

Gestes, verbalisations et combinaisons bimodales dans les productions d'enfants français âgés de 18 mois à 4 ans et demi

Aurore Batista¹ & Marie-Thérèse Le Normand²

¹ Laboratoire de Linguistique et Didactique des Langues Etrangères et Maternelles
Université de Grenoble,
aurore.batista@u-grenoble3.fr

² Laboratoire de Psychologie et Neurosciences Cognitives (LPN cog)
Université Paris Descartes,
marie-therese.lenormand@inserm.fr

RESUME

Dans le présent article, nous étudions les productions langagières de jeunes enfants français en situation de jeu quasi-naturelle avec un adulte. Nous savons que les enfants âgés de 17 à 41 mois qui commencent à mobiliser leurs ressources verbales produisent des combinaisons geste-mot. Le rôle de ces combinaisons bimodales a déjà été bien étudié chez les enfants italiens et américains mais ce n'était pas encore le cas chez l'enfant français. Notre premier objectif est donc d'analyser les productions langagières verbales, gestuelles et bimodales d'enfants français pour vérifier qu'on y retrouve bien les évolutions constatées ailleurs. Partant des mêmes données, notre deuxième objectif consiste à réfléchir à une représentation du développement langagier précoce qui, à la différence d'indices telles la LME, prenne en compte l'ensemble des productions gestuelles, bimodales et verbales des jeunes enfants. A cette fin, nous proposons une analyse qui met sur le même plan gestualité et verbalisations, ainsi que l'ébauche d'un modèle bimodal qui reste à affiner par la suite.

ABSTRACT

The current study focuses on language production in young French children during a free play session with an adult. Children aged 17 to 41 months beginning to develop speech also produce gesture-word combinations. Although the role of these bimodal combinations has been well studied for Italian and American children, little is known about these combinations in young French children.

Our first goal is to analyse the verbal, gestural and bimodal language production of French children and verify that the changes which have been found for Italian and American are similar for French. Our second goal is to go beyond structural description of early language acquisition such as MLU types of description, integrating gestural and bimodal aspects of communication. Our proposal is to present a description of language behaviour based on all its aspects, and suggesting a tentative bimodal model that needs to be refined in our future work.

1. INTRODUCTION

L'étude des productions langagières enfantines s'est longtemps focalisée uniquement sur les aspects linguistiques de l'acquisition du langage, ainsi que l'attestent les méthodes proposées pour évaluer les capacités langagières, qu'il s'agisse de questionnaires basés sur le compte rendu parental (Kern, 2007), de batteries de tests (ELO : Khomsi, 2001 ; Isadyle : Piérart, 2005), ou encore, du calcul d'indicateurs tels que la « longueur moyenne des énoncés ou LME » (Parisse et Le Normand 2001, 2006 ; Le Normand, 2007 ; Rondal, 1997).

Pourtant, de plus en plus d'études (McNeill, 1992, 2000 ; Calbris, 2000, Kendon, 2004) tendent à montrer que le langage parlé est multimodal et incorpore des signaux non linguistiques (vocalité et gestualité), et que le développement langagier gagne à être décrit dans ses aspects multimodaux (Morgenstern et al. 2008 ; Rowe et al., 2008 ; Colletta et al., 2010). Par exemple, les observations de l'équipe de Goldin-Meadow concernant les enfants

américains, nous dévoilent clairement que le développement communicationnel de l'enfant ne se résume pas au passage du « stade un mot » (période où les productions linguistiques de l'enfant se résument à des énoncés à un mot) au « stade deux mots » (période où on voit apparaître dans ses productions des énoncés à deux mots). En fait, d'après leurs observations (Butcher et Goldin-Meadow, 2000 ; Goldin-Meadow et Butcher, 2003), vers 12 mois environ, l'enfant commence par utiliser des gestes seuls, qui sont d'ailleurs la plupart du temps des gestes déictiques (exemple : l'enfant pointe du doigt un objet et regarde alternativement la mère et l'objet afin de signifier à sa mère qu'il le veut), puis, à partir de 14 mois, des combinaisons bimodales composées d'un geste et d'un mot (exemple : l'enfant pointe du doigt vers un référent tout en prononçant le mot correspondant), et quelques mois plus tard des énoncés composés de deux mots (exemple relevé dans notre corpus « *voiture chien* »). Il semble donc que le geste apparaisse comme un moyen d'expression transitoire, un tremplin qui permet à l'enfant d'entrer dans la communication verbale à un âge où cette tâche est encore trop complexe pour lui.

Les combinaisons bimodales produites par l'enfant au cours de la seconde année ont donné lieu à des observations plus précises. Ainsi Volterra et son équipe (Volterra et Erting, 1994 ; Volterra et al., 2004 ; Capirci et al., 1996), analysant les premières productions langagières de jeunes enfants italiens, en ont proposé une typologie détaillée, reprise par Morgenstern et al., 2008 ; Estève, 2009 ; Ducey-Kaufmann et al., 2004). En analysant ce que le geste apporte à la parole, les auteurs considèrent trois types de combinaisons :

- soit le geste entretient une relation de « redondance » avec la parole (exemple relevé dans notre corpus : l'enfant pointe du doigt une figurine représentant un chien et dit « chien », ou secoue la tête de droite à gauche et dit « non ») ;
- soit le geste entretient une relation « complémentaire » avec la parole, nécessaire pour comprendre celle-ci (exemple : l'enfant utilise une forme déictique telle « ça » ou « lui » et pointe du doigt le référent, seul le geste permettant d'identifier celui-ci) ;
- soit le geste entretient une relation « supplémentaire » avec la parole car il apporte une information non présente dans la verbalisation (exemple relevé dans notre corpus : l'enfant pointe du doigt une figurine représentant un personnage et ajoute « *tombé* ») ;

Or les chercheurs italiens ont observé que si à 16-20 mois, âge où l'enfant est encore dans la période des énoncés à un mot, l'enfant priviliege les combinaisons redondantes, après 20 mois, alors qu'il a commencé à produire des énoncés à deux mots, il utilise presque uniquement des combinaisons supplémentaires et complémentaires. Ils en concluent que cette évolution va de pair avec l'entrée de l'enfant dans la syntaxe, laquelle se produit en général aux alentours de 18-20 mois, et que cela manifeste l'émergence de la capacité à dire « *quelque chose au sujet de quelque chose* ». Le même phénomène a été attesté ailleurs, chez l'enfant américain (Butcher et Goldin-Meadow, 2000 ; Goldin-Meadow et Butcher, 2003). Les chercheurs américains sont d'ailleurs parvenus à montrer que l'utilisation des combinaisons bimodales à valeur complémentaire et supplémentaire est prédictive de l'entrée dans la période des énoncés à deux mots, autrement dit dans la syntaxe.

La question qui se pose à l'issue de toutes ces observations est de savoir ce qu'il advient des combinaisons gestes-mots au-delà de la deuxième année, lorsque le geste accompagne non plus un seul mot, mais des verbalisations comportant deux mots ou davantage. L'un des objectifs de notre travail est d'apporter des éléments de réponse à cette question, grâce à un vaste corpus de productions langagières d'enfant âgés de 17 à 41 mois. Toutefois, dans le cas présent, nos objectifs sont à la fois plus modestes et circonscrits. Notre premier objectif consiste à utiliser les données provenant de l'exploitation de ce corpus pour vérifier si, chez l'enfant français, on retrouve la même évolution des conduites verbales et bimodales que celle mise au jour chez l'enfant italien et l'enfant américain. Notre second objectif consiste à réfléchir à un indice du développement langagier du jeune enfant qui permette d'intégrer les

aspects gestuels de ses productions, à un âge où celles-ci participent pleinement et nécessairement à l'expression de ses intentions de communication. Commençons par décrire le corpus, ses conditions de recueil ainsi que les données que nous en avons extraites et leur exploitation.

2. METHODE

2.1. Le corpus : conditions de recueil et étendue

Le corpus qui sert de base empirique à la présente étude provient d'un programme de recherche-action soutenu et financé par les localités du bassin grenoblois¹ et intitulé « Parler » (Parler Apprendre Réfléchir Lire Ensemble pour Réussir), programme auquel participe Marie-Thérèse Le Normand, co-auteure de cet article. Les acteurs de ce projet tentent de lutter contre l'échec scolaire des enfants issus de familles défavorisées en intervenant le plus tôt possible (dans le cas présent, en crèche) afin de prévenir les difficultés liées à l'apprentissage du langage oral et écrit. L'action concerne 154 enfants, répartis dans 6 crèches de l'agglomération grenobloise accueillant en majorité des enfants dont la C.S.P. des parents est basse et pour lesquels on peut soupçonner un retard de langage², enfants auxquels s'adresse le programme d'aide et d'entraînement mis au point par les chercheurs et réalisé au cours d'une période de 6 mois.

Afin de mesurer l'effet du programme, la quasi-totalité des enfants a été filmée deux fois (une fois avant le démarrage du programme, une fois après la dernière séance) en situation de jeu et en interaction avec un adulte, selon le protocole de Parisse et Le Normand (2006). Le matériel utilisé en guise de support est la maison de jeu *Fisher Price*. Celle-ci comporte deux chambres, une salle à manger, une cuisine, un garage séparé, ainsi que divers objets (un cheval à bascules, une table, des chaises, une poussette) et des figurines (deux parents, deux enfants, un clown et un chien). L'utilisation de ce support avec tous les enfants homogénéise le format de l'interaction et l'orientation pragmatique de l'activité langagière tout en garantissant le contrôle des entrées lexicales reliées à l'activité ludique. En outre, l'attitude de l'adulte est invariante d'une interaction à l'autre : celui-ci se conduit comme un partenaire de jeu bienveillant qui recentre l'attention de l'enfant sur la panoplie en posant toujours le même genre de questions (exemples : « *Tu mets la voiture dans le garage ?* », « *Ils vont manger les petits hommes ?* », etc.), et se borne à reprendre les verbalisations de l'enfant en proposant des corrections et des expansions (exemple : enfant : « *afien* » ; adulte : « *C'est le chien ? C'est le petit chien de la maison* »). Au total, ce sont 154 interactions d'une durée moyenne de 15mn qui ont été filmées. Au final, le corpus est constitué de 154 séquences de 11 minutes (cette durée correspond à la plus petite interaction enregistrée), collectées au début de chaque interaction, afin de rendre les comparaisons possibles.

Afin d'analyser les productions multimodales en fonction de l'âge, nous avons distingué quatre classes d'âges déjà documentées dans les études précédentes (Volterra et Ertong, 1994 ; Volterra et al., 2004 ; Goldin-Meadow et Butcher, 2003 ; Kern, 2007 ; Morgenstern et al., 2008). Le Tableau 1 présente les effectifs correspondant à chaque classe d'âge.

Classes	17-23 mois	24-29 mois	30-35 mois	36-41 mois
Effectifs	32	55	50	17

¹ Il implique le Conseil Régional Rhône-Alpes, Grenoble-Alpes Métropole et les municipalités d'Echirolles, de Fontaine et de Grenoble, ainsi que L'Inspection Académique de l'Isère et le Laboratoire des Sciences de l'Education de l'Université Pierre Mendès France de Grenoble.

² En raison de leur résultat à une évaluation linguistique. La méthodologie du programme « Parler » est consultable à l'adresse suivante : <http://www.programme-parler.fr/>

Tableau 1: Nombre d'enfants dans chaque classe d'âge

2.2. La transcription et l'annotation des données

L'exploitation des données du corpus a été réalisée à l'aide du logiciel *ELAN®*³, avec lequel nous avons créé une grille d'annotation *ad hoc* pour transcrire et annoter les variables pertinentes. Dans les lignes qui suivent, nous évoquons les choix opérés en matière de transcription, puis d'annotation des verbalisations, des gestes et des combinaisons bimodales. Nous avons choisi de transcrire sur deux lignes séparées les paroles de l'adulte et de l'enfant afin de pouvoir annoter séparément ces dernières. Chaque annotation correspond à un « groupe de souffle », c'est-à-dire à une succession de signaux *voco-verbaux* ininterrompue entre deux pauses, silences ou changements de locuteur. Les pauses chez l'enfant (30 millisecondes en moyenne) étant plus longues que chez l'adulte (20 millisecondes en moyenne), nous avons considéré qu'il n'y avait pas d'interruption de la parole lorsque les pauses étaient inférieures à 30 ms, et avons annoté les groupes de souffle en conséquence, imitant en cela d'autres chercheurs⁴.

La transcription des paroles est orthographique, mais tient compte des hésitations, faux-départs et reprises (exemple : « *le ch le ch ch le chien* ») ainsi que de la prononciation réelle de l'enfant (exemples : l'altération consonantique « *fien* » pour « chien », l'erreur de segmentation aboutissant à « *zoiseau* »). Les marques de remplissages ou *fillers* (Veneziano et Sinclair, 2000) tels « *a* » ou « *è* » placés en début de mot sont transcrits (exemple : si l'enfant dit /apuset/ nous transcrivons « *apoussette* »), de même que les groupes de mots « *formules* » (Chevrier-Muller et Narbonna, 2007) : expressions figées utilisées par l'adulte et que l'enfant s'approprie telles quelles comme /se/ que nous transcrivons « *c'est* » ou /apu/ qui devient « *apu* » ou encore /ela/ qui devient « *est là* ».

Le nombre de mots composant les énoncés étant une variable cruciale lorsqu'on s'intéresse aux premières productions linguistiques enfantines, nous avons choisi de catégoriser les verbalisations de manière très précise, à partir des indications qui suivent. A chaque verbalisation correspondant à un énoncé, est attribué, sur la piste suivante, un « nombre de mots » équivalent à un, deux ou trois mots. Les « énoncés à un mot » comportent une seule unité de sens et peuvent être composés de :

- un mot, c'est-à-dire une unité lexicale (nom, verbe ou autre) dont la forme est conforme à la prononciation adulte ou standard du mot ;
- un quasi-mot, c'est-à-dire une unité lexicale (nom, verbe ou autre) dont la forme est légèrement différente de la prononciation standard (altération phonologique, erreur de segmentation)⁵,
- un ensemble *filler* + mot ou un mot-formule. Comme nous l'avons indiqué ci-avant, ces verbalisations sont à considérer comme une seule unité lexicale.

Les « énoncés à deux mots » comportent deux unités de sens et comportent eux-aussi trois catégories :

- l'assemblage de deux mots dont la forme est conforme à la prononciation standard ;
- l'assemblage d'un mot et d'une unité lexicale autre (quasi-mot, ensemble *filler* + mot ou mot-formule).
- l'assemblage de deux unités lexicales telles celles désignées ci-avant.

Enfin, les « énoncés à trois mots » comportent trois unités de sens : soit trois unités lexicales (quelle que soit leur forme), soit deux unités lexicales assorties d'un morphème (déictique,

³ Disponible à l'adresse : <http://www.lat-mpi.eu/tools/elan/>

⁴ Voir Lacheret-Dujour (2003) ainsi que les conventions de transcriptions ICOR de l'UMR 5191 ICAR (CNRS – Lyon 2 – ENS-LSH).

⁵ L'intérêt de cette catégorie est qu'elle permet de disposer d'informations phonologiques pouvant se révéler cruciales en cours d'analyse.

pronom, déterminant ou autre) dont le statut morphémique est avéré, soit une seule unité lexicale assortie de deux morphèmes (exemples : « *mets bébé là* », « *ça va là* »).

Précisons, pour terminer sur les annotations linguistiques, que la répétition de mots n'est pas comptabilisée dans la catégorisation des énoncés, et que lorsque la verbalisation n'est pas conforme à la syntaxe standard, elle n'est pas normalisée mais prise en compte telle quelle. À titre d'exemple, pour un énoncé ayant la forme « *i i i pati du petit fien* », l'analyse aboutit à considérer qu'il s'agit d'un énoncé à 5 mots : « *i + pati + du + petit + fien* » comportant 5 unités de sens.

En ce qui concerne les gestes, nous avons uniquement pris en compte les gestes déictiques ou de pointage, ce qui exclut les saisies d'objets issues des actions de jeu (comme, par exemple, lorsque l'enfant saisit une figurine pour la mettre dans la chaise ou dans la voiture) et les autres catégories de gestes communicationnels définis par Colletta (2004) et Kendon (2004). Précisons que les gestes déictiques représentent 76% de l'intégralité de la production gestuelle enfantine relevée dans le corpus, ce qui s'explique à la fois par la nature de la situation, qui génère des besoins de désignation, et par l'âge des enfants, âge où les pointages constituent le mode dominant de production gestuelle⁶. Parmi ces gestes, nous avons étiqueté comme tels les gestes de « saisie communicationnelle » (ce que Volterra et al., 2004 nomment les gestes de *ritualized request*), saisies associées au fait de brandir un objet afin de le montrer à l'adulte et accompagnées d'un regard fixe de l'enfant vers l'adulte ou d'un regard alterné de l'enfant vers l'adulte et l'objet. L'annotation des gestes s'effectue sur une quatrième piste. Ceux-ci sont ensuite catégorisés comme des productions isolées (« geste seul ») ou comme intégrés à des « combinaisons » bimodales geste + mot.

S'agissant de ces combinaisons bimodales, nous avons conservé la même typologie que celle proposée par Volterra et al. (2004) et présentée dans la section 1, et en conséquence, nous avons distingué trois types de combinaisons : à valeur redondante, à valeur complémentaire et à valeur supplémentaire. Les annotations correspondantes sont réalisées sur la cinquième piste du fichier de transcription.

2.3. Les variables indépendantes : âge réel et âge linguistique

Rappelons que l'objectif premier de cette étude est de vérifier que l'évolution des conduites langagières enfantines bimodales mise en évidence dans les travaux italiens et américains vaut également pour l'enfant français. Cette évolution peut être mise en évidence de deux façons : soit en comparant les productions à partir de l'âge réel, biographique, des enfants, soit en les comparant à partir de leur âge linguistique. Dans le cas présent, les enfants sont supposés venir de familles défavorisées et présenter davantage de retard de langage, à un âge précoce, que dans d'autres couches de la population, puisqu'un programme d'aide leur est destiné. Il était donc nécessaire de prendre en compte leur âge linguistique, à la fois pour vérifier que les capacités linguistiques de ces enfants se situent effectivement en-dessous du niveau attendu, et également pour disposer d'une variable plus fiable que l'âge réel afin d'atteindre l'objectif rappelé au début de cette section.

De ce point de vue, la LME (Longueur Moyenne d'Enoncé) est un indice intéressant car il est corrélé à l'âge jusqu'à 48 mois (Klee et al., 1989) et peut être calculé en mots (MLU_w ou « *mean length of utterance in words* » ou en morphèmes (MLUM ou « *mean length of utterance in morphemes* »), selon l'objectif visé. Brown (1973) insistait sur l'importance de la calculer en nombre de morphèmes car son intérêt allait au développement grammatical. Or pour notre étude, ce n'est pas tant le développement grammatical qui nous intéresse que le

⁶ Les autres gestes correspondent pour l'essentiel à des emblèmes d'acquiescement et de négation (gestes performatifs) et à de la gestualité représentationnelle (représentation d'actions et d'objets concrets : monter, tourner, pousser, etc.). Leur exploitation est prévue dans la suite de ce travail.

développement lexical, et la manière dont nos jeunes sujets combinent des mots et des gestes. Par conséquent, nous avons opté pour le calcul de la LME en fonction du nombre de mots⁷ selon l'opération suivante :

$$\frac{\text{Nombre total de mots}}{\text{Nombre total d'énoncés verbaux}} \quad ^8$$

Grâce à l'étude princeps de Le Normand (1991) ainsi qu'aux enquêtes de Parisse & Le Normand (2001) et de Le Normand et al. (2008), nous disposons aujourd'hui de valeurs moyennes de la LME par âge pour l'enfant français. Pour apprécier les capacités linguistiques des sujets de notre étude, nous avons donc comparé leur LME à celle obtenue en 2008 par Le Normand et ses collègues pour 316 sujets. Les résultats apparaissent dans la Figure 1 ci-dessous.

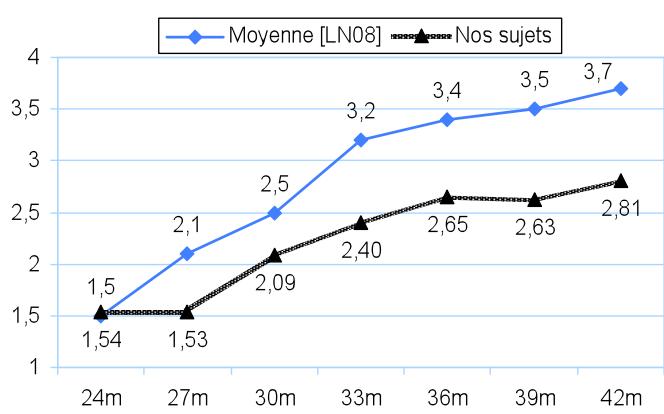


Figure 1 : Valeurs de LME de nos sujets comparées aux valeurs moyennes (Le Normand et al., 2008).

De manière générale, on observe que les valeurs de LME sont plus basses pour nos sujets. Si à 24 mois l'indice est le même pour nos sujets et ceux étudiés par Le Normand et ses collègues, par la suite les valeurs se différencient :

- à 27 mois, la LME de nos sujets stagne aux alentours de 1,5 alors qu'elle dépasse déjà deux mots par énoncés chez l'enfant tout venant ;
- à 30 mois, l'écart se maintient, toujours en faveur de l'enfant tout venant ;
- à 36 mois, nos sujets présentent un indice moyen de 2,5 alors que ceux étudiés par Le Normand et ses collègues produisent en moyenne 3,4 mots par énoncé ;
- enfin, à 42 mois l'enfant tout venant produit toujours en moyenne un mot de plus par énoncé que nos sujets.

En conclusion, on observe effectivement un décalage entre les capacités linguistiques de nos sujets et les capacités linguistiques de sujets tout venant. Le décalage est suffisamment important et constant pour légitimer la nécessité, pour la présente étude, de s'appuyer aussi sur l'âge linguistique des sujets plutôt que sur leur seul âge réel.

⁷ Le calcul de l'indice a été effectué à partir de la transcription des verbalisations sous CLAN®, disponible sur <http://childe.psy.cmu.edu/clan/>

⁸ Dans notre cas, nombre total de groupes de souffles.

3. RESULTATS

3.1. Evolution des productions en fonction de l'âge : des gestes vers les mots

Avant de présenter le détail de nos analyses, il paraît important d'avoir une vue d'ensemble des résultats.

Age	G	E1M	C	E2M	E3M+	Total
17-23	194	248	236	120	186	984
24-29	134	666	521	718	1071	3110
30-35	126	592	400	573	1569	3260
36-41	16	242	90	202	809	1359
Total	470	1748	1247	1613	3635	8713

Tableau 2: Résultats bruts par classe d'âge.

Comme le montre le Tableau 2, dans les données collectées, nous avons identifié 8713 énoncés parmi lesquels 470 gestes seuls (G), 1247 combinaisons geste + mot (C), 1748 énoncés à un mot (E1M), 1613 énoncés de deux mots (E2M) et 3635 énoncés de trois mots et plus (E3M+). Pour vérifier la présence d'un effet de l'âge, il est nécessaire de prendre en compte les différences d'effectifs d'une classe d'âge à l'autre. Nous avons donc calculé, pour chaque catégorie, la valeur moyenne obtenue dans chaque groupe d'âge en divisant le nombre total d'occurrences par le nombre d'enfants correspondant à ce groupe d'âge.

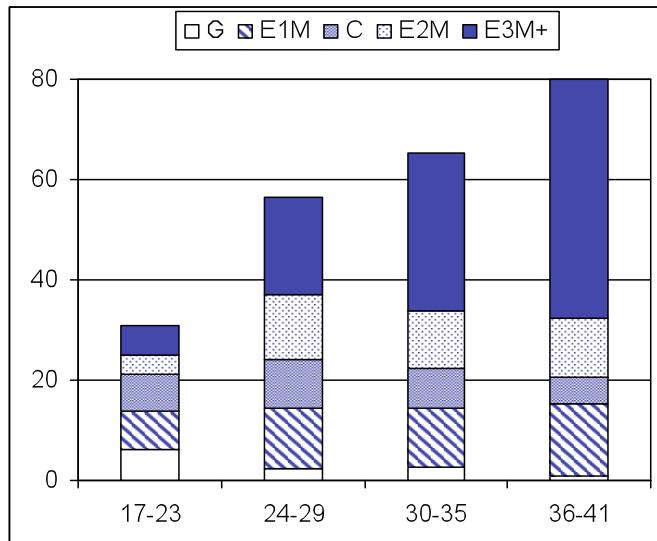


Figure 2 : Moyenne des types de productions par enfant dans chaque classe d'âge.

Commençons par une observation générale : la moyenne des productions (toutes modalités d'émission confondues) augmente régulièrement sur la période observée, elle passe de 31 énoncés chez les plus jeunes à 80 chez les plus âgés.

En observant le détail des productions sur la figure 2, nous pouvons voir que la proportion d'énoncés purement gestuels (G) diminue avec l'âge et tend à disparaître dans la dernière classe d'âge au profit des énoncés à deux mots (E2M), dont la part augmente brusquement à 24 mois (de 3 à 13) mais reste stable ensuite, ainsi que des énoncés à trois mots et plus (E3M+), qui connaissent une croissance très importante et constante dès 24 mois.

En revanche, la part des combinaisons geste + Mot (C) et des énoncés à un mot (E1M) reste à peu près stable de 17 à 41 mois. D'ailleurs, ce phénomène nous interpelle : si l'enfant devient capable de produire des énoncés à plusieurs mots, pourquoi continue-t-il à produire des gestes en accompagnement des mots dans les combinaisons bimodales ? Certes, leur part n'évolue

pas, mais la nature de ces combinaisons change. Nous notons en effet que si elles sont généralement redondantes chez les plus jeunes de nos sujets (45% de combinaisons redondantes contre 22% de supplémentaires et 23% de complémentaires dans la première classe d'âge), elles deviennent majoritairement supplémentaires (77% chez les plus âgés), comme nous l'avons constaté ailleurs (Batista et Colletta, 2009). On retrouve ici l'évolution attestée dans les données italiennes et américaines, une évolution qu'on peut analyser comme signalant l'émergence des capacités syntaxiques, puisque avec les combinaisons supplémentaires, l'enfant devient en mesure de « dire quelque chose au sujet de quelque chose » (Volterra, 2004), c'est-à-dire d'exprimer une prédication, comme il le fait ensuite dans les énoncés à deux mots.

Autre phénomène à noter : en examinant les données recueillies dans chaque groupe d'âge, il apparaît qu'à aucun moment les énoncés à deux mots ne sont majoritaires. Tout se passe comme si, à la différence de la période des énoncés à un mot, celle des énoncés à deux mots n'existe pas. En d'autres termes, une fois capable d'enchaîner deux mots dans sa production, l'enfant produit très vite des énoncés plus longs comportant trois mots ou davantage, et au vu de ces données, on peut difficilement parler d'un « stade des deux mots » (Goldin-Meadow, 2003). Mais une autre lecture des mêmes données est possible, cette fois en faisant abstraction de la modalité pour ne considérer que le nombre d'éléments composant l'énoncé.

3.2. La période des énoncés à deux éléments comme période de transition

Si à l'instar d'Estève (2009) ou Millet et Estève (2009) nous considérons que les deux modalités expressives (verbale/gestuelle) contribuent de façon identique à la composition de l'énoncé, une autre lecture des données devient possible en comptabilisant simplement le nombre d'éléments composant l'énoncé. Les énoncés à un élément (1EL) comprennent donc les énoncés à un mot (E1M) et les gestes seuls (G) ; et les énoncés à deux éléments (2EL) englobent les combinaisons geste-mot (C) et les énoncés à deux mots (E2M). En ce qui concerne les énoncés à trois éléments et plus (3EL+), ils comprennent pour l'instant seulement les énoncés à trois mots et plus (E3M+)⁹.

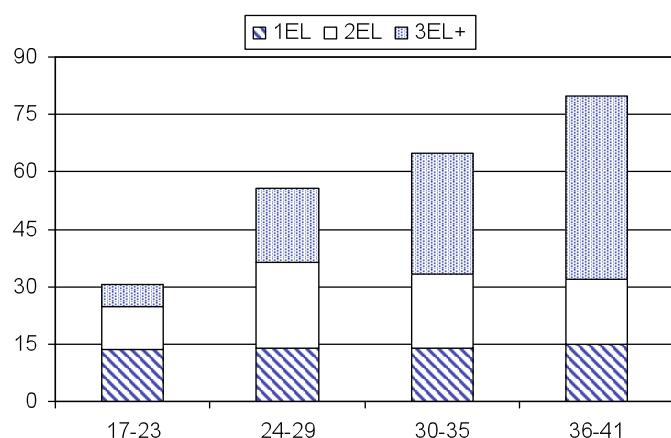


Figure 3 : Moyenne des types de production (nombre d'éléments) par enfant dans chaque classe d'âge.

À la lecture de la figure 3, on voit apparaître trois périodes distinctes caractérisant les productions de nos sujets :

⁹ À terme, ils comprendront les combinaisons comportant une verbalisation de deux mots ou plus accompagnée d'un geste.

- les enfants de la première classe d'âge communiquent majoritairement grâce à des énoncés composés d'1EL. Le nombre d'énoncés à un élément reste ensuite stable tout au long de la période étudiée ;
- chez les enfants de la seconde classe d'âge, ce sont les énoncés à deux éléments (2EL) qui deviennent majoritaires. Par la suite, le nombre de ceux-ci n'augmente pas et se stabilise, lui aussi.
- Enfin, à partir de 30 mois (troisième classe d'âge), les énoncés produits par l'enfant sont en majorité des énoncés comportant au minimum trois éléments (3EL+). Le nombre de ceux-ci augmente ensuite, chez les enfants de la quatrième classe d'âge.

Les différences observées sont-elles statistiquement significatives ? La comparaison des variances ne montre pas de différence significative en ce qui concerne l'emploi des 1EL à travers les quatre classes d'âge. Par contre elle met en évidence des différences significatives en ce qui concerne les 2EL et les 3EL+. Ainsi, pour les 2EL, l'augmentation de leur emploi est significative (voir le tableau 3) entre la première et la seconde classe d'âge : les enfants âgés de moins de 24 mois produisent moins d'énoncés à deux éléments que leurs pairs plus âgés. Pour les (3EL+), on note un accroissement significatif de leur emploi dès 24 mois, et bien sûr à 30 mois, dans la classe d'âge où ils deviennent majoritaires en nombre, ainsi qu'à 36 mois, bien que la différence soit moins nette au plan statistique (Tableau 4, dernière ligne : $p=.066$). Contrairement à ce qui se passe pour les 2EL, l'évolution de l'usage des 3EL est visible tout au long de la période étudiée, et les énoncés longs sont de plus en plus employés au fil de l'âge.

2EL		
17-23 vs 24-29 F(85)=13,831		p=.000

Tableau 3 : Comparaison significative pour les 2EL.

17-23 vs 24-29	F(85)=7,893	p=.006
24-29 vs 30-35	F(104)=4,695	p=.033
30-35 vs 36-41	F(66)=3,493	p=.066

Tableau 4 : Comparaisons significatives pour les 3EL+.

Ainsi donc, si on ne peut, au vu de nos données, accorder de réalité empirique au « stade des deux mots », la présente analyse aboutit à mettre en évidence une période au cours de laquelle les énoncés à deux éléments, qu'ils soient de nature verbale ou de nature bimodale, sont effectivement majoritaires parmi les productions de l'enfant. Par la suite, nous projetons de réaliser une analyse qualitative pour comprendre plus précisément ce qu'il se passe avant et après cette période de transition où émergent les premières constructions syntaxiques, entre les productions initiales de l'enfant, brèves et limitées à un mot ou un geste, et le moment où la plupart de ses énoncés correspondent à des verbalisations longues. À présent, examinons nos données en partant non plus de l'âge des sujets, mais de leurs performances linguistiques effectives, mesurées à l'aide de la LME comme on l'a expliqué dans la section 2.

3.3. Vers un modèle bimodal du développement langagier ?

La comparaison des productions dans les quatre classes d'âge a fait apparaître une nette évolution de celles-ci en même temps que la part importante qu'y prennent la gestualité et les combinaisons bimodales. Mais les performances linguistiques de nos sujets sont en moyenne inférieures à celles de sujets tout venant. Pour disposer d'un tableau évolutif plus général, il nous faut désormais analyser leurs productions à partir de leur âge linguistique. À cette fin,

nous avons répartis les enfants dans trois groupes correspondant à trois niveaux d'habileté linguistique :

- le premier groupe rassemble les enfants ayant une LME comprise entre 1 et 1,5 (non inclus), et qui en sont au stade initial, avec une production linguistique majoritaire d'énoncés à un mot ;
- le second groupe rassemble les enfants ayant une LME comprise entre 1,5 et 2,5 (non inclus), et dont les productions sont majoritairement, au plan linguistique, des énoncés à deux mots ;
- le troisième groupe rassemble les enfants ayant une LME supérieure à 2,5 et qui produisent en majorité des verbalisations plus longues.

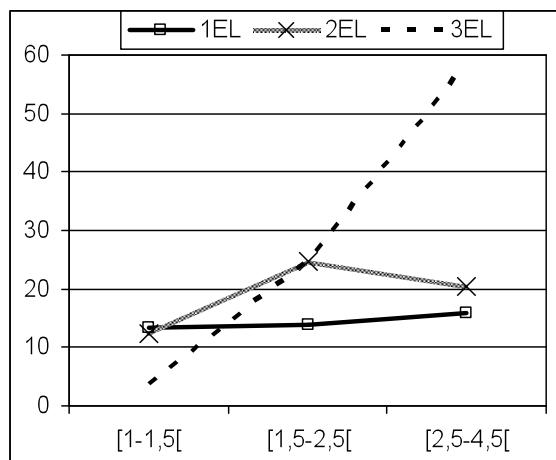


Figure 4 : Moyenne des types de production (nombre d'éléments) par enfant en fonction de la LME.

Comme cela apparaît sur la figure 4, les enfants ayant une LME inférieure à 1,5 produisent autant d'énoncés à un (1EL) qu'à deux éléments (2EL), mais très peu d'énoncés à trois éléments et plus (3EL). Les enfants dont la LME est comprise entre 1,5 et 2,5 (groupe intermédiaire), quant à eux, communiquent en utilisant autant d'énoncés à deux éléments qu'à trois éléments et plus, mais moins d'énoncés à un élément (même si le nombre de ces derniers reste stable). Enfin, les enfants dont la LME est supérieure à 2,5 communiquent essentiellement à l'aide de verbalisations longues, même si le nombre des 1EL et des 2EL reste stable dans leurs productions.

Avec cette analyse, on retrouve évidemment l'évolution des productions constatée à partir de l'âge de nos sujets, mais en partant de l'âge linguistique et non plus de l'âge biographique, elle prend une portée plus générale. Il reste à présent à affiner à nouveau l'analyse pour tenter de proposer un modèle de développement bimodal des productions langagières qui intègre les éléments gestuels. La figure 5 présente l'évolution de la répartition des cinq types de productions que nous avons distingués au départ en fonction du niveau de LME de l'enfant et fait apparaître trois profils langagiers :

1. Le premier profil (LME inférieure à 1,5) correspondrait à l'enfant qui communique encore majoritairement à l'aide d'énoncés comportant un seul élément (un mot ou un geste) et de combinaisons, même si des verbalisations plus longues (2M ou 3M+) peuvent apparaître dans son répertoire. À ce stade, l'enfant utilise et combine des éléments issus de différents modes communicationnels pour créer du sens et exprimer ses intentions de communication.
2. Le second profil (LME comprise entre 1,5 et 2,5) correspondrait à l'enfant qui communique majoritairement à l'aide d'énoncés comportant un ou deux éléments et chez qui les verbalisations longues (3M+) restent encore assez peu employées. À ce

stade, l'enfant privilégie les verbalisations courtes, même s'il continue à faire usage des éléments gestuels.

3. Le troisième profil (LME supérieure à 2.5) correspondrait à l'enfant qui communique majoritairement à l'aide de verbalisations longues et chez qui les moyens de communication gestuels (employés seuls ou combinés) deviennent minoritaires. À ce stade, la plupart des énoncés de l'enfant correspondent à des verbalisations sans le geste.

Issue de l'analyse encore partielle de nos données, il manque encore à cette tentative de modélisation la prise en compte des productions précoces, avant 18 mois, période au cours de laquelle il est possible d'isoler un stade antérieur au cours duquel l'enfant ne sait pas encore combiner deux éléments. Il lui manque aussi l'analyse bimodale des verbalisations longues, puisqu'on peut trouver des combinaisons comportant un geste et plusieurs mots. Il lui manque enfin la prise en compte non plus seulement des pointages gestuels, mais des autres gestes que peut produire l'enfant de cet âge, en particulier des emblèmes sociaux et des gestes à visée représentationnelle. C'est donc seulement à l'issue de cette extension des observations qu'on sera en mesure de proposer et tester un modèle bimodal du développement langagier.

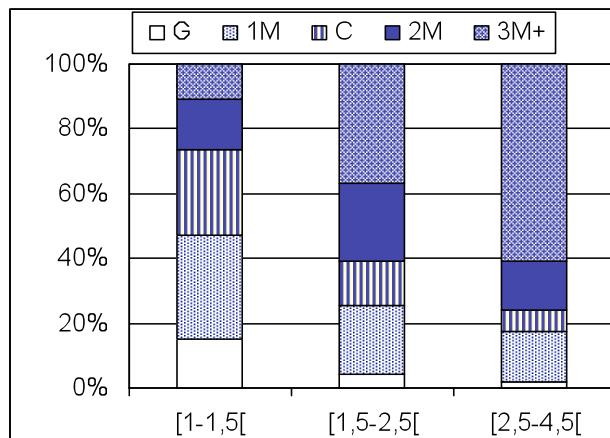


Figure 5 : Répartition des types de production en fonction de la LME (%).

4. CONCLUSION

Deux objectifs majeurs guidaient cette étude :

1. Etudier l'évolution des productions langagières enfantines entre 18 mois et quatre ans et demi afin de vérifier que celle-ci est conforme, au moins pour la période correspondante, aux évolutions mises en évidence dans les travaux italiens et américains comparables et qui ont pris en compte les moyens gestuels de communication utilisés par les jeunes enfants.
2. Réfléchir à une nouvelle lecture de cette évolution dans le but de proposer un modèle bimodal du développement langagier qui intègre l'ensemble des moyens de communication utilisés par le jeune enfant : verbalisations, mais aussi gestes et combinaisons bimodales.

Concernant le premier objectif, l'évolution mise en évidence, faite d'une diminution de l'emploi des gestes seuls, d'une utilisation relativement stable des énoncés à un mot et des combinaisons bimodale, et d'un accroissement important des énoncés à deux mots (vers deux ans) puis des verbalisations longues (au-delà), confirme celle attestée dans les données italiennes et américaines. Elle signale tout à la fois la préférence de l'enfant, au fil de l'âge,

pour les verbalisations, et le rôle important joué par les combinaisons bimodales. En effet, ces dernières évoluent nettement au plan qualitatif, comme nous l'avons montré ailleurs (Batista et Colletta, 2009), et signalent ainsi l'émergence des capacités syntaxiques dès avant deux ans. Il reste à présent à compléter ce tableau évolutif en intégrant les changements qualitatifs qui se produisent dans les combinaisons geste-mot lorsque le geste accompagne des énoncés comportant plus d'un mot, et nous allons nous y employer dans la suite de notre travail, en focalisant nos observations sur les traces de l'émergence d'une syntaxe plus complexe.

Concernant le second objectif, nous avons d'abord procédé à une relecture des données qui neutralise la différence de modalité et met sur le même plan gestualité et verbalisation. Celle-ci fait apparaître une période de transition : le moment où les productions combinant deux éléments (qu'il s'agisse d'énoncés à deux mots ou de combinaisons bimodales) sont majoritaires dans les conduites langagières de l'enfant. Puis, après avoir neutralisé la variable sociologique inhérente au choix de l'échantillon et pris comme base d'analyse l'âge linguistique plutôt que l'âge biographique, nous avons proposé l'ébauche d'un modèle bimodal du développement langagier précoce, en trois phases. Comme toute ébauche, cette tentative de représentation reste à compléter et affiner avant de pouvoir être testée. Mais nous pouvons d'ores et déjà préciser que, comme nous disposons de deux observations à six mois de distance pour la quasi-totalité de nos sujets, il sera possible de commencer à tester ce modèle à partir de nos données. L'apparition récente de tests d'évaluation des capacités langagières intégrant des éléments gestuels (Coquet & al., 2007) montre en tous les cas qu'il est temps de réfléchir à une modélisation de ce type.

REMERCIEMENTS

Nous remercions les protagonistes du programme « Parler » ainsi que Jean-Marc Colletta pour sa relecture attentive de l'article.

REFERENCES

BATISTA A. et COLLETTA J.-M. (2009) : « Analyse des productions multimodales d'enfants français âgés de dix-sept à quarante et un mois en situation de jeu », /Actes du colloque AcquisiLyon, 3-4 décembre 2009/, Disponible à l'adresse : http://www.ddl.ish-lyon.cnrs.fr/colloques/AcquisiLyon/pageweb/Fichier/ALY_Complet.pdf

BROWN R.-W. (1973) : *A first language: the early stages*, Cambridge, Harvard University Press.

BUTCHER C. et GOLDIN-MEADOW S. (2000) : « Gesture and the transition from one- to two-word speech: When hand and mouth come together », McNeill, D. (Ed.) /Language and gesture/, pp. 235-257, New York, Cambridge University Press.

Calbris G. (2000) : « Espaces symboliques révélés par l'étude sémantique du geste coverbal », /Cahier du C.I.E.L./, 1.

Capirci O., Iverson J., Pizzuto E. et Volterra V. (1996) : « Gesture and words during the transition to two-word speech », /Journal of Child Language/, 23, pp. 645-673.

Chevrier-Muller C. et Narbonna J. (2007) : *Le langage de l'enfant : aspects normaux et pathologiques*, Paris : Masson.

Colletta J.-M. (2004) : *Le développement de la parole chez l'enfant âgé de 6 à 11 ans*, Hayen, Belgique : Mardaga.

Colletta J.-M., Pellenq C. et Guidetti M. (2010) : « Age-related changes in co-speech gesture and narrative: Evidence from French children and adults », /Speech Communication/, 52, pp. 565-576. doi:10.1016/j.specom.2010.02.009.

Coquet F., Roustit J. et Jeunier B. (2007) : « La Batterie Evalo 2-6. Évaluation du langage oral et des comportements non verbaux du jeune enfant », /Rééducation Orthophonique/, 231, pp. 203-226.

Ducey-Kaufmann V., Abry C., Vilain A. et Lalevée C. (2004) : « Les segments du corps en interaction dans la naissance de la parole », /Interacting bodies [Corps en interaction], Festschrift in Honour to Peter MacNeilage/

(Barcelona ICPHS Satellite Conference 1-2 August 2003), 388, Disponible à l'adresse : http://gesture-lyon2005.ens-lyon.fr/article.php3?id_article=227.

Estève I. (2009) : « Approches bilingue et multimodale de la construction langagière chez l'enfant sourd : quelles pistes pour l'évaluation ? », /Communication au Colloque Multimod 2009, 9-10-11 juillet 2009/, Toulouse.

Goldin-Meadow S. (2003) : /The resilience of language/, New-York, Psychology Press.

Goldin-Meadow S. et Butcher C. (2003) : « Pointing toward two-word speech in young children », In S. Kita (Ed.), /Pointing: Where language, culture, and cognition meet/, pp. 85-107, Mahwah, Earlbaum Associates.

Kendon A. (2004) : /Gesture. Visible action as utterance/, Cambridge, Cambridge University Press.

Kern S. (2007) : « Le compte-rendu parental comme méthode d'évaluation du développement langagier précoce », /Rééducation orthophonique/, 231, pp. 139-153.

Khoms A. (2001) : /Evaluation du langage oral/, Paris, ECPS Editions.

Klee T., Schaffer M., May S., Membrino S., et Mougey K. (1989) : « A comparison of the age-MLU relation in normal and specifically language-impaired preschool children », /Journal of Speech and Hearing Research/, 54, pp. 226-233.

Lacheret-Dujour A. (2003) : « Structure communicative et géométrie intonative : que nous dit la synthèse de la parole ? », /Cahiers de Linguistique de Louvain/, pp. 88-114.

Le Normand M.-T. (1991) : « Individual differences in the production of word classes in eight specific language-impaired preschoolers », /Journal of Communication Disorders/, 24, pp. 331-351.

Le Normand M.-T. (2007) : « Evaluation de la production spontanée du langage oral et de l'activité sémantique du récit chez l'enfant d'âge préscolaire », /Rééducation Orthophonique/, 231, pp. 53-72.

Le Normand M.-T., Parisse C. et Cohen H. (2008) : « Lexical diversity and productivity in French preschoolers: Developmental, Gender and Sociocultural factors », /Clinical Linguistics and Phonetics/, 22, pp 47-58.

MacWhinney B. (2009) : « Volume 2: Transcription Format and Programs, Part 2: The CLAN Programs », /Tools for Analyzing Talk/, Carnegie Mellon University.

Mc Neill D (1992) : /Hand and mind. What gestures reveal about thought/, Chicago, University of Chicago Press.

McNeill D. (2000) : « Growth Points, Catchments, and Contexts », Kita S. ed., /Japanese Journal of Cognitive Science/ (special issue on gesture), 7, pp. 22-36.

Millet A. et Estève I. (2009) : « Contacts de langues et multimodalité chez des locuteurs sourds : concepts et outils méthodologiques pour l'analyse », /Journal of Language Contact/, Disponible à l'adresse : <http://www.jlcjournal.org/>

Morgenstern A., Leroy M. et Mathiot E. (2008) : « Le pointage chez l'enfant : origines et fonctions ? », /Congrès mondial de linguistique française. Psycholinguistique, Acquisition/, pp. 1805-1818, Disponible à l'adresse : http://www.linguistiquefrancaise.org/index.php?option=com_article&access=standard&Itemid=129&url=/articles/cmlf/pdf/2008/01/cmlf08170.pdf.

Parisse C. et Le Normand M.-T. (2001) : « Local and global characteristics in the development of morphosyntax by French children », /First Language/, 21 (62), pp 187-203.

Parisse C. et Le Normand M.-T. (2006) : « Une méthode pour évaluer la production du langage spontané chez l'enfant de 2 à 4 ans », /Glossa/, 97, 20-41.

Piérart B. (2005) : /Le langage de l'enfant. Comment l'évaluer ?, Bruxelles, De Boeck.

Pinker S. et Jackendoff R. (2005) : « The faculty of language : what's special about it? », /Cognition/, 95, pp. 201-236.

Rondal J.-A. (1997) : /L'évaluation du langage/, Hayen, Mardaga.

Rowe M. L., Özçaliskan S. et Goldin-Meadow S. (2008) : « Learning words by hand: Gesture's role in predicting vocabulary development », /First language/, 28, pp. 182-199.

Veneziano E. et Sinclair H. (2000) : « The changing status of “filler syllables” on the way to grammatical morphemes », /Journal of Child Language/, 27(3), pp. 461-500.

Volterra V., Caselli M.C., Capirci O., et Pizzuto E. (2004) : « Gesture and the emergence and development of language », M. Tomasello & D. Slobin (Eds.), */Beyond Nature-Nurture. Essays in Honor of Elizabeth Bates/*, pp. 3-40. London : Erlbaum.

Volterra V. et Erting C.-J. (1994) : */From gesture to language in hearing and deaf children/*, Gallaudet University Press.