

Des mots pour se dire : Développement de la fonction narrative chez l'enfant

Aurore Boulard¹, Céline Poudat², Jean-Marie Gauthier³

¹ Université de Liège – Aurore.Boulard@ulg.ac.be

² UMR 7187 LDI, Université de Paris 13 – celine@poudat.fr

³ Université de Liège – jmgauthier@ulg.ac.be

Abstract

The present study is based on a corpus made of spontaneous oral speech: 136 children aged from 3 to 13 have been recorded in three different environments that are favorable to the production of self-narratives. Using linguistic annotations and tried and tested text statistics methods, we have been able to show different aspects of the development of narratives and self-narratives in children speeches, including the impact of age and environment on the construction of narratives.

The results we obtained are significant, and we hope we succeeded in demonstrating how relevant text statistics methods are when applied to clinical psychology.

Keywords : narrative function, text statistics, children, text typologies

Résumé

Grâce à un corpus constitué de transcriptions de discours oral spontané de 136 enfants de 3 à 13 ans enregistrés dans trois environnements propices au développement de la narration de soi, nous avons pu mettre en évidence certains aspects du développement de la fonction narrative chez l'enfant. Comment l'enfant commence-t-il à parler de lui? Son discours est-il sensible au contexte de l'énonciation et à partir de quel âge?

Enrichi d'une annotation morphosyntaxique, le corpus a été soumis à un ensemble de méthodes de statistiques multidimensionnelles qui nous ont permis de mettre au jour des attractions et des oppositions potentiellement caractéristiques que la présente communication se propose de décrire et de spécifier, en résonance avec les théories cliniques sur la question.

Au vu des résultats particulièrement significatifs obtenus, nous espérons avoir démontré l'intérêt de l'analyse de données textuelles (ADT) pour la psychologie clinique. Parce qu'elle objective le discours, l'analyse statistique permet au clinicien une interprétation au plus près du texte et ouvre le champ à de nouvelles pistes de recherche ainsi qu'à la vérification de nombreuses hypothèses cliniques tant en psychologie développementale qu'en psychopathologie en ce qui concerne l'individuation et l'identité.

Mots-clés : fonction narrative, analyse de données textuelles, enfants, typologies textuelles

1. Introduction

Margareth Somers (1994) écrit que « c'est à travers la narration et le narrativité que nous constituons notre identité sociale. Chacun de nous doit devenir ce qu'il est en se situant dans

des narrations sociales rarement personnelles ». C'est ce que nous pouvons observer dès le plus jeune âge chez les enfants qui sont tout d'abord baignés dans les narrations parentales avant de rencontrer à l'école à la fois d'autres enfants du même âge qu'eux et des adultes avec qui ils partageront d'autres histoires. À travers tous ces récits, l'enfant va commencer à développer dès le plus jeune âge ce sentiment d'identité, de continuité de Soi à Soi, qui lui permettra à l'adolescence de raconter son histoire avec un *moi* enfin unifié par toutes ces expériences discursives. Avant cette période l'enfant n'est pas capable d'ordonner ses expériences en une histoire cohérente. Il raconte des bribes d'histoires ou des événements non reliés dans le temps et où la causalité est absente (Habermas & Bluck, 2000). Jusqu'à 10-12 ans, l'enfant raconte des événements en commençant à intégrer les normes de sa culture, il raconte les histoires qui lui arrivent dans sa famille et d'éventuels déplacements. Ce faisant, il acquiert petit à petit ce qu'Habermas et Bluck vont appeler une cohérence narrative (McAdams, 2001).

Puisque la narration de soi n'est pas directement observable avant l'adolescence, nous avons voulu en observer les prémisses à travers l'analyse linguistique empirique d'un corpus de transcriptions de discours oral spontané de 136 enfants de 3 à 13 ans enregistrés dans trois environnements propices au développement de la narration (avec le(s) parent(s), un psychologue et en groupe). L'expérience que nous avons menée est inédite, dans le sens où les travaux de psycholinguistique travaillent sur des ensembles de données beaucoup plus restreints, tandis que les psychologues travaillent peu sur corpus, le traitement des questions ouvertes étant encore très problématique dans ces communautés scientifiques. Le présent article mobilise ainsi des méthodologies de recherche très actuelles en analyse de données textuelles et des stratégies d'annotation linguistique et d'exploration de corpus qui sont peu appliquées en lien avec des questions de recherche en psychologie du développement.

Le corpus sera exploré à travers deux lunettes d'observation : son lexique et sa morphosyntaxe. Nous nous inscrivons ici dans la continuité des travaux initiés par Biber (1988), et poursuivis par Karlgren (1994), Malrieu et Rastier (2001) et Poudat (2006) autour de la question des genres textuels. Nous avons ainsi pu montrer (voir Poudat, 2006) l'intérêt de la morphosyntaxe pour faire émerger des modes d'énonciation et de narration pour caractériser les genres et les styles ; nous entendons donc appliquer des développements méthodologiques comparables à ce corpus d'enfants, afin de mettre au jour les oppositions discursives et lexicales les plus manifestes d'un âge et d'un environnement à l'autre, en lien avec l'observation du développement de la fonction narrative chez l'enfant.

Nous tenterons donc de répondre aux questions qui suivent : Comment se développe la narration chez l'enfant ? Dans quelle mesure son discours varie-t-il avec son environnement ? Quels sont les effets de l'âge sur le développement de cette fonction narrative ? Observe-t-on des paliers de développement ?

Cette étude s'inscrit dans une recherche plus large sur le développement de la fonction narrative autobiographique chez l'enfant et notamment sur l'utilisation des pronoms personnels chez l'enfant comme marqueurs discursifs de l'identité (voir Boulard, Gauthier, 2012 ; Poudat, Gauthier, Boulard, 2012).

Enrichi d'une annotation morphosyntaxique, le corpus a ainsi été soumis à un ensemble de méthodes de statistiques multidimensionnelles qui nous ont permis de mettre au jour des

attractions et des oppositions potentiellement caractéristiques que la présente communication se propose de décrire et de spécifier, en résonance avec les théories cliniques sur la question.

2. Corpus

Le corpus qui fonde cette étude est constitué d'un ensemble de 408 transcriptions du discours oral spontané de 136 enfants et préadolescents âgés de 3 à 13 ans (de la première année maternelle à la deuxième année d'humanité) issus de 12 écoles de la Communauté Française de Belgique. L'échantillon est composé de 62 garçons et 74 filles. Afin que cet échantillon puisse être représentatif de la population générale, nous avons effectués ces enregistrements dans les deux types d'enseignement fondamental (public et privé) et dans des écoles situées pour huit d'entre elles en ville tandis que les quatre autres étaient situées dans un environnement plus rural.

Les enfants ont été enregistrés dans trois conditions de socialisation différentes : (i) jeu libre individuel avec un(e) psychologue, appelée condition "psy"; (ii) jeu libre avec les parents, à la maison, appelée condition "parents"; et (iii) jeu libre collectif, appelée condition "groupe". Nous avons donc trois enregistrements d'une durée de 50 minutes pour chaque enfant.

Lors de la situation de jeu libre avec le psychologue, la consigne donnée à l'enfant était de choisir une des activités proposées par le chercheur : soit dessiner, soit jouer avec de la pâte à modeler. Ces activités ont été pré testées et choisies car elles permettaient de ne pas orienter le discours de l'enfant et de le laisser parler librement de sujets de son choix. Le jeu libre collectif avec des enfants du même âge consistait en la construction d'objets en pâte à modeler. Cette activité ludique permettait les échanges entre enfants et ne demandait pas une concentration élevée, rendant ainsi possible les interactions de toutes sortes autour de sujets divers. Pour la condition familiale, nous avons enregistré le discours de l'enfant à son domicile avec les parents lors d'un moment privilégié passé en famille. Pour les plus petits, le bain ou le repas du soir ont été choisis, pour les plus grands (à partir de 6 ans), le repas du soir a été choisi dans la plupart des familles. Ces trois situations en environnement soutenant (parents, psy, groupe) permettent de poser des hypothèses intéressantes sur le développement de la narration de soi et comme l'ont montré Boulard et Gauthier (2011), donnent une image dynamique du discours de l'enfant.

Le corpus comprend donc 408 transcriptions dans lesquelles n'ont été extraites que les productions de l'enfant ; les tours de parole du praticien, des parents et des autres enfants selon la situation considérée ne sont pas conservés. Au total, il comptabilise 553 006 mots et ponctèmes.

L'originalité de ce corpus réside à la fois dans sa taille importante par rapport aux autres études du même type et dans la méthodologie utilisée. A notre connaissance, aucune étude ne compare trois situations discursives différentes chez un même sujet. Colletta (1998) met d'ailleurs en évidence le peu d'attention au développement narratif chez l'enfant scolarisé et préconise justement cette comparaison.

3. Méthodologie

3.1. Observables linguistiques

Deux niveaux de l'analyse linguistique seront considérés : le lexique, à travers l'examen du vocabulaire du corpus, et la morphosyntaxe, appréhendée au moyen de l'analyseur (morpho) syntaxique Cordial (Société Synapse Développement), qui effectue une analyse syntaxique (syntagmatique) des textes et affecte à chaque unité segmentée une catégorie morphosyntaxique et un lemme. Cordial a également été retenu du fait de ses performances très honorables sur le français. Les catégories ont été rigoureusement sélectionnées suivant les spécificités des textes et les caractéristiques du corpus. Ainsi, les descripteurs relatifs aux « dialogues » dans les textes (séquences précédées d'un tiret qui abondent dans le discours littéraire) ont été écartés, de même que certains ponctèmes qui ne sauraient être relevés dans les transcriptions (*e.g.* crochets, parenthèses ou points-virgules). En outre, les catégories sémantiques de l'analyseur les plus fantaisistes (*e.g.* la catégorie *humanoïde*) n'ont pas été retenues. Au final, un ensemble de 136 descripteurs morphosyntaxiques et syntaxiques a été conservé, qui nous permettra de typer les regroupements textuels obtenus, et de mettre au jour des systèmes d'opposition narratifs et énonciatifs.

3.2. Méthodes statistiques

Afin de mettre en évidence la structure du corpus, nous avons mobilisé deux méthodes factorielles : nous avons appliqué la méthode classique de l'Analyse Factorielle des Correspondances (AFC) au vocabulaire du corpus afin d'en dégager sa structure lexicale, et nous avons mené une Analyse en Composantes Principales (ACP) combinée à une Classification Ascendante Hiérarchique¹ (CAH) pour mettre au jour les grandes oppositions morphosyntaxiques à l'œuvre dans les textes, suivant la méthodologie décrite dans Poudat (2006).

Ces deux méthodes ont enfin été combinées en projetant le vocabulaire du corpus a posteriori sur les axes de l'ACP : l'ACP a été complétée d'une classification et les classes obtenues ont été décrites a posteriori par le lexique, convoqué à titre illustratif².

Toutes les analyses menées ont été réalisées au moyen du logiciel DtmVic³, développé par Ludovic Lebart.

1 Le choix d'une CAH, et non pas d'une CDH (Classification Descendante Hiérarchique) est d'abord justifié par la pratique, parce que la combinaison AFC (ou ACP)-CAH est la plus répandue en ADT et dans la littérature en général ; seul *Alceste* (Société Image) recourt à une CDH en scindant suivant les premiers axes de l'AFC. Globalement, CAH et CDH obtiennent des résultats comparables, à quelques différences près : si la CDH permet d'obtenir des résultats peut-être plus directement interprétables pour ce qui concerne le haut de l'arbre de classification, la CAH est plus intéressante pour dégager les regroupements de forte densité situés au bas de l'arbre.

2 Nous remercions particulièrement Ludovic Lebart de son aide pour mettre en œuvre cette dernière méthode.

3 DtmVic (Data and Text Mining : Visualisation, Inference, Classification), logiciel de statistique exploratoire multidimensionnelle dédié au traitement des données numériques et textuelles. Version 5.2. http://www.dtmvic.com/05_SoftwareF.html

4. Résultats

4.1. Exploration de la structure morphosyntaxique et lexicale du corpus

La figure 1 nous permet d'observer les deux premiers plans factoriels de l'ACP ; la première valeur propre, comprise entre 1 et 136 (soit le nombre total des descripteurs observés) est ici égale à 22,37. De manière générale, plus la valeur propre est importante, plus elle résume de variables et plus le facteur est intéressant en termes de synthèse. À chaque facteur (ou valeur propre) est associé un pourcentage d'inertie, ou de variance expliquée, exprimé relativement à l'inertie totale du nuage observé, qui permet d'apprécier l'importance relative du facteur. Le premier facteur correspond à un pourcentage d'inertie de 11,36%, ce qui est significatif si l'on considère les 136 descripteurs pris en compte : il rendrait ainsi compte de 15,45 variables, tandis que les quatre premières valeurs propres en rendraient compte de 36, soit 35% des variables prises en compte. Soulignons également que le nuage des individus, que nous ne rendons pas ici, était particulièrement bien dispersé, et qu'aucun texte outlier⁴ n'a été détecté.

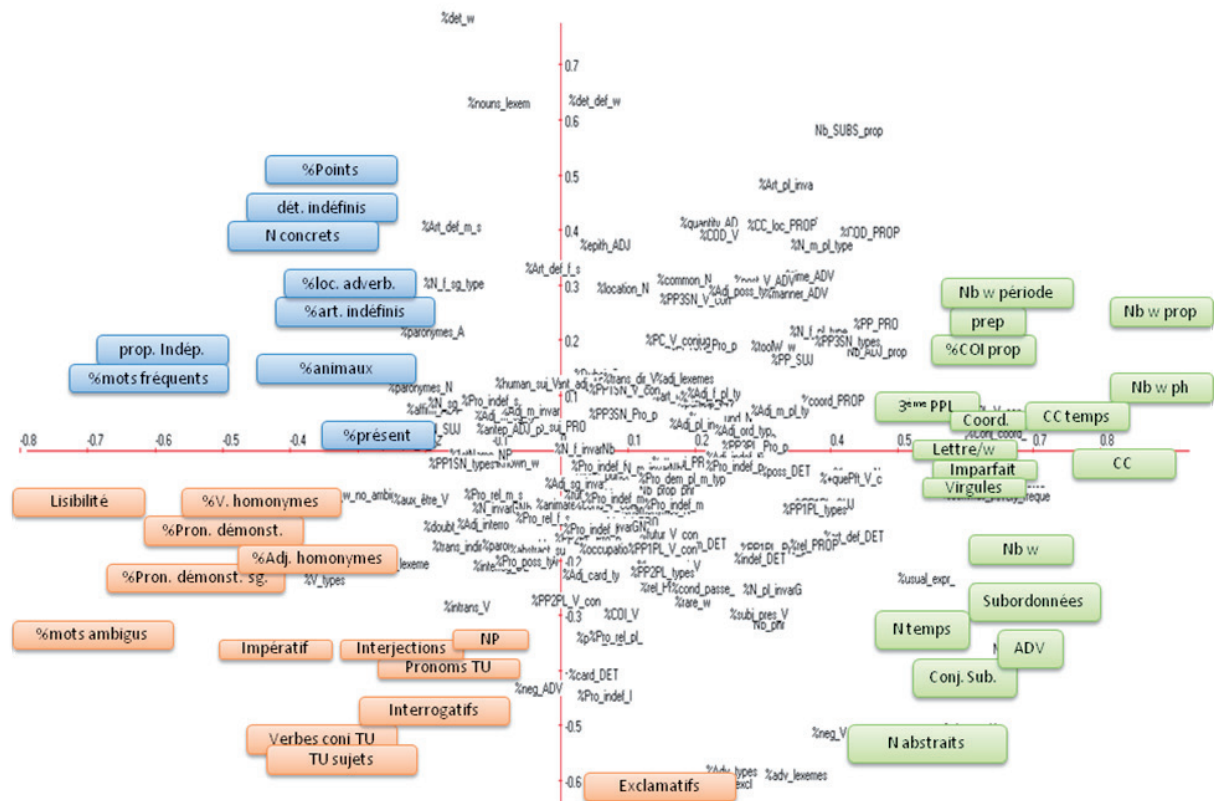


Figure 1 : Carte du premier plan factoriel de l'ACP – Morphosyntaxe – 408 textes / 136 descripteurs morphosyntaxiques

Le premier facteur oppose d'abord deux styles narratifs marqués : il est positivement corrélé aux indices d'une élaboration narrative plus développée (en vert), comme le montre

4 Un outlier est une observation (ici un texte) qui se démarque fortement des autres observations constituant l'échantillon étudié. La présence d'un outlier dans un échantillon peut avoir une incidence non négligeable sur les résultats des analyses effectuées.

l'intercorrélation des phrases et des propositions plus longues, le recours à des coordinations (versant positif de l'axe 2) et des subordinations (versant négatif de l'axe 2). L'enfant mobilise conjointement des mots plus longs et plus « abstraits », ce qui atteste d'une maîtrise plus aboutie de la langue française, tandis que la narration semble se structurer par l'emploi d'adverbes et de compléments circonstanciels (CC), et plus spécifiquement de temporels (CC de temps, et noms « temporels »). Ce style plus développé qui contextualise et cadre les événements du récit est naturellement corrélé à l'usage de l'imparfait, temps imperfectif qui saisit le procès passé en cours de déroulement et qui requiert une certaine maîtrise des microsystèmes verbaux et des plans du récit.

Quand j'ai été à Plopsa land à Noël je crois l'année passée. Aussi je crois on avait on avait été dans tu vois un train fantôme. En fait c'était une rivière on va dire. On va appeler ça comme ça et on était dans un bateau et il y avait. Tu avais plein de squelettes et tout ça et alors à un moment on monte et on monte fort et je dis à ma maman fais gaffe on va se faire tremper et maman elle dit à papa protège l'appareil photo et alors à un moment quand on redescend on voit on voyait pas encore l'eau et on redescend et on avait été tout mouillé. (extrait du texte 285 – Martin, 9 ans, situation PSY)

À l'opposé, le versant négatif du premier axe factoriel laisse entrevoir un style discursif moins élaboré et plus lapidaire (en bleu), caractérisé par la présence de phrases courtes (points, propositions indépendantes) et de déterminants indéfinis, indiquant un degré de cohésion textuelle limité avec peu de chaînes de références. Associé à un lexique représentant un univers plus concret et enfantin, notamment peuplé d'*animaux*, ce style moins construit est également observable sur le deuxième axe, qui extrait 7,11% du nuage de points. Il s'oppose d'une part au style plus élaboré précédemment décrit, et plus spécifiquement aux noms abstraits et à l'emploi de subordonnées, indices d'une compétence syntaxique plus développée ; et d'autre part à un groupement de traits morphosyntaxiques qui renvoie nettement au style conversationnel familier (en orange), autour des déictiques *tu* et des pronoms démonstratifs (*ceci, cela, ça*). Corrélés à la présence d'interjections et de ponctuations interrogatives, auxquelles on peut rajouter les exclamatifs (également positionnés sur l'extrémité négative de l'axe 2), ce troisième style témoigne d'un degré d'interactivité plus élevé, et surtout plus animé. Si le lexique employé paraît plus ambigu, peut-être du fait des délais de sélection plus courts dont dispose le locuteur lors d'une conversation animée, il semble également plus « lisible » d'après l'analyseur sémantique de Cordial⁵.

La deuxième expérience menée sur le vocabulaire du corpus, en nous restreignant aux mots les plus fréquents (avec un seuil de 50 occurrences), soit 519 149 occurrences conservées, confirme et précise les régularités observées sur le plan morphosyntaxique, ce qui souligne encore la stabilité de la structure mise au jour. Comme le montre la figure 2, l'axe 1 oppose encore une fois deux niveaux d'élaboration⁶ : le versant positif de l'axe fédère un ensemble de mots formes caractéristiques de l'univers enfantin, de son bestiaire (du *nounours* au *crocodile*) à ses couleurs et sa gastronomie (*saucisse, pomme, gâteau*), tandis que le versant négatif de l'axe met en scène des personnages quotidiens (la *meilleure amie*, la *mère*, le *professeur*) et des événements (la *dispute*, ils *disent*, *énervé*, etc.) qui montrent comment se met en place de la

5 Avec toutes les réserves que ce type de catégorie requiert.

6 L'orientation des axes est ici inversée, ce qui n'a aucune incidence puisqu'on sait bien que les facteurs ne s'observent qu'en termes d'oppositions.

fonction narrative qui donnera lieu, un peu plus tard, au récit de soi. De manière intéressante, on retrouve également ce pôle conversationnel détecté précédemment, sur le versant négatif de l'axe 2, avec la présence d'interjections, de prénoms et de verbes à l'impératif ou conjugués à la seconde personne du singulier (e.g. *attends, aimes, arrête*).

Hé, tu peux faire encore avec les chewing-gums là ? Tu peux faire avec le mauve. Avec le mauve alors. Du rose. Comme ça tu feras un chiquelet, oui ? Non, quand même pas. Hé ! Il veut le dessiner. Ah oui. Mais j'ai mal à la jambe. (Extrait du texte 325, Marie, 4 ans, situation GROUPE).

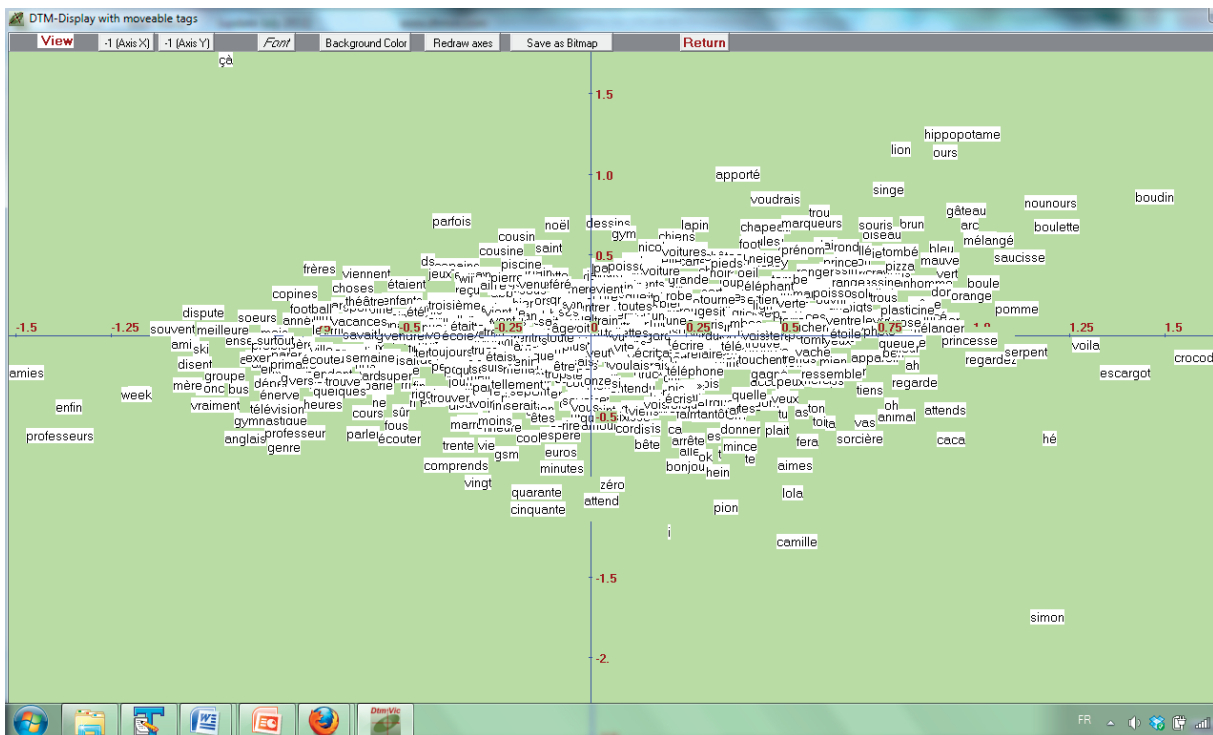


Figure 2 : Carte du premier plan factoriel de l'AFC – Lexique– 408 textes / seuil de 50 occ.

4.2. Âges et situations projection des variables catégorielles sur les axes

L'examen de la structure morphosyntaxique et lexicale du corpus nous a permis de mettre au jour trois styles discursifs, manifestant des degrés d'élaboration et des modes conversationnels distincts. Afin de préciser cette structure, nous avons projeté les variables catégorielles illustratives *âge* et *situation* sur le premier plan factoriel (Figure 3) en mobilisant la méthode du bootstrap⁷ (Lebart, 2004), afin de tracer des zones de confiance autour des 11 modalités de la variable *âge* (de 3 à 13 ans), et des trois modalités de la variable *situation* (*psy, parent, groupe*).

La projection de l'âge sur le premier plan factoriel s'avère d'abord particulièrement intéressante : si l'enfant de trois ans se distingue spécifiquement (zone plus petite et non recouverte) à l'extrémité négative de l'axe 1 qui renvoie au style discursif moins développé que nous avons

⁷ La méthode de bootstrap, ou de rééchantillonnage, est un test extrêmement sévère pour évaluer la stabilité d'une structure. Elle consiste à faire de l'inférence statistique sur de « nouveaux » échantillons tirés à partir d'un échantillon initial.

mis au jour *supra*, les âges représentés se distribuent presque régulièrement jusqu'à l'âge de 9 ans des deux côtés de l'axe 1, qui figure bien une échelle d'élaboration narrative. Sur la base des données dont nous disposons, il semblerait que le développement de la fonction narrative chez l'enfant connaisse une certaine stabilisation vers l'âge de 10 ans : les zones de confiance autour des préadolescents de 11 à 13 ans se recouvrent en effet largement, signalant une relative homogénéité des discours produits à ces âges, qui se caractérisent notamment par une maîtrise de la subordination syntaxique.

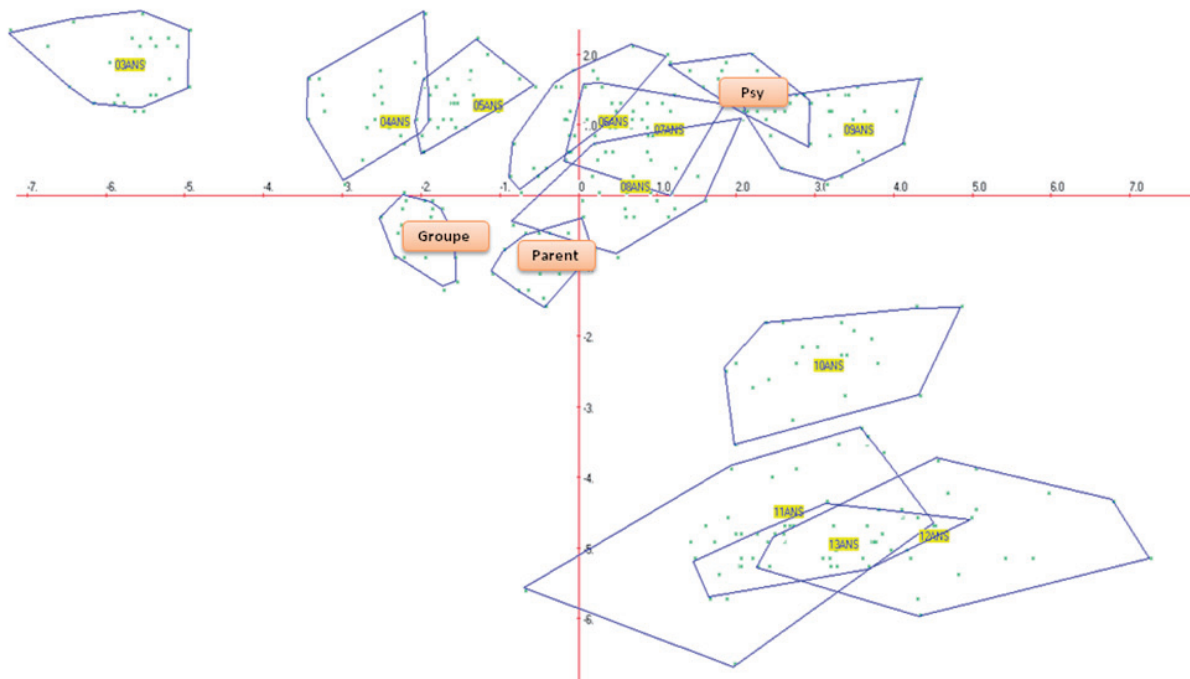


Figure 3 : Zones de confiance (enveloppes convexes) autour des 3 situations et des 11 âges observés sur le premier plan factoriel – variables catégorielles illustratives de l'ACP Morphosyntaxe, voir Fig. 1

Les situations de production des discours sont tout aussi significatives, d'autant que les enveloppes convexes tracées sont plus petites. Si les interactions en groupe et avec les parents tendent naturellement vers le pôle conversationnel mis au jour, on observe que les conversations avec les parents sont significativement plus animées (davantage de ponctèmes interrogatifs et exclamatifs) tandis que les interactions entre enfants semblent plus régressives sur le plan linguistique (vocabulaire simplifié, pronoms démonstratifs), dans la mesure où elles se situent sur le versant négatif de l'axe 1, qui distingue les enfants les plus jeunes (3 à 5 ans) des autres. La situation d'entretien avec le psychologue est en revanche plus propice à la mise en place d'une narration plus élaborée : on sort en effet de la conversation quotidienne, avec un degré d'interaction plus faible et une orientation de l'écoute vers l'enfant. En outre, la situation est inédite pour l'enfant qui se trouve confronté à une personne non familière, ce qui exige un contrôle linguistique plus important. La socialisation met en effet en jeu la question de l'identité et de la différence ; l'acquisition des compétences sociales n'est pas le résultat d'une passivité mais la capacité de mettre en jeu sa propre identité qui est bien cette continuité de Soi à Soi. La compétence narrative autorise la socialisation en quelque sorte, puisque la confrontation à l'autre développe un discours différent.

Ces résultats se trouvent confirmés par la projection du lexique a posteriori sur le premier plan factoriel de l'ACP, par le biais d'une classification ascendante hiérarchique. Les deux niveaux d'analyse morphosyntaxique et lexicale convergent et illustrent une fois encore cet ergodisme déjà décrit par Brunet (2011), puisque le style moins élaboré observé précédemment est bien caractérisé par le lexique de l'univers enfantin observable sur la figure 2 tandis que la stabilisation de la fonction narrative s'évaluerait notamment par la maîtrise des cadratifs temporels. Si l'on se concentre sur la situation « psy » qui s'est avérée propice à la construction d'une narration plus élaborée, on peut observer sur la figure 4 deux niveaux d'élaboration qui donnent lieu à deux classes, distinguant les enfants âgés de 6 ans (classe 9) et les enfants âgés de 10 à 13 ans (classe 12).

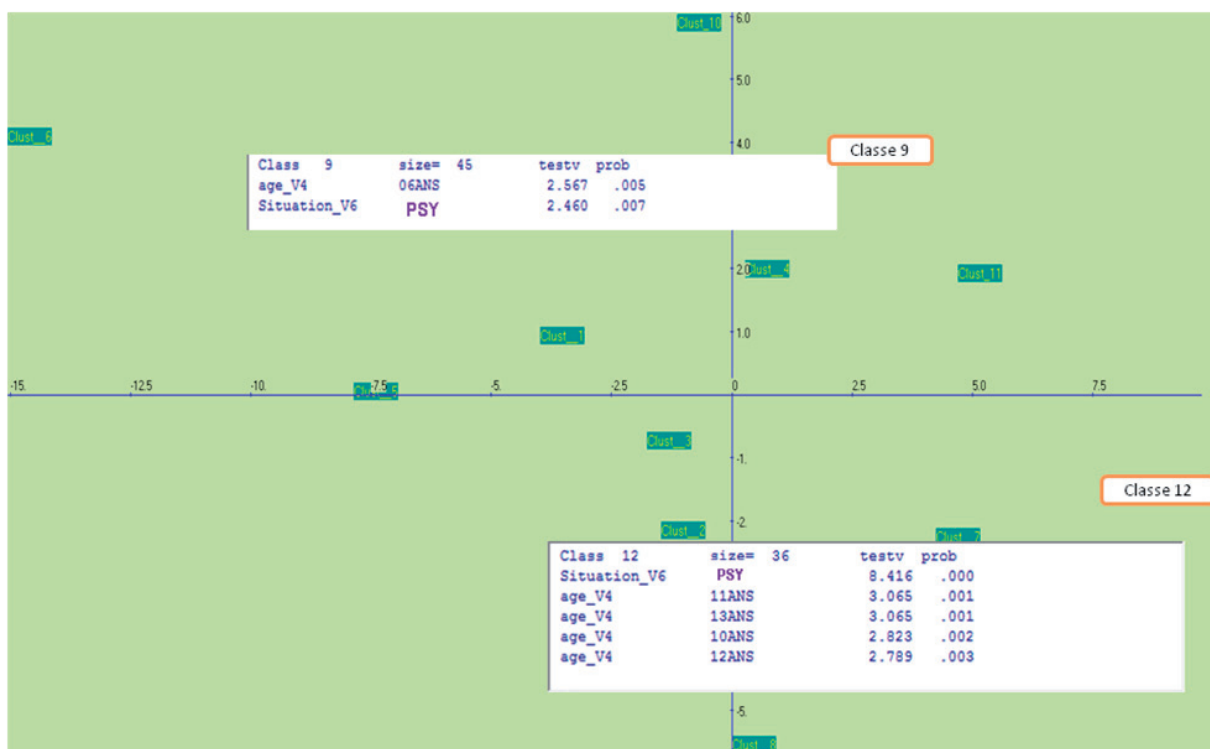


Figure 4 : Carte des deux premiers plans factoriels – AFC – Lexique– 408 textes / seuil de 50 occ.

Les productions de la classe 9 se caractérisent par un usage significatif de « non personnes » (singulier et pluriel), pour reprendre les termes de Benveniste, de propositions mobilisant des compléments d'objet direct et indirect et d'adverbes de quantité et de temps, comme *parfois*. Les adverbes s'avèrent être empiriquement des indices importants pour évaluer le développement langagier : si les enfants plus jeunes de la classe 9 positionnent plus volontiers l'adverbe après le verbe, les 10-13 ans en utilisent un plus grand nombre et leurs productions sont plus longues (nombre total de mots, phrases et propositions plus longues) ; c'est également à ces âges que l'on observe un emploi significatif de l'imparfait et des cadratifs temporels – et des compléments circonstanciels en général, de même qu'un recours aux noms plus abstraits et humains : la *mère* fait ainsi partie des mots les plus spécifiques de cette classe, tandis que le lexique des enfants plus jeunes reste polarisé autour du dessin demandé (*dessins, marqueurs, jeux...*), dont les préadolescents s'émancipent complètement.

5. Discussion

Nos résultats confirment d'abord de manière empirique que le développement du langage se manifeste bien par un développement rapide du lexique (Bideaud, Houdé & Pedinielli, 2003) et une augmentation de la longueur des énoncés qui reflètent à la fois l'accroissement de la complexité de l'énoncé et l'accroissement de son organisation (Gratiot-Aphandéry & Zazzo, 1976). Parallèlement à cet accroissement de la longueur des énoncés, nous observons l'arrivée de marqueurs temporels. « Le langage de la temporalité se développe donc, non seulement comme l'expression des conquêtes effectuées dans l'organisation des réactions dans le temps, mais encore comme un instrument de celle-ci, en permettant à l'enfant d'intérioriser les modèles de temporalité que lui offre la culture de sa société, au cours des dialogues qui l'informent sur la signification de ses actes socialisés » (Malrieu, 1973, p. 341).

La méthodologie que nous avons mise en œuvre s'est avérée particulièrement productive et réaffirme l'intérêt d'examiner le système morphosyntaxique endogène au corpus pour mettre au jour les dispositifs énonciatifs et les modes de narration qui coexistent dans les textes, ce qui nous laisse à penser qu'une systématisation de l'approche et une explicitation de ses règles d'interprétation doit être développée.

L'analyse de nos résultats nous montre ainsi que les enfants s'expriment de manière différente selon la condition. Le discours en présence des parents est empreint d'interrogations, d'interpellations tandis que la condition groupe met en évidence un discours moins « ponctué ». Nous pouvons y voir la différence mise en évidence par Vygotsky(1984) entre le langage égocentrique où l'enfant parle pour lui-même, n'a pas ou peu d'intérêt pour l'interlocuteur, n'essaye pas vraiment de communiquer, ne prend pas ou peu d'intérêt à l'interlocuteur, et le langage social où l'enfant entre en échange avec l'autre, en lui posant des questions ou en s'imposant. Si pour Piaget (1989) le langage égocentrique ne remplit pas de fonction utile, il est pour Vygotsky plus qu'un simple accompagnement de l'activité. Il devient un réel instrument de pensée en ce sens qu'il permet d'accéder à une temporalité par la planification de l'action par exemple : « je vais faire une chenille ». Pour Vygotsky, le langage égocentrique est très important du point de vue méthodologique car selon lui, il se rapproche du langage intérieur de l'enfant, qui est par essence inaccessible. Comme il a la même fonction que le langage intérieur, c'est en quelque sorte le révélateur du développement de la pensée de l'enfant. Le langage social avec les parents met en évidence une « présence fréquente des impératifs, des vocatifs, des moyens linguistiques et extralinguistiques utilisés pour attirer l'attention ». Il dénote la présence d'un langage affectif et en même temps dynamique, rattaché à l'action tel que l'a observée Slama-Cazacu (1966). « C'est de ce double substrat – nécessité d'exprimer directement certains états affectifs, mais aussi nécessité de communiquer aux autres les contenus psychiques personnels – que dérivent les particularités d'ordre lexical et grammatical des interpellations des enfants en bas âge. Les nécessités et les intérêts qui les déterminent, de même que la conscience que leur satisfaction dépend de la communication par le langage et de son efficacité, imposent à leurs expressions des formes d'ordre, de demande, de prière et sont la cause d'un pourcentage très élevé d'impératifs et de vocatifs ». Lorsque nous observons nos données, nous constatons que ces deux formes de discours peuvent coexister au même âge. L'enfant est dès le plus jeune âge capable d'adapter son discours lorsqu'il est présence d'un adulte inconnu.

De 2 ans et demi à 6 ans l'enfant va se constituer comme une personne autonome, de moins en moins dépendante de son environnement. C'est le stade du personnalisme (cf. Wallon, 1934) : à 2 ans l'enfant semble se confondre avec son entourage et reste prisonnier des situations dans l'appréciation de la connaissance des personnes. Lui-même n'est pas toujours très sûr de sa propre identité. De 3 à 5 ans, l'enfant reste profondément inséré dans son milieu familial. Ses rapports avec les siens, la place qu'il occupe parmi ses frères et sœurs, font partie de sa propre identité personnelle. Il ne sait pas se distinguer du sort qui lui échoit dans la constellation domestique. Avec le stade suivant, de 6 à 11 ans, le syncrétisme⁸ de la personne et celui de l'intelligence vont pouvoir se résoudre en faisant place aux différenciations nécessaires. À l'école, les rapports exigés par la camaraderie et par la discipline sont bien plus variables que dans la famille et doivent pouvoir se moduler suivant un entourage et les circonstances, eux-mêmes changeants. L'enfant apprendra à se connaître comme une personnalité polyvalente. En ajustant ses conduites à des circonstances particulières, loin de se dissiper sans fin, il aura de lui-même une connaissance plus précise et plus complète. L'enfant dans sa famille peut d'une part se confondre avec elle et d'autre part, la présence d'un Autre, inconnu de l'enfant amène celui-ci à une plus grande explication de lui-même mise en évidence dans notre corpus par la présence de noms abstraits, de marqueurs temporels ou de subordonnées. Nous pouvons également observer que lorsque l'enfant grandit, l'influence des conditions sur le discours de l'enfant se fait moindre. Nous voyons un recouvrement des classes d'âge de 10, 11 et 12 ans. Ceci est une belle démonstration d'une évolution quasi linéaire entre 4 et 10 ans de la compétence narrative. La stabilisation observée à l'âge de 10 ans et plus pourrait nous conduire à penser que l'enfant aurait acquis dès cet âge les compétences suffisantes et indispensables à l'élaboration d'un récit. Ce qui confirme que c'est vers la fin de l'enfance et au début de l'adolescence que l'enfant devient capable de structurer son récit, prérequis nécessaire à la construction d'une identité stable (MacAdams, 2001 ; Mac Lean, 2005).

Nous avons analysé le discours tout venant de 136 enfants âgés de 3 à 13 ans. Bien que certains auteurs (Gini et al, 2007) n'observent pas de différence liée au sexe de l'enfant dans les productions discursives, d'autres auteurs (Bauer, Stennes & Haight, 2003) montrent des différences dans les narrations autobiographiques dues au genre. Il serait donc intéressant d'introduire cette variable dans nos prochaines analyses.

Afin de mieux cerner le développement de la fonction narrative chez les enfants, une méthodologie plus contraignante pourrait être mise en place. Comme par exemple demander aux enfants de nous raconter un évènement important de leur vie. Colletta (1998) obtient des résultats prometteurs en utilisant cette méthodologie. Cependant, et c'est ici la force et l'originalité de notre recherche, nous perdrons alors la comparaison des conditions de production discursive. Car si les activités diffèrent entre les conditions de groupe et la condition parent, ce n'est pas tant l'activité et les mots utilisés qui nous intéressent mais les catégories employées et la structure morphosyntaxique qui à notre sens permettent une généralisation en fonction de la condition discursive.

Sur le plan méthodologique, mentionnons également que nous prévoyons d'accorder une attention toute particulière à la transcription des données orales, en tirant acquis des recherches

8 Le stade du syncrétisme selon Wallon est un stade où l'enfant perçoit de manière globale et encore avec une certaine confusion les différents éléments (dont lui) de son entourage.

actuelles et des développements considérables qui sont menés actuellement. Enfin, une sélection plus précise des descripteurs au regard des résultats obtenus dans le cadre de la présente étude sera envisagée dans nos recherches ultérieures.

Références

- Bauer, P.J., Stennes, L., & Haight, J-C. (2003). Representation of the inner self in autobiography. Women's and men's use of internal states languages in personal narratives. *Memory, 11*, 27-42.
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge : Cambridge University Press.
- Bideaud, J., Houdé, O., & Pardinielli, J-L. (2003). *L'homme en développement*. Paris : Presses Universitaires de France.
- Boulard, A., Gauthier, J-M. (2012). Quand l'enfant dit « je ». *Enfance*. (paper accepted, December, 2011).
- Brunet, E. (2011). *Ce qui compte. Écrits choisis, tome II – Méthodes statistiques*. Ed C. Poudat. par Céline Poudat. Collection *Lettres numériques*, vol. 11. Paris : Honoré Champion.
- Colletta, J-M. (1998). Les conduites narratives chez l'enfant : une approche étho-linguistique et développementale. *6ème Congrès de l'International Pragmatics Association, Reims, juillet 1998*.
- Gini, M., Oppenheim, D., & Sagi-Schwartz, A. (2007). Negotiation styles in mother-child narrative co-construction in middle childhood: Associations with early attachment. *International Journal of Behavioral Development, 31*(2), 149-160.
- Gratiot-Aphandéry, H., & Zazzo, R. (1976). *Traité de Psychologie de l'Enfant*. Paris : Presses Universitaires de France.
- Habermas, T., & Bluck, S. (2000). Getting a life : The emergence of life story in adolescence. *Psychological Bulletin, 126*, 748-769.
- Karlgren, J. (1994). Text genre recognition using discriminant analysis in International Conference on Computational Linguistics.
- Lebart, L. (2004). *Validité des visualisations de données textuelles* in Actes des 7èmes Journées internationales d'Analyse statistique des Données Textuelles, Louvain-la-Neuve, 10-12 mars 2004, pp.708-715.
- MacAdams, D.P. (2001). The psychology of live stories. *Review of general Psychology, 5*, 100-122.
- Mac Lean, K.C.(2005). Late adolescent identity development: narrative meaning making and memory telling. *Developmental Psychology, 41*(4), 683-691.
- Malrieu, P. (1973). L'expression verbale de la temporalité avant quatre ans. *Bulletin de psychologie, 26*, 340-349.
- Malrieu, D., & Rastier, F.(2001). Genres et variations morphosyntaxiques. *TAL, 42*(2).
- Piaget, J.(1989). *Le langage et la pensée chez l'enfant*. Neuchatel : Delachaux et Niestlé.
- Poudat, C. (2006). Étude contrastive de l'article scientifique de revue linguistique dans une perspective d'analyse des genres. *Texto!*, *11*(3-4). <http://www.revue-texto.net/1996-2007/Corpus/Corpus.html>
- Poudat, C, Gauthier, J.-M, & Boulard, A. (2012, February 08). *Cooccurrences des personnes dans le discours de l'enfant: une approche statistique de la construction de l'identité*. Paper presented at La Cooccurrence: Du fait statistique au fait textuel, Besançon, France.
- Somers, M. (1994). The narrative construction of identity : A relational and network approach. *Theory and society, 23*, 605-649.
- Slama-Cazacu, T. (1966). Le dialogue chez les petits enfants: sa signification et quelques unes de ses particularités. *Bulletin de Psychologie, 19*, 688-697.
- Vygotsky, L.S. (1997). *Pensée et Langage*. Paris: La Dispute.