
- Bellert, I. (1971). On a condition of the coherence of texts, *Semiotica*, 4, 253-288.
- Benoît, J. et Fayol, M. (1989). Le développement de la catégorisation des textes, *Pratiques*, 62, 71-85.
- Bentivoglio, P. (1996). Acquisition of Preferred Argument Structure in Venezuelan Spanish., *seventh international congress for the study of child language* Istanbul, Turquie.
- Benton, S., Kraft, R., Glover, J. et Plake, B. (1984). Cognitive Capacity Differences Among Writers, *Journal of Educational Psychology*, 76, 820-834.

- Bereiter, C., Burtis, P. et Scardamalia, M. (1988). Cognitive Operations in Constructing Main Points in Written Composition, *Journal of Memory and Language*, 27, 261-278.
- Bereiter, C. et Scardamalia, M. (1987). *The psychology of written composition*, Hillsdale: Lawrence Erlbaum.
- Bereiter, M. et Scardamalia, C. (1998). L'expertise en lecture-rédaction, In Piolat, A., Pélissier, A. (eds.), *La rédaction de textes, Approche cognitive*, 13-50, Lausanne: Delachaux et Niestlé.
- Berman, R. (1995). Narrative competence and storytelling performance: how children tell stories in different contexts, *Journal of Narrative and Life History*, 5, 285-313.
- Berman, R. (1997). Preliterate knowledge of language, In Portecovo, C. (eds.), *Writing development : an interdisciplinary view*, 61-76, Amsterdam: John Benjamins.
- Berman, R. (1998). Typological perspectives on connectivity, In Penner, Z., Dittmar, N. (eds.), *Issues in the theory of language acquisition*, 203-224, Bern: Peter Lang.
- Berman, R. (1999a). Form/function relations in narrative, *Developing Literacy Across Genres, Modalities, And Languages*, 1, 80-92.
- Berman, R. (1999b). Genre and modality in developing discourse abilities, *Developing Literacy Across Genres, Modalities and Languages*, 1, 93-113.
- Berman, R. (2000). Thematic perspectives on how children talk about interpersonal conflict, *Developing Literacy Across Genres, Modalities, And Languages*, 3, 65-76.
- Berman, R. (Eds.) (2001). *Setting the narrative scene: how children begin to tell a story*, Vol. 10. Mahwah, NJ : Erlbaum.
- Berman, R. (2004). Between emergence and mastery: The long developmental route of language acquisition, In Berman, R. (eds.), *Language Development across Childhood and Adolescence*, Vol. 3, 9-34, Amsterdam/Philadelphia : John Benjamins.
- Berman, R. (2005a). Introduction: Developing discourse stance in different text types and languages, *Journal of pragmatics*, 37, 105-124.
- Berman, R. (2005b). Preface, *Journal of Pragmatics*, 37, 103-104.
- Berman, R. (2006). Developing linguistic knowledge and language use across adolescence, In Hoff, E., Shatz, M. (eds.), *Handbook of Language Development*, 92-111, Blackwell.
- Berman, R. (2008). The psycholinguistics of text construction, *Journal of Child Language*, 35, 1-37.
- Berman, R. et Katzenberger, I. (2004). Form and function in introducing narrative and expository texts: A developmental perspective, *Discourse Processes*, 38, 57-94.
- Berman, R. et Nir-Sagiv, B. (2004). Linguistic indicators of inter-genre differentiation in later language development, *Journal of Child Language*, 31, 339-380.
- Berman, R. et Nir-Sagiv, B. (2007). Comparing narrative and expository text construction across adolescence: A developmental paradox, *Discourse Processes*, 43, 79-120.
- Berman, R., Ragnarsdottir, H. et Strömquist, S. (2002). Discourse stance, *Written Languages and Literacy*, 5, 255-290.

- Berman, R. et Ravid, D. (2009). Becoming a literate language user oral and written text construction across adolescence, In Olson, D., Torrance, N. (eds.), *The Cambridge Handbook of Literacy*, 92-111, Cambridge University Press.
- Berman, R. et Slobin, D. (1994a). Narrative structure, In Berman, R., Slobin, D. (eds.), *Relating events in narrative: A crosslinguistics developmental study*, 39-84, Hillsdall, NJ: Lawrence Erlbaum Associates.
- Berman, R. et Slobin, D. (1994b). *Relating events in narrative: a crosslinguistic developmental study*, Hillsdale: Erlbaum.
- Berman, R. et Strömquist, S. (1999). Typological perspectives on developing literacy, *Developing Literacy Across Genres, Modalities, And Languages*, 1, 21-22.
- Berman, R. et Verhoeven, L. (2002). Cross linguistic perspectives on the development of text-production abilities in speech and writing, *Written Language and Literacy*, 5, Vol.1.
- Berman, R. et Verhoeven, L. (2002). Cross linguistic perspectives on the development of text-production abilities in speech and writing, *Written Language and Literacy*, 5, Vol.2.
- Bernicot, J., Lacroix, A. et Reilly, J. (2003). La narration chez les enfants atteints du syndrome de Williams : aspects structuraux et pragmatiques, *Enfance*, 55, 265-281.
- Berninger, V., Fuller, F. et Whitaker, D. (1996). A process model of writing development across the life span, *Educational Psychology Review*, 8, 193-218.
- Berninger, V., Cartwright, A., Yates, C., Swanson, L. et Abbott, R. (1994). Developmental skills related to writing and reading acquisition in the intermediate grades, *Reading and Writing : A, Interdisciplinary Journal*, 6, 161-196.
- Berninger, V. et Swanson, H. (1994). Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writing, In Carlson, J., Butttterfly, E. (eds.), *Advances in Cognition and Educational Practice, Children's Writing: Toward a Process Theory of the Development of Skilled Writing*, Vol. 2, 57-81, Greenwich: J.A.I. Press.
- Berninger, V., Vaughan, K., Abbott, R., Abbott, S., Rogan, L., Brooks, A., Reed, E. et Graham, S. (1997). Treatment of hanwriting problems in beginning writers: Transfer from handwriting to composition, *Journal of*, 89, 652-666.
- Berninger, V., Vaughan, K., Abbott, R., Begay, C., Coleman, K., Curtin, G., Hawkins, J. et Graham, S. (2002). Teaching spelling and composition alone or together : Implications for the simple view of writing, *Journal of Educational Psychology* , 94, 291-304.
- Berninger, V., Vaughan, K., Abbott, R., Brooks, A., Abbott, S., Reed, L., Rogan, E. et Graham, S. (1998). Early intervention for spelling problems: Teaching spelling units of varying size within a multiple connection framework, *Journal of Educational Psychology*, 90, 587-605.
- Berninger, V., Yates, C., Cartwright, A., Rutberg, J., Remy, E. et Abott, R. (1992). Lower level developmental skill in beginning writing, *Reading and writing, An interdisciplinary Journal*, 4, 257-280.
- Bernstein, B. (1971). *Class, codes and control, Theoretical studies towards a sociology of language*, Londres: Routledge et Kegan Paul.

- Bernstein, B. (1973). *Class, codes and control, Applied studies towards a sociology of language*, Londres: Routledge et Kegan Paul.
- Bernstein, B. (1975a). *Class, codes and control, Towards a theory of educational transmissions*, Londres: Routledge et Kegan Paul.
- Bernstein, B. (1975b). *Langage et classes sociales*, Paris: Les éditions de minuit.
- Bernstein, B. (2000). *Pedagogy, symbolic control and identity, Theory, research, critique, Edition révisée*, Lanham: Rowman et Littlefield.
- Bernstein, B. (2007). *Pédagogie, contrôle symbolique et identité, Théorie, recherche, critique, Edition française*, Lévis: Les Presses de l'Université Laval.
- Berrendonner, A. (1986). Notes sur la contre inférence, *Cahiers de linguistique française*, 7, 250-277.
- Berrendonner, A. (1991). Pour une macro-syntaxe, *Travaux de linguistique, Revue internationale de linguistique française*, 21, 25-31.
- Berrendonner, A. (1994). Anaphores confuses et objets indiscrets, In Schnedecker, C., Charolles, M., Kleiber, G., David, J. (eds.), *L'anaphore associative (Aspects linguistiques, psycholinguistiques et automatiques)*, 209-230, Paris: Klincksieck.
- Berrendonner, A. et Reichler Béguelin, M. (1989). Décalages : les niveaux de l'analyse linguistique, *Langue Française*, 81, 99-125.
- Biber, D. (1988). *Variations across speech and writing*, Cambridge: Cambridge University Press.
- Biber, D., Conrad, S. et Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*, Cambridge: Cambridge University Press.
- Blanche-Benveniste, C. (1983). L'importance du français parlé pour la description du français tout court, *Recherches sur le Français Parlé*, 5, 23-45.
- Blanche-Benveniste, C. (1985). La langue du dimanche, *Reflets*, 14, 42-43.
- Blanche-Benveniste, C. (1990). *Le français parlé*, Paris: Edition du Centre National de la Recherche Scientifique.
- Blanche-Benveniste, C. (1995). Le semblable et le dissemblable en syntaxe, *Recherches sur le Français Parlé*, 13, 7-32.
- Blanche-Benveniste, C. (2000). *Approche de la langue parlée en français*, Gap: Ophrys.
- Bloom, L., Miller, P. et Hood, L. (1975). Variation and reduction in the development of organization in language development, In Pick, A. (eds.), *Minnesota symposium on child psychology*, Vol. 9, 88-142, Minneapolis: University of Minnesota Press.
- Bock, K. et Warren, R. (1985). Conceptual accessibility and syntactic structure in sentence formulation, *Cognition*, 21, 47-67.
- Bock, K. et Levelt, W. (1994). Language Production: Grammatical encoding, In Gernsbacher, M. (eds.), *Handbook of Psycholinguistics*, 945-984, San Diego, California: Academic Press.
- Boersma P. et Weenink D. (2009). *Praat: doing phonetics by computer (Version 5.1.05)* [Programme informatique], site internet : <http://www.praat.org/>.
- Bonin, P. (2003). *Production verbale de mots : approche cognitive*, Bruxelles: DeBoeck, Université.

- Bosch, P. (1985). Constraints, coherence, comprehension, reflexion on anaphora, In Sözer, E. (eds.), *Text connexity, text coherence: Aspects, methods, results*, 299-319, Hamburg: Buske.
- Boscolo, P. (1990). The construction of expository text, *First Language*, 10, 217-230.
- Bourdin, B. (1994). *Coût cognitif de la production verbale. Étude comparative oral / écrit chez l'enfant et l'adulte*, Thèse de Doctorat, Université de Bourgogne.
- Bourdin, B. (2002). Gestion de la production et capacité, In Fayol, M. (eds.), *Production du langage*, 148-169, Paris: Lavoisier.
- Bower, G. et Morrow, D. (1990). Mental models in narrative comprehension, *Science*, 247, 44-48.
- Branigan, H., McLean, J. et Reeve, H. (2003). *Something old, something new: addressee knowledge and the given-new contract*, Paper presented at the Proceedings of the Twenty-fifth Annual Conference of the Cognitive Science Society, Boston.
- Brewer, W. (1980). Literacy theory, rhetoric, and stylistics, implications for psychology. , In Spiro, R., Bruce, B., Brewer, W. (eds.), *Theoretical issues in reading comprehension*, 221-237, Hillsdale: Lawrence Erlbaum.
- Britton, B. (1994). Understanding expository text: Building mental structures to induce insights, In Gernsbacher, M. (eds.), *Handbook of Psycholinguistics*, 641-674, San Diego, California: Academic Press.
- Bruner, J. (1986). *Actual Minds, possible worlds*, Cambridge: Harvard University Press.
- Cahana-Amitay, D. et Berman, R. (1999). Initial ideas on L(onger)-Units, *Developing Literacy Across Genres, Modalities, And Languages*, 1, 114-127.
- Caron, J. (1989/2001). *Précis de psycholinguistique*, Paris: Presse Universitaire de France.
- Cautain, C. (2007). *Étude préliminaire des évaluations dans les productions narratives d'élèves de ZEP*, Mémoire de Master 2, Université de Lyon.
- Cautain, C., Chenu, F. et Jisa, H. (2008). Revisions in the written discourse of French children and adolescents, *11th International Conference of the EARLI-SIG Writing* Lund.
- Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics and point of view., In Li, C. (eds.), *Subject and topic*, 25-55, New York: Academic Press.
- Chafe, W. (1994). *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*, Chicago: University of Chicago Press.
- Campbell, A., Brooks, P. et Tomasello, M. (2000). Factors affecting young children's use of pronouns as referring expressions, *Journal of Speech, Language, and Hearing Research*, 43, 1337-1349.
- Chanquoy, L., Foulin, J.N. et Fayol, M. (1990). Temporal management of short text writing by children and adults, *Cahiers de Psychologie Cognitive*, 10, 513-540.
- Chanquoy, L. et Alamargot, D. (2002). Mise en place et développement des traitements rédactionnels : le rôle de la mémoire de travail, *Le Langage et L'homme*, 38, 171-190.
- Charolles, M. (1978). Introduction aux problèmes de la cohérence des textes, *Langue Française*, 38, 7-41.
- Charolles, M. (1988). Les études sur la cohérence et la connexité textuelles depuis la fin des années 1960, *Modèles Linguistiques*, 10, 45-66.
- Charolles, M., et Ehrlich, M.F. (1991). Aspects of textual continuity, linguistic approaches, In Denhière, G., Rossi, J.P. (eds.), *Text and text processing*, 251-267, Amsterdam: North-Holland.

- Chaudron, C. et Richards, J. (1986). The effects of discourse markers on the comprehension of lectures, *Applied Linguistics*, 7, 113-127.
- Chauveau, G. (2000). « Réussir en ZEP : l'exemple de la lecture », *Comment réussir en ZEP : vers des zones d'excellence pédagogique*, Paris: Retz.
- Chesnet, D., Guillabert, F. et Esperet, E. (1994). G-Studio : Un logiciel pour l'étude en temps réel des paramètres temporels de la production écrite, *L'Année psychologique*, 94, 283-294.
- Chesnet, D. et Alamargot, D. (2005). Analyses en temps réel des activités oculaires et graphomotrices du scripteur: intérêt du dispositif 'Eye and Pen', *L'Année Psychologique*, 105, 477-520.
- Cho, S. (2004). *Argument ellipsis in Korean-Speaking children's early speech*, Harvard University.
- Chomsky, N. (1957). *Syntactic structures*, The Hague: Mouton.
- Clancy, P. (1993). *Preferred argument structure in Korean acquisition proceedings of the*, Paper presented at the 25th Annual Child Language Research Forum, Stanford University.
- Clancy, P. (1997). Discourse motivations for referential choice in Korean Acquisition, In Sohn, H.M., Haig, J. (eds.), *Japanese/Korean Linguistics*, Vol. 6, 639-59, Stanford: CSLI Publications.
- Clancy, P. (2003). The lexicon in interaction: developmental origins of Preferred Argument Structure in Korean, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 81-108, Amsterdam: John Benjamins.
- Clark, H. (1996). *Using language*, Cambridge: Cambridge University Press.
- Clark, H. et Haviland, S. (1977). Comprehension and the Given-New contract, In Freedle, R. (eds.), *Discourse Production and Comprehension*, 1-40, Norwood, NJ: Ablex.
- Clark, H. et Murphy, G. (1982). Audience design in meaning and reference, In LeNy, J., Kintsch, W. (eds.), *Language and comprehension*, 287-299, Amsterdam: North-Holland.
- Clark, H. et Schaefer, E. (1989). Contributing to Discourse, *Cognitive Science*, 13, 259-294.
- Clark, H. et Wilkes-Gibbs, D. (1986). Referring as a collaborative process, *Cognition*, 22, 1-40.
- Coirier, P., Gaonac'h, D. et Passerault, J.M. (1996). *Psycholinguistique textuelle*, Paris: Armand Colin Édition.
- Coirier, P., et Marchand, E. (1994). Writing argumentative texts, a typological and structural approach, In Eigler, G., Jechle, T. (eds.), *Writing, Current Trends in European Research*, 163-181, Freiburg: Hochschul Verlag.
- Collectif, (1982). Circulaire n° 81-536 du 28 décembre 1981 (B.O.spécial n°1 du 21-01- 1982).
- Collectif, (1989). *Apprendre à lire et à écrire : dix ans de recherche sur la lecture et la production des textes*, dans la revue française de pédagogie, INRP (Institut National de la Recherche Pédagogique), Paris.
- Collectif, (1991). *Les cycles à l'école primaire : Une école pour enfant des outils pour les maîtres*, Ministère de l'Éducation Nationale de la Jeunesse et des Sports et Direction des Écoles, CNDP (Centre National de Recherche Pédagogique), Paris : Hachette École.
- Collectif, (1992). *La maîtrise de la langue à l'école : Une école pour enfant des outils pour les maîtres*, Ministère de l'Éducation Nationale et de la Culture et Direction des Écoles, CNDP (Centre National de Recherche Pédagogique), Paris : Savoir Livre.
- Collectif, (1994). *Lire et écrire à l'école primaire*, INRP (Institut National de la Recherche Pédagogique), Paris.

- Collectif, (1995). *Programmes à l'école primaire : Une école pour enfant des outils pour les maîtres*, Ministère de l'Éducation Nationale et Direction des Écoles, CNDP (Centre National de Recherche Pédagogique), Paris : Savoir Livre.
- Collectif, (1997). *La maîtrise de la langue au collège*, Ministère de l'Éducation Nationale de l'Enseignement Supérieur et de la Recherche, Direction des lycées et des collèges CNDP (Centre National de Recherche Pédagogique), Paris : Savoir Livre.
- Collectif, (1997). *Le système éducatif de la France*, Ministère de l'Éducation Nationale de l'Enseignement Supérieur et de la Recherche et de la Technologie, Réimprimé avec l'autorisation du CRDP (Centre Régional de Documentation Pédagogique), Lyon.
- Corblin, F. (1991). *Sujet impersonnel et sujet indistinct : il et ça*, Paper presented at the L'impersonnel, Mécanismes linguistiques et fonctionnements littéraires, Grenoble.
- Corston-Oliver, S. (2003). Core arguments and the inversion of the nominal hierarchy in Roviana, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 273-300, Amsterdam: John Benjamins.
- Coveney, A. (2004). The alternation between l'on and on in spoken French, *Journal of French Language Studies*, 14, 91-112.
- Craig, C. (1987). The Rama language: A text with grammatical notes, *Journal of Chibchan Studies* 5.
- Creissels, D. (2004). Cours de Syntaxe de Sciences Du Langage, Université Lumière, Lyon 2.
- Creissels, D. (2006a). *Syntaxe générale une introduction typologique 1 : catégories et constructions*, Vol. 1. Paris: Hermès - Lavoisier.
- Creissels, D. (2006b). *Syntaxe générale une introduction typologique 2 : la phrase*, Vol. 2. Paris: Hermès - Lavoisier.
- De Weck, G. (1991). *La cohésion dans les textes d'enfants, étude du développement des processus anaphoriques*, Neuchâtel: Delachaux et Niestlé.
- De Weck, G. et Schneuwly, B. (1994). Anaphoric procedures in four text types written by children, *Discourse Processes*, 17, 101-113.
- Demont, E. (2007). L'acquisition du langage, *Psychologie Française*, 52, 9-12.
- Dockrell, J. et Messer, D. (2004). Lexical acquisition in the early school years, In Berman, R. (eds.), *Language Development across Childhood and Adolescence*, Vol. 3. 35-52, Amsterdam/Philadelphia: John Benjamins.
- Downing, P. (1985). *Classifier constructions and referential marking in Japanese*, Paper presented at the Conference on Japanese Language and Linguistics, UCLA.
- Du Bois, J. (1987). The discourse basis of ergativity, *Language*, 63, 805-855.
- Du Bois, J. (2003). Argument Structure: Grammar in use, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 11-60, Amsterdam: John Benjamins.
- Du Bois, J., Kumpf, L., Ashby, W. (2003a). Introduction, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 1-10, Amsterdam: John Benjamins.
- Du Bois, J., Kumpf, L. et Ashby, W. (Eds.) (2003b). *Preferred Argument Structure: Grammar as architecture for function*, Amsterdam: John Benjamins.

- Duez, D. (1991). *La pause dans la parole de l'homme politique*, Paris: Edition du CNRS.
- Duez, D. (2001). Caractéristiques acoustiques et phonétique des pauses remplies dans la conversation en français, *Travaux Interdisciplinaires du Laboratoire Parole et Langage*, 20, 31-48.
- Durie, M. (2003). New light on information pressure: Information conduits, "escape valves", and role alignment stretching, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 159-196, Amsterdam: John Benjamins.
- Durie, M. (1985). *A grammar of Acehnese*, Dordrecht: Foris Publications.
- Dutra, R. (1987). The hybrid S category in Brazilian Portuguese: some implications for word order, *Studies in Language*, 11, 163-181.
- England, N. (1983). *Mamean voice : Syntactic and narrative considerations*, University of Iowa.
- England, N. et Martin, L. (2003). Issues in the comparative argument structure analysis in Mayan narratives, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 131-158, Amsterdam: John Benjamins.
- Favart, M. (2005). Les marques de cohésion : leur rôle fonctionnel dans l'acquisition de la production écrite de texte, *Psychologie française*, 50, 305-322.
- Favart, M. et Olive, T. (2005). Modèles et méthodes d'étude de la production écrite, *Psychologie Française*, 50, 273-285.
- Favart, M. et Passerault, J. (1996). Functionality of Cohesion devices in the management of Local and Global coherence: two studies in children's written production of narratives, In Rijlaarsdam, R., Van den Berg, H., Jechle, T. (eds.), *Theories, Models and Methodologies in Writing*, 349-365, Amsterdam: Amsterdam University Press.
- Favart, M. et Coirier, P. (2006). Acquisition of the linearization process in text composition in third to ninth graders: effects of textual superstructure and macrostructural organization, *Journal of Psycholinguistic Research*, 35, 305-328.
- Fayol, M. (1984). La distanciation dans le langage, *Enfance*, 1, 5-19.
- Fayol, M. (1985). *Le récit et sa construction, une approche de psychologie cognitive*, Paris: Delachaux et Niestlé.
- Fayol, M. (1987). Vers une psycholinguistique textuelle génétique, L'acquisition du récit, In Piérait-Le Bonniec, G. (eds.), *Connaître et le dire*, 223-238, Bruxelles: Pierre Mardaga.
- Fayol, M. (1989). Une approche psycholinguistique de la ponctuation. Etude en production et en compréhension, *Langue Française*, 81, 21-39.
- Fayol, M. (1990). La production de textes écrits, Introduction à l'approche cognitive, *Education Permanente*, 102, 21-30.
- Fayol, M. (1991). Text typologies: a cognitive approach, In Denhière, G., Rossi, J. (eds.), *Text and text processing*, 61-76, Amsterdam : North-Holland.
- Fayol, M. (1996). La production d'écrits narratifs : Approche de psycholinguistique textuelle chez l'enfant et l'adulte, In David, J., Plane, S. (eds.), *L'apprentissage de l'écriture de l'école au collège*, 9-36, Paris: Presses Universitaires de France.
- Fayol, M. (1997). *Des idées au texte : psychologie cognitive de la production verbale orale et écrite*, Paris: Presse Universitaire de France.

- Fayol, M. (2000). Comprendre et produire des textes écrits : l'exemple du récit, In Kail, M., Fayol, M. (eds.), *L'acquisition du langage, le langage en développement au-delà de trois ans*, 183-214, Paris: Presses Universitaires de France.
- Fayol, M. (2002). Conclusion, In Fayol, M. (eds.), *Production du langage*, 287-301, Paris: Lavoisier.
- Fayol, M. (2002). Introduction, In Fayol, M. (eds.), *Production du langage*, 17-24, Paris: Lavoisier.
- Fayol, M. (Eds.) (2002). *Production du langage*, Paris: Lavoisier.
- Fayol, M. et Abdi, H. (1988). The influence of script structure on punctuation, *CPC/European Bulletin of Cognitive Psychology*, 8, 269-279.
- Fayol, M. et Abdi, H. (1990). Ponctuation et connecteurs : étude expérimentale de leur fonctionnement dans les paires de propositions, In Charolles, M., Fisher, S., Jayez, J. (eds.), *Le discours*, 167-180, Nancy: P.U.N.
- Fayol, M., et Heurley, L. (1995). Des modèles de production du langage à l'étude du fonctionnement du scripteur, enfant et adulte, In Boyer, J.Y., Dionne, J.P. (eds.), *La production de textes : vers un modèle d'enseignement de l'écriture*, 17-48, Montréal: Logiques.
- Fayol, M., Jisa, H. et Mazur-Palandre, A. (2007). A developmental study of referential cohesion, *The first conference of the swedish association for language and cognition* Lund.
- Fayol, M., Jisa, H. et Mazur-Palandre, A. (2008). Information flow in written and spoken French: A developmental perspective, *Writing Research Across Borders* University of California, Santa Barbara.
- Fayol, M. et Miret, A. (2005). Ecrire, orthographier et rédiger des textes, *Psychologie Française*, 50, 391-402.
- Ferrand, L. (2002). Les modèles de la production de la parole, In Fayol, M. (eds.), *Production du langage*, 27-44, Paris: Lavoisier.
- Ferreira, F. et Anes, M. (1994). Why study spoken language?, In Gernsbacher, M. (eds.), *Handbook of Psycholinguistics*, 33-56, San Diego, California: Academic Press.
- Flower, L. et Hayes, J. (1981). A cognitive process theory of writing, *College Composition and Communication*, 32, 385-387.
- Fodor, J. (1983/1986). *La modularité de l'esprit, Essai sur la psychologie des facultés*, Paris: Les éditions de minuit.
- Fodor, J. (2000). *The mind doesn't work that way*, Massachusetts: MIT.
- Fodor, J. (2003). *L'esprit, ça ne marche pas comme ça*, Paris: Odile Jacob.
- Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology*, Cambridge, MA: MIT Press.
- Foulin, J.N. (1995). Pauses et débits : les indicateurs temporels de la production écrite, *L'année Psychologique*, 95, 483-504.
- Foulin, J.N. (1998). To what extent does pause location predict pause duration in adults and children's writing, *Cahiers de Psychologie Cognitive*, 17, 601-620.
- Foulin, J.N. et Fayol, M. (1988). Étude en temps réel de la production écrite chez des enfants de sept et huit ans, *European Journal of Psychology of Education*, 3, 461-475.
- Fox, B. (1987). *Discourse structure and anaphora*, Vol. 48. Cambridge: Cambridge University Press.
- Francis, G. (1986). *Anaphoric nouns*, Vol. 11. English Language Research: University of Birmingham.

- François, D. (1974). *Français parlé. Analyse des unités phoniques et significatives d'un corpus recueilli dans la région parisienne*, Paris: S.E.L.A.F.
- Gaatoone, D. (1991, 1991). *Il y a impersonnel et impersonnel : sur les contraintes et les finalités de la construction impersonnelle*, Paper presented at the L'impersonnel, Mécanismes linguistiques et fonctionnements littéraires, Grenoble.
- Gadet, F. (1989). *Le français ordinaire*, Paris: Armand colin éditeur.
- Gadet, F. (1996). Une distinction bien fragile : oral/écrit, *TRANEL (Travaux neuchâtelois de linguistique)*, 25, 13-27.
- Garrett, M. (1975). The analysis of sentence production, In Bower, G. (eds.), *The psychology of learning and motivation*, 133-177, New York: academic Press.
- Garrett, M. (1982). Production of speech : Observation from normal and pathological use, In Ellis, A. (eds.), *Normality and pathology in cognition functions*, 19-76, London: Academic Press.
- Garrett, M. (1984). The organisation of processing structure for language production: Applications to aphasic speech, In Caplan, D., Lecours, A., Smith, A. (eds.), *Biological Perspectives on Language*, 172-193, Cambridge : MIT Press.
- Garrett, M. (1988). Processes in language production, In Newmayer, J. (eds.), *Language: Psychological and biological aspects*, 69-95, Cambridge: Cambridge University Press.
- Garrett, M. (1980a). Levels of processing in sentence production, In Butterworth, B. (eds.), *Language Production*, Vol. 1, 117-220, New York : Academic Press.
- Garrett, M. (1980b). The limits of accommodation: Arguments for independent processing levels in sentence production, In Fromkin, V. (eds.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand*, 263-271, New York: Academic Press.
- Gathercole, S. et Baddeley, A. (1993). *Working memory and language*, Hove, UK ; Hillsdale, USA: L. Erlbaum Associates.
- Gayraud, F. (2000). *Le développement de la différenciation oral/écrit vu à travers le lexique*, Thèse de Doctorat, Université Lumière, Lyon 2.
- Gayraud, F., Jisa, H. et Viguié, A. (1999). The development of syntactic packaging in French children's written and spoken texts, *Developing Literacy Across Genres, Modalities, And Languages*, 1, 169-181.
- Gayraud, F., Jisa, H. et Viguié, A. (2001). Utilisation des outils cohésifs comme indice de sensibilité au registre : une étude développementale, *Acquisition et Interaction en Langue Étrangère*, 14, 3-24.
- Genetti, C. et Crain, L. (2003). Beyond Preferred Argument Structure: Sentences, pronouns, and given, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 197-224, Amsterdam: John Benjamins.
- Gernsbacher, M. (1990). *Language comprehension as structure building*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gernsbacher, M. (1994). *Handbook of psycholinguistics*, San Diego, California: Academic Press.
- Gernsbacher, M. et Givón, T. (1995). Introduction, coherence as a mental entity, In Gernsbacher, M., Givón, T. (eds.), *Coherence in spontaneous text*, VII-X, Philadelphia: John Benjamins.

- Gernsbacher, M. et Hargreaves, D. (1988). Accessing sentence participants ; the advantage of first mention, *Journal of Memory and Language*, 27, 699-717.
- Gineste, M.D. et Le Ny, J.F. (2002). *Psychologie cognitive du langage*, Paris: Dunod.
- Givón, T. (1995). Coherence in text vs. Coherence in mind, In Gernsbacher, M., Givón, T. (eds.), *Coherence in spontaneous text*, 59-116, Philadelphia: John Benjamins.
- Goosse, A. (1991). *Mélanges de grammaire et de lexicologie françaises*, Leuven et Louvain-la-neuve: Peeters et Publications linguistiques de Louvain.
- Goosse, A. (1993/2004). *Le bon usage : grammaire française*, Paris: De Boeck Duculot.
- Graham, S., Harris, K. et Chorzempa, B. (2002). Contribution of spelling instruction to the spelling, writing, and reading of poor spellers, *Journal of Educational Psychology*, 94, 669–689.
- Graham, S., Harris, K. et Kink, B. (2000). Is handwriting causally related to learning to write? Treatment of handwriting problems in beginning writers, *Journal of Educational Psychology*, 92, 620–633.
- Gregory, M. (1967). Aspects of varieties differentiation, *Journal of Linguistics*, 3, 177-98.
- Gregory, M. et Carroll, S. (1978). *Language and situation : language varieties and their social contexts*, London/Boston: Routledge and Kegan Paul.
- Guerin, E. (2006). *Introduction de la notion de variation situatiolectale dans la grammaire scolaire par la caractérisation de deux opérateurs pragmatiques : on et ça*, Paris X - Nanterre, Paris.
- Guerreiro, S., Oshima-Takane, Y. et Kuriyama, Y. (2006). The development of referential choice in english and japanese : a discourse-pragmatic perspective, *Journal of child language*, 33, 823-57.
- Halliday, M. (1985). *An introduction to functional grammar*, London: Edward Arnold.
- Halliday, M. (1989). *Spoken and written language (2nd ed.)*, Oxford: Oxford University Press.
- Halliday, M. et Hasan, R. (1976/1989). *Cohesion in English*, Londres: Longman group limited.
- Halliday, M. et Hasan, R. (1985). *Language, context, and text: Aspects of language in a social-semiotic perspective*, Victoria: Deakin University.
- Haviland, S. et Clark, H. (1974). What's New? Acquiring New Information as a Process in Comprehension, *Journal of verbal learning and verbal behavior*, 13, 512-521.
- Hawkins, P. (1973). Social class, the nominal group and reference, In Bernstein, B. (eds.), *Class, codes and control, Applied studies toward a sociology of language*, 81-92, Londres: Routledge et Kegan Paul.
- Hawkins, J. (1983). *Word order universals*, New York: Academic press.
- Hayes, J. (1995). Un nouveau modèle du processus d'écriture, In Boyer, J.Y., Dionne, J.P. (eds.), *La production de textes : vers un modèle d'enseignement de l'écriture*, 49-72, Montréal: Logiques.
- Hayes, J. (1996). A new framework or understanding cognition and affect in writing, In Levy, M., Ransdell, S. (eds.), *The Science of Writing: Theories, Methods, Individual differences*, 1–27, Mahwah: Lawrence Erlbaum Associates.
- Hayes, J. (1998). Un nouveau cadre pour intégrer cognition et affect dans la rédaction, In Piolat, A., Pélissier, A. (eds.), *la rédaction de textes*, 51-101, Lausanne: Delachaux et Niestlé.

- Hayes, J. et Flower, L. (1980). Identifying the organization of writing processes, In Gregg, L., Steinberg, E. (eds.), *Cognitive processes in writing*, 3-30, NJ : Hillsdale.
- Hayes, J. et Flower, L. (1986). Writing research and the writer, *American psychologist*, 41, 1106-1113.
- Herslund, M. (2003). Article et pronom. Réflexions sur le syntagme nominal, In Hadermann, P., Van Slijcke, A., Berré, M. (eds.), *La syntaxe raisonnée. Mélanges de linguistique générale et française offerts à Annie Boone à l'occasion de son 60e anniversaire*, 105-116, Bruxelles: De Boeck et Larcier.
- Hickmann, M. (1980). Creating referents in discourse: A developmental analysis of linguistic cohesion, In Kreiman, J., Ojeda, E. (eds.), *Papers from the Parasession on pronouns and Anaphora*, 192-203, Chicago: Chicago Linguistics Society.
- Hickmann, M. (1987). Ontogenèse de la cohésion dans le discours, In Piérait-Le Bonniec, G. (eds.), *Connaître et le dire*, 239-262, Bruxelles: Pierre Mardaga.
- Hickmann, M. (1995). Discourse organisation and the development of reference to person, Space and time, In Fletcher, P., MacWhinney, B. (eds.), *The Handbook of child language*, 194-218, Oxford: Blackwell.
- Hickmann, M. (2000). Le développement de l'organisation discursive, In Kail, M., Fayol, M. (eds.), *L'acquisition du langage, le langage en développement au-delà de trois ans*, 83-115, Paris: Presses Universitaires de France.
- Hickmann, M. (2002). Développement de la production verbale orale, In Fayol, M. (eds.), *Production du langage*, 173-203, Paris: Lavoisier.
- Hickmann, M. (2003). *Children's discourse : person, space and time across languages*, Cambridge ; New York: Cambridge University Press.
- Hickmann, M. (2004). Coherence, cohesion and context, Some comparative perspectives in narrative development, In Strömquist, S., Verhoeven, L. (eds.), *Relating events in narrative, Typological and contextual perspectives*, 281-306, Mahwah: Lawrence Erlbaum Associates.
- Hickmann, M. et Hendriks, H. (1999). Cohesion and anaphora in children's narratives : a comparison of English, French, German and Mandarin Chinese, *Journal of Child Language*, 26, 419-452.
- Hickmann, M., Hendriks, H., Roland, F. et Liang, J. (1996). The marking of new information in children's narratives: a comparison of English, French, German, and Mandarin Chinese, *Journal of child Languages*, 23, 591-619.
- Hickmann, M., Kail, M. et Roland, F. (1995a). Cohesive anaphoric relations in french children's narratives as a function of mutual knowledge, *First Language*, 15, 277-300.
- Hickmann, M., Kail, M. et Roland, F. (1995b). L'organisation référentielle dans les récits d'enfants en fonction des contraintes contextuelles, *Enfance*, 2, 215-226.
- Hoc, J.M. (1992). *Psychologie cognitive de la planification*, Grenoble: Presse Universitaire de Grenoble.
- Hofling, C. (2003). Trackling the deer: Nominal reference, parallelism and preferred Argument Structure in Itzaj Maya narrative genres, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 131-158, Amsterdam: John Benjamins.
- Hopper, P. (1998). Emergent grammar and the a priori grammar postulate, In Tannen, D. (eds.), *Linguistics in Context: Connecting Observation and Understanding*, 117-34, Norwood : Ablex.
- Hudelot, C. (1980). Qu'est-ce que la complexité syntaxique? L'exemple de la relative, *La Linguistique*, 16, 5-41.

- Hughes, M., Allen, S. (2006). *A discourse-pragmatic analysis of subject omission in child English*, Paper presented at the 30th Annual Boston University Conference on Language Development, Boston.
- Hunt, K. (1970). Syntactic Maturity in School children and Adults, *Monographs of the Society for Research in Child Development*, 35, iii-67.
- Hurell, L. (1956). *A comparison of the development of oral and written language in school age children*, Indiana: Purdue University, Lafayette.
- Iwasaki, S. (1985, 1985). *The "Given A Constraint" and the Japanese particle ga*, Paper presented at the First Annual Pacific Linguistics Conference, University of Oregon.
- Jeanjean, C. (1980). *Les formes sujets de type nominal : étude sur le français contemporain*, Université de Provence.
- Jisa, H. (1998a). Relative clauses in French children's narrative text, *Journal of Child Language*, 25, 623-652.
- Jisa, H. (2000a). Development of subject ellipsis in French narrative discourse: reference in syntax, *Linguistics*, 38, 591-620.
- Jisa, H. (2000b). Increasing cohesion in narratives: a developmental study of maintaining and reintroducing subjects in French, *Linguistics*, 38, 591-620.
- Jisa, H. (2004). Growing into academic French, *Later Language Development: Typological and psycholinguistic Perspectives. TILAR*, 3, 135-162.
- Jisa, H. et Kern, S. (1998b). Relative clauses in French children's narrative texts, *Journal of Child Language*, 25, 623-652.
- Jisa, H. et Mazur, A. (2005). The structure and content of noun phrases in French, *Xth International Congress for the Study of Child Language*, Berlin.
- Jisa, H. et Mazur, A. (2006). Les sujets lourds: considérations développementales en français oral et écrit, *Méthodologies actuelles dans les recherches sur les acquisitions tardives en langue maternelle*, Université René Descartes, Paris 5.
- Jisa, H. et Mazur, A. (2006). *L'expression de la causalité : une étude développementale*, Paper presented at the Journée d'étude : Des savoirs savants aux savoirs enseignés, Université Paris X, Nanterre, 33-60.
- Jisa, H., Reilly, J., Verhoeven, L., Baruch, E. et Rosado, E. (2002). Cross-linguistic perspectives on the use of passive constructions in written texts, *Journal of Written Language and Literacy*, 5, 163-181.
- Jisa, H., Strömquist, S. (2002). Pragmatique et acquisition tardives chez l'enfant : langage oral et langage écrit, In Bernicot, J., Trognon, A., Guichetti, M., Musiol, M. (eds.), *La pragmatique : raisonnement, développement de l'enfant et pathologies*, 257-265, Nancy: Presses Universitaires de Nancy.
- Jisa, H. et Viguié, A. (2005). Developmental perspectives on the role of French on in written and spoken expository texts, *Journal of Pragmatics*, 37, 125-142.
- Johansson, V. (1999). Word frequencies in speech and Writing: a study of expository discourse, *Genre and modality in developing discourse abilities*, 1, 182-198.
- Johnson-Laird, P. (1980). Mental models in cognitive science, *Cognitive Science*, 4, 71-115.
- Johnson-Laird, P. (1983). *Mental models*, Cambridge: Cambridge University Press.
- Jones, M. (1996). *Foundations of French Syntax*, Cambridge: Cambridge University Press.

- Jones, D. et Christensen, C. (1999). Relationship between automaticity in handwriting and students' ability to generate written text, *Journal of Educational Psychology*, 91, 44-49.
- Just, M. et Carpenter, P. (1992). A capacity theory of comprehension : Individual differences in working memory, *Psychological Review*, 99, 122-149.
- Kail, M. et Fayol, M. (2000a). Introduction, In Kail, M., Fayol, M. (eds.), *L'acquisition du langage, le langage en développement au-delà de trois ans*, 1-8, Paris: Presses Universitaires de France.
- Kail, M. et Fayol, M. (Eds.) (2000b). *L'acquisition du langage, Le langage en développement au-delà de trois ans*, Paris: Presses Universitaires de France.
- Kail, M. et Hickmann, M. (1992). French children's ability to introduce referents in narratives as a function of mutual knowledge, *First Language*, 12, 637-62.
- Kail, M. et Hickmann, M. (1992). On French children's ability to introduce referents in discourse as function of mutual knowledge, *First Language*, 12, 73-94.
- Karmiloff-Smith, A. (1979a). *A functional approach to language: a study of determiners and reference*, Cambridge: Cambridge University Press.
- Karmiloff-Smith, A. (1979b). Micro- and Macrodevelopmental changes in language acquisition and other representational systems, *Cognitive science*, 3, 91-118.
- Karmiloff-Smith, A. (1981a). Getting developmental differences or studying child development?, *Cognition*, 10, 151-158.
- Karmiloff-Smith, A. (1981b). The grammatical marking of thematic structure in the development of language production, In Deutsch, W. (eds.), *The child's construction of language*, 121-147, New York: Academic Press.
- Karmiloff-Smith, A. (1985). A constructivist approach to modelling linguistic and cognitive development, *Archives de Psychologie*, 53, 113-126.
- Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data, *Cognition*, 23, 95-147.
- Karmiloff-Smith, A. (1990). Constraint on representational change: Evidence from children's drawing, *Cognition*, 34, 57-83.
- Karmiloff-Smith, A. (1993/1995). *Beyond Modularity: A developmental perspective on cognitive science*, Massachusetts: The MIT Press.
- Karmiloff, M. et Karmiloff-Smith, A. (2001). *Pathways to language: from fetus to adolescent*, Harvard University Press.
- Katzenberger, I. (1999). Further ideas on (l)onger-Units, *Developing Literacy Across Genres, Modalities, And Languages*, 1, 128-141.
- Katzenberger, I. (2004). The development of clause packaging in spoken and written texts, *Journal of Pragmatics*, 36, 1921-1948.
- Katzenberger, I. (2005). The super-structure of written expository texts – a developmental perspective, In Ravid, D., Shyldkrot, H.B. (eds.), *Perspectives on language and language development. Essays in Honor of Ruth A. Berman*, 327-336, Dordrecht: Kluwer Academic Publishers.
- Keenan, E. et Comrie, B. (1977). Noun phrase accessibility and universal grammar, *Linguistic Inquiry*, 8, 63-99.

- Kellogg, R. (1987). Effects of topic knowledge on the allocation of processing time and cognitive effort to writing processes, *Memory and Cognition*, 15, 256-266.
- Kellogg, R. (1994). *The psychology of writing*, New York: Oxford University Press.
- Kellogg, R. (1998). Un modèle de la mémoire de travail dans la rédaction, In Piolat, A., Pélissier, A. (eds.), *La rédaction de textes, Approche cognitive*, 102-135, Lausanne: Delachaux et Niestlé.
- Kellogg, R. (2008). Training writing skills: A cognitive developmental perspective, *Journal of*, 1, 1-26.
- Kern, S. (1997). *Comment les enfants jonglent avec les contraintes communicationnelles, discursives et linguistiques dans la production d'une narration*, Thèse de Doctorat, Université Lumière, Lyon.
- Khorounjaia, E. et Tolchinsky, L. (2004). Discursive constraints on the lexical realization of arguments in Spanish, In Berman, R. (eds.), *Language Development across Childhood and Adolescence*, Vol. 3, 83-110, Amsterdam/Philadelphia: John Benjamins.
- Kintsch, W. (1974). *The representation of meaning in memory*, New York: Wiley and Sons.
- Kociemba, V. (2004). *Aider les élèves en difficulté, Enseigner en ZEP. Sous la direction de La Bordeire R.*, Paris: Bordas pédagogie.
- Kohn, S. et Cragolino, A. (2003). The role of Preferred Argument Structure for understanding aphasic sentence planning, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 339-352, Amsterdam: John Benjamins.
- Kumpf, L. (2003). Genre and Preferred Argument Structure: Sources of argument structure in classroom discourse, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure, Grammar as architecture for function*, 109-130, Amsterdam: John Benjamins.
- Labov, W. (1976). *Sociolinguistique*, Paris: Les éditions de minuit.
- Labov, W. (1978/1993). *Le parler ordinaire, la langue dans les ghettos noirs des états-unis*, Paris: Les éditions de minuit.
- Labov, W. et Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience, In Helms, J. (eds.), *Essays on the verbal and visual arts*, 12-44, Seattle: University of Washington Press.
- Lambrecht, K. (1987). On the Status of SVO Sentences in French Discourse, In Tomlin, R. (eds.), *Coherence and Grounding in Discourse*, 217-261, Amsterdam: John Benjamins.
- Lambrecht, K. (1994). *Information structure and sentence form: topic, focus and mental representations of discourse referents*, Cambridge: Cambridge University Press.
- Le Ny, J.F. (1979). *La sémantique psychologique*, Paris: Presses Universitaires de France.
- Leeman-Bouix, D. (1994). *Grammaire du verbe français, des formes au sens*, Tours : Nathan Université.
- Léger, A. (non publié). Le débat Bernstein/Labov : différences langagières ou inégalité, Université de Caen.
- Lehmann, A. et Martin-Berthet, F. (2000). *Introduction à la lexicologie, Sémantique et morphologie*, Liège: Nathan Université.
- Lemaire, P. (1999). *Psychologie cognitive*, Bruxelles: De boeck Université.
- Leuwens, C. (2002). La production de phrases, In Fayol, M. (eds.), *Production du langage*, 107-130, Paris: Lavoisier.

- Levelt, W. (1989). *Speaking: From intention to articulation*, Cambridge/Massachusetts: The MIT Press.
- Levelt, W. (1995). The ability to speak: from intentions to spoken words, *European Review*, 3, 13-23.
- Levelt, W. (1999). Producing spoken language: a blueprint of the speaker, In Brown, C., Hagoort, P. (eds.), *The neurocognition of language*, 94-122, Oxford: Oxford University Press.
- Loban, W. (1976). Language Development : Kindergarten through Grade Twelve, *National Council of Teachers of English*, 18.
- MacArthur, C., Graham, S. et Fitzgerald, J. (Eds.) (2006). *Handbook of Writing Research*, New York: The Guilford Press.
- Maggio, S. (2007). *La dynamique de la production verbale écrite chez les enfants de 10 à 16 ans*, Mémoire de Master, Université Blaise Pascal.
- Maggio, S., Jisa, H., Lété, B. et Fayol, M. (2008). On managing spelling when writing texts, *Writing Research Across Borders*, Santa Barbara.
- Maggio, S., Lété, B. et Fayol, M. (2008). On managing spelling when writing texts, *Sig writing*, Lund.
- Maillard, M. (1985). L'impersonnel français de « il » à « ça », In Chocheyras, J. (eds.), *Autour de l'impersonnel*, 63-118, Grenoble: Ellug.
- Mandler, J. (1982). Recent research on story grammars, In Le Ny, J. et Kintsch, W. (eds.), *Language and comprehension*, 207-219, Amsterdam: North Holland.
- Martin, L. (2003). Narrator virtuosity and the strategic exploitation of Preferred Argument Structure in Mocho: Repetition and constructed speech in Mocho narrative, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 411-436, Amsterdam: John Benjamins.
- Matsushashi, A. (1981). Pausing and planning: the tempo of written discours production, *Research in the teaching of english*, 15, 113-134.
- Matthews, D., Lieven, E., Theakston, A. et Tomasello, M. (2006). The effect of perceptual availability and prior discourse on young children's use of referring expressions, *Applied Psycholinguistics*, 27, 403-422.
- Matthey, M. (2003). *Apprentissage d'une langue et interaction verbale*, Bern: Peter lang SA.
- Mazur, A. (2005). *Etude développementale de la complexité des syntagmes nominaux lexicaux en position sujet*, Mémoire de DEA, Université Lumière, Lyon 2.
- Mazur-Palandre, A. (2007). *Analyse syntaxique de textes produits dans différents contextes de production*, Paper presented at the Colloque jeunes chercheurs de l'école doctorale "Langage et langues", "Normes, variations, identité et altérité", Institut de Linguistique et Phonétique Générale et Appliquée.
- Mazur-Palandre, A. (2007). *Later language development: syntactic packaging in written and spoken French*, Paper presented at the The Second Oxford Linguistics Postgraduate Conference, Oxford.
- Mazur-Palandre, A. (2007). *Les connexions syntaxiques : considérations contextuelles et développementales*, Paper presented at the Colloque jeunes chercheurs Praxiling Analyses du discours et contextes : quelles relations?, Université Paul Valéry, Montpellier III.
- Mazur-Palandre, A. (2008a). Le flux de l'information: une étude développementale, *Dans le cadre des séminaires internes de DDL*, Lyon.

- Mazur-Palandre, A. (2008b). Referential cohesion in written expository and narrative: A developmental study, *Sigwriting 2008, The 11th international conference of the EARLI, social interest group on writing*, Lund University.
- McCutchen, D. (1988). "Functional automaticity" in children's writing: A problem of metacognitive control, *Written Communication*, 5, 306-324.
- McCutchen, D. (1996). A capacity theory of writing: Working memory in composition, *Educational Psychology Review*, 8, 299-325.
- McCutchen, D. (1998). Une théorie de la capacité pour la rédaction : La mémoire de travail dans la rédaction, In Piolat, A., Pélissier, A. (eds.), *La rédaction de textes, Approche cognitive*, 183-224, Lausanne: Delachaux et Niestlé.
- McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing, *Educational Psychologist*, 35, 13-23.
- Mishina-Mori, S. (2007). *Argument representation in Japanese/English simultaneous bilinguals: Is there a transfer?*, Paper presented at the 30th Annual Boston University Conference on Language Development, Boston.
- Moirand, S. (1975). Le rôle anaphorique de la nominalisation dans la presse écrite, *Langue Française*, 28, 60-78.
- Monnier, C. (2008). La mémoire de travail : approche théorique et développement, *Psychologie française*, 53, 279-280.
- Mosenthal, P. (1985). Defining the expository discourse continuum, toward a taxonomy of expository text, *Poetics*, 14, 387-414.
- Mounin, G. (Eds.) (1974/2004). *Dictionnaire de la linguistique*, Paris: PUF.
- Narasimhan, B., Budwig, N. et Murty, L. (2005). Argument Realization in Hindi Caregiver-Child Discourse, *Journal of Pragmatics*, 37, 461-95.
- Nichols, S., Jones, W., Roman, M., Wulfeck, B., Delis, D., Reilly, J. et Bellugi, U. (2004). Mechanisms of verbal memory impairment in four neurodevelopmental disorders, *Brain and Language*, 88, 180-189.
- Nippold, M. (2002). Lexical learning in school-age children, adolescents, and adults: a process where language and literacy converge, *Journal of Child Language*, 29, 474-478.
- Nippold, M. (2004). Research on later language development: International perspectives, In Berman, R. (eds.), *Language Development across Childhood and Adolescence*, Vol. 3, 1-8, Amsterdam/Philadelphia: John Benjamins.
- O'Donnell, R., Griffin, W. et Norris, R. (1967). *Syntax of kindergarden and elementary school children : a transformational analysis*, Illinois: National Council of Teachers of English.
- Olive, T. (2004). Working Memory in Writing: Empirical Evidence From the Dual-Task Technique, *European Psychologist*, 9, 32-42.
- Olive, T. et Piolat, A. (2003). Activation des processus rédactionnels et qualité des textes, *Le Langage et l'Homme*, XXXVIII.
- Olive, T. et Piolat, A. (2005). Le rôle de la mémoire de travail dans la production de textes, *Psychologie Française*, 50, 373-390.
- Ong, W. (1982). *Orality and Literacy: The Technologizing of the Word*, Routledge.

- Paradis, J. et Navarro, S. (2003). Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: what is the role of input?, *Journal of Child Language*, 30, 371-393.
- Payne, D. (1987). Information structuring in Papago narrative discourse, *Language*, 63, 783-804.
- Pellegrini, A., Galda, L. et Rubin, D. (1984). Context in text: The development of oral and written language in two genres, *Child Development*, 55, 1549-1555.
- Péry-Woodley, M.P. (1993). *Les écrits dans l'apprentissage*, Vanves: Hachette.
- Peterson, C. et McCabe, A. (1982). *Development Psycholinguistics: three ways of looking at a child's narrative*, New York: Plenum.
- Piaget, J. (1923). *Le langage et la pensée chez l'enfant*, Neuchâtel: Delachaux et Niestlé.
- Piolat, A. (2004). Approche cognitive de l'activité rédactionnelle et de son acquisition. Le rôle de la mémoire de travail, *LINX (Linguistique Institut Nanterre Paris X)*, 51, 55-74.
- Piolat, A. et Fruttero, L. (1999). *Effects of memory span on cognitive load and composing strategies related to writing processes*, Paper presented at the Proceedings of the Writing Conference, MSHS, Poitiers.
- Piolat, A., Isnard, N. et Della-Valle, V. (1993). Traitement de texte et stratégies rédactionnelles, *Le Travail Humain*, 56, 79-99.
- Piolat, A., Olive, T. et Kellogg, R. (2005). Cognitive effort of note taking, *Applied Cognitive Psychology*, 19, 291-312.
- Piolat, A. et Pélissier, A. (Eds.) (1998). *La rédaction des textes. Approche cognitive*, Lausanne: Delachaux et Niestlé.
- Piolat, A. et Roussey, J.Y. (1992). Rédaction de textes. Eléments de psychologie cognitive, *Langages*, 106, 106-125.
- Piolat, A., Roussey, J., Olive, T. et Amada, M. (2004). Processing time and cognitive effort in revision: effects of error type and of working memory capacity, In Allal, L., Chanquoy, L., Largy, P. (eds.), *Revision. Cognitive and Instructional Processes*, 21-38, Dordrecht: Kluwer Academic Publishers.
- Prince, E. (1981). Toward a taxonomy of given-new information, In Cole, P. (eds.), *Radical Pragmatics*, 223-56, NY: Academic Press.
- Prince, E. (1992). The ZPG letter: subjects, definiteness, and information-status, In Thompson, S., Mann, W. (eds.), *Discourse description: diverse analyses of a fund raising text*, 295-325, Philadelphia/Amsterdam: John Benjamins.
- Ragnarsdottir, H. et Strömquist, S. (2005). The development of generic maður/man for the construction of discourse stance in Icelandic and Swedish, *Journal of Pragmatics*, 37, 143-155.
- Ravid, D. (2000). NP complexity in the development of text writing, *Developing Literacy Across Genres, Modalities and Languages*, 3, 163-170.
- Ravid, D. (2002). Paper about NP analysis: conceptualization and methodological issues, Tel Aviv University.
- Ravid, D. (2005). Emergence of linguistic complexity in later language development: evidence from expository text construction, In Ravid, D., Shyldkrot, H.B. (eds.), *Perspectives on language and language development. Essays in Honor of Ruth A. Berman*, 337-355, Dordrecht: Kluwer Academic Publishers.

- Ravid, D. (2006). Semantic development in textual contexts during the school years: Noun Scale analyses, *Journal of Child Language*, 33, 791–821.
- Ravid, D. et Berman, R. (1999). A clause-level analysis of 4th grade texts, *Developing Literacy Across Genres, Modalities, And Languages*, 1, 208-219.
- Ravid, D. et Berman, R. (sous press). Developing noun phrase complexity across adolescence: A text-embedded analysis, *First Language*.
- Ravid, D. et Cahana-Amitay, D. (2005). Verbal and nominal expressions in narrating conflict situations in Hebrew, *Journal of Pragmatics*, 37, 157-183.
- Ravid, D. et Katzenberger, I. (1999). Global text evaluation, *Developing Literacy Across Genres, Modalities, And Languages*, 1, 220-223.
- Ravid, R. et Shlesinger, Y. (1995). Factors in the selection of compound-types in spoken and written hebrew, *language Sciences*, 17, 147-179.
- Ravid, D. et Shyldkrot, H.B. (Eds.) (2004). *Perspectives on language and language development. Essays in Honor of Ruth A. Berman*, Dordrecht: Kluwer Academic Publishers.
- Ravid, D. et Tolchinsky, L. (2000). Developing linguistic literacy: A comprehensive model, *Developing Literacy Across Genres, Modalities and Languages*, 3, 1-31.
- Ravid, D. et Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model, *Journal of Child Language*, 29, 417-447.
- Ravid, D., Van Hell, J., Rosado, E. et Zamora, A. (2002). Subject NP patterning in the development of text production, *Written Language and Literacy*, 5, 69-94.
- Ravid, D., Shyldkrot, H.B. et Berman, R. (2005). *Perspectives on language and language development : essays in honor of Ruth A. Berman*, Dordrecht ; Boston: Kluwer Academic.
- Reichler-Béguelin, M.J. (1989). Anaphores, connecteurs, et processus inférentiels, In Rubattel, C. (eds.), *Modèles du discours. Recherches actuelles en Suisse romande*, 303-336, Berne: Peter Lang.
- Reichler-Béguelin, M.J. (1993). *Anaphores associatives non lexicales: incomplétude macrosyntaxique? Complétude et incomplétude dans les langues romanes et slaves*, Paper presented at the Actes du 6e Colloque international de Linguistique romane et slave (Septembre/Octobre 1991), Cracovie.
- Reilly, J. et Bernicot, J. (2003). Que nous apprennent les développements atypiques sur l'acquisition du langage ?, *Enfance*, 55, 205-207.
- Reilly, J., Losh, M., Bellugi, U. et Wulfeckd, B. (2004). Brain and Language 88 (2004) 163–166 "Frog, where are you?" Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome, *Brain and Language*, 88, 229–247.
- Reilly, J., Weckerly, J. et Bates, E. (2003). Neuroplasticité et développement : morphosyntaxe chez des enfants atteints de lésions focales précoces, *Enfance*, 55, 209-222.
- Reilly, J. et Wulfeck, B. (2004). Issues in plasticity and development: Language in atypical children, *Brain and Language*, 88, 163–166.
- Reilly, J., Zamora, A. et McGivern, R. (2005). Acquiring perspective in English: the development of stance, *Journal of Pragmatics*, 37, 185-208.

- Reinhart, T. (1980). Conditions for text coherence, *Poetics Today*, 1, 161-180.
- Rey, A. (1972). Usages, jugements et prescriptions linguistiques, *Langue française*, 16, 4-28.
- Rey-Debove, J. (2001). De on à je vers le nom propre : des pronoms personnels en français, In Bogaards, P. (eds.), *Quitte ou double sens*, 279-304, Amsterdam / New York: Rodopi.
- Riegel, M., Pellat, J.C. et Rioul, R. (1994/2002). *Grammaire méthodique du français*, Paris: Presses Universitaires de France.
- Salas, N., Tolchinsky, L., Mazur Palandre, A., Jisa, H. et Chenu, F. (2008). The effect of age, genre, and modality on Preferred Argument Structure in French and Spanish, *4th International Conference on Speech, Writing and Context (ICSWC)* Querétaro, Mexico.
- Salles, M. (2006). Cohésion-cohérence: accords et désaccords, *Corela (revue électronique)*, Numéros Spéciaux Organisation des Textes et Cohérence des Discours.
- Sampson, O. (1956). A study of speech development in children 18-30 months, *British Journal of Educational Psychology*, XXVI, 194-202.
- Sanford, A. et Garrod, S. (1981a). *Understanding written language*, New York: Wiley.
- Sanford, A. et Garrod, S. (1981b). Vers la construction d'un modèle psychologique de la compréhension du langage écrit, *Bulletin de psychologie*, 35, 643-648.
- Schilperoord, J. (1996). *It's about time. Temporal aspects of cognitive processes in text production*, Amsterdam: Rodopi.
- Schneuwly, B. (1988). *Le langage écrit chez l'enfant*, Neuchâtel: Delachaux et Niestlé.
- Schneuwly, B. et Bronckart, J.P. (1986). Connexion et cohésion dans quatre types de textes d'enfants, *Cahier de linguistique française*, 7, 279-294.
- Schneuwly, B., Rosat, M.C. et Dolz, J. (1989). Les organisateurs textuels dans 4 types de textes écrits. Étude chez des élèves de 10, 12 et 14 ans, *Langue Française*, 81, 40-58.
- Schuetze-Coburn, S. (1987). *Topic management and the lexicon: a discourse profile of three-argument verbs in German*, Master thesis, Los Angeles.
- Schober, M. et Clark, H. (1989). Understanding by addressees and overhearers, *Cognitive Psychology*, 21, 211-232.
- Seigneuric, A., De Guibert, M., Megherbi, H., Potier, D. et Picard, A. (2008). La MTVE : une épreuve de mémoire de travail verbale chez l'enfant adaptée du Listening Span Task de Siegel et Ryan (1989), *L'Année Psychologique*, 108, 273-308.
- Seigneuric, A. et Ehrlich, M.F. (2005). Contribution of working memory capacity to children's reading comprehension : a longitudinal investigation, *Reading and writing*, 18, 617-656.
- Seigneuric, A., Ehrlich, M.F., Oakill, J. et Yuill, N. (2000). Working memory resources and children's reading comprehension, *Reading and writing*, 13, 81-103.
- Seigneuric, A. et Megherbi, H. (2008). Capacité de mémoire de travail et traitement du pronom chez l'enfant, *Psychologie française*, 53, 281-306.
- Serratrice, L. (2005). The role of discourse pragmatics in the acquisition of subjects in Italian, *Applied Psycholinguistics*, 26, 437-462.

- Serratrice, L., Sorace, A. et Paoli, S. (2004). Crosslinguistic influence at the syntax-pragmatics interface: subjects and objects in Italian-English bilingual and monolingual acquisition *Bilingualism, Language and Cognition*, 7, 183-206.
- Siegel, L. et Ryan, E. (1989). The development of working memory in normally achieving and subtypes of learning disabled children, *Child Development*, 60, 973-980.
- Skarabela, B. (2006). *The role of social cognition in early syntax: The case of joint attention in argument realization*, Boston University, Doctoral dissertation.
- Skarabela, B. (2007). Signs of early social cognition in children's syntax: The case of joint attention in argument realization in child Inuktitut, *Lingua*, 117, 1837-1857.
- Skarabela, B., Allen, S. (2002). *Joint attention and argument realization in child Inuktitut*, Paper presented at the 26th Boston University Conference on Language Development, Boston.
- Smith, W. (1987). *Preferred Argument Structure in spoken Hebrew discourse*, Master thesis, Los Angeles.
- Smith, W. (1996). Spoken narrative and preferred clause structure: evidence from modern Hebrew discourse, *Studies in Language*, 20, 163-89.
- Stewart, A. (1984). *A glance at Preferred Argument Structure in Conchucos Quechua*, Master thesis, Los Angeles.
- Strömquist, S. (2000). A note on pauses in speech and in writing, *Developing Literacy Across Genres, Modalities and Languages*, 3, 211-225.
- Strömquist, S., Holmqvist, K., Johansson, V., Karlsson, H. et Wenglin, A. (2006). What key-logging can reveal about writing, In Sullivan, K., Lindgren, E. (eds.), *Computer Key-stroke Logging and Writing: Methods and Applications*, 45-73, Amsterdam: Elsevier.
- Thal, D., Reilly, J., Seibert, L., Jeffries, R. et Fenson, J. (2004). Language development in children at risk for language impairment: Cross-population comparisons, *Brain and Language*, 88, 163-166.
- Tijus, C. (2001). *Introduction à la psychologie cognitive, comment la pensée et l'intelligence se construisent en interaction avec le monde*, Paris: Nathan Université.
- Tolchinsky, L. (2006). The emergence of writing, In MacArthur, C., Graham, S., Fitzgerald, J. (eds.), *Handbook of Writing Research*, 83-95, New York: The Guilford Press.
- Tolchinsky, L., Johansson, V. et Zamora, A. (2002). Text openings and closings: Textual autonomy and differentiation, *Written Language and Literacy*, 5, 219-254.
- Tolchinsky, L., Perera, J., Argerich, N. et Aparici, M. (1999). Rhetorical moves and noun phrase complexity in written expository texts, *Developing Literacy Across Genres, Modalities and Languages*, 1, 257-278.
- Tolchinsky, L. et Rosado, E. (2005). The effect of literacy, text type, and modality on the use of grammatical means for agency alternation in Spanish, *Journal of Pragmatics*, 37, 209-237.
- Torrance, M. et Galbraith, D. (2006). The processing demands of writing, In MacArthur, C., Graham, S., Fitzgerald, J. (eds.), *Handbook of Writing Research*, 67-82, New York: The Guilford Press.
- Traxler, M. et Gernsbacher, M. (1995). Improving coherence in written communication, In Gernsbacher, M., Givón, T. (eds.), *Coherence in spontaneous text*, 215-238, Philadelphia: John Benjamins.
- Tuaillon, G. (1983). *Matériaux pour l'étude des régionalismes du français*, Paris: Klincksieck.
- Van Dijk, T. et Kintsch, W. (1983). *Strategies of discourse comprehension*, New York: Academic Press.

- Van Galen, G. (1990). Phonological and motoric demands in handwriting: Evidence for discrete transmission of information, *Acta Psychologica*, 74, 259-75.
- Van Hell, J., Verhoeven, L., Tak, M. et Van Oosterhout, M. (2005). To take a stance: a developmental study of the use of pronouns and passives in spoken and written narrative and expository texts in Dutch, *Journal of Pragmatics*, 37, 239-273.
- Verdelhan-Bourgade, M. (2002). *Le Français de Scolarisation pour une didactique réaliste*, Paris : Presses Universitaires de France.
- Viguié, A. et Jisa, H. (1999). Schooling in France, *Developing Literacy Across Genres, Modalities, And Languages*, 1, 291-298.
- Viguié-Simon, A. (2001). *Développement de la connectivité selon le type textuel et le mode de production au cours de l'adolescence*, Thèse de Doctorat, Université Lumière, Lyon.
- Weber, E. (2003). Nominal information flow in the talk of two boys with autism, In Du Bois, J., Kumpf, L., Ashby, W. (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 353-384, Amsterdam: John Benjamins.
- Wigglesworth, G. (1991, 1991). *Reference tracking strategies of young children in extended narrative discourse*, Paper presented at the Working papers in Linguistics, La Trobe University.
- Wittek, A. et Tomassello, M. (2005). Young children's sensitivity to listener knowledge and perceptual context in choosing referring expressions, *Applied Psycholinguistics*, 26, 541-558.
- Wulfeck, B., Bates, E., Krupa-Kwiatkowski, M. et Saltzman, D. (2004). Grammaticality sensitivity in children with early focal brain injury and children with specific language impairment, *Brain and Language*, 88, 215-228.
- Zamuner, V. (1986). Children's coherence in discourse: A review, In Petöfi, J. (eds.), *Text connectedness from psychological point of view*, 24-63, Hamburg: Buske.
- Zamora, A. et Kriz, S. (2000). Stance in written and spoken texts: a developmental study, *Developing Literacy Across Genres, Modalities, And Languages*, 3, 31-42.
- Zellner, B. (1994). Pauses and the temporal structure of speech, In Keller, E. (eds.), *Fundamentals of speech synthesis and speech recognition*, 41-62, Chichester: John Wiley.