

# La prosodie chez le bébé au stade pré-linguistique : premières formes stables

Christelle Dodane\* & Karine Martel\*\*

\*Laboratoire Dipralang, Université Paul Valéry Montpellier 3  
Route de Mende, 34 000 MONTPELLIER - France  
[christelle.dodane@univ-montp3.fr](mailto:christelle.dodane@univ-montp3.fr)

\*\*Laboratoire PALM, Université de Caen Basse-Normandie  
Esplanade de la Paix, 14032, CAEN – France  
[karine.martel@unicaen.fr](mailto:karine.martel@unicaen.fr)

## ABSTRACT

This research investigated the role of prosody in the pre-linguistic development of children. Longitudinal case-study of two monolingual French babies in spontaneous interaction with their parents is presented. Children were videotaped and audio-recorded between ages of ten and twelve months in their home. Each utterance was coded in function of its global melodic contour and its interaction type (dialogue vs. monologue). Two categories of recurring prosodic patterns emerge: 1°) self-centred vocal productions, characterized by falling or flat melodic contours; 2°) vocal productions with specific melodic contours associated to specific situations of interaction with adult (call, initiation of communication, etc.). We assume that these prosodic forms could constitute the earliest units allowing the transition from pre-linguistic to linguistic level.

## 1. INTRODUCTION

Au cours des stades précédant l'apparition du premier mot (stade holophrastique), l'étude des productions vocales chez les bébés représente une véritable gageure en raison de la difficulté à isoler des unités pertinentes. Ces difficultés sont accrues par la grande variabilité intra et inter-individuelle qui caractérise l'entrée des jeunes enfants dans le langage. Ainsi, les questions se posent de savoir 1°) quelles unités vont émerger de cette variabilité et se stabiliser progressivement ? et 2°) comment déterminer l'existence et la nature de ces unités ? Parmi les différentes possibilités qui s'offrent à l'enfant pour « structurer » ses premières productions vocales et interagir avec son environnement, il semble que la prosodie représente un chemin privilégié en raison de son statut tout-à-fait particulier au début de la vie. Les éléments musicaux et prosodiques se trouvent favorisés dans le monde sonore intra-utérin, la paroi abdominale se comportant comme un filtre passe-grave [1], [2]. Les voix, bien que déformées, émergent clairement du bruit de fond et, en particulier, la voix maternelle [3]. Dès la naissance, les bébés développent une très grande sensibilité aux caractéristiques prosodiques de leur langue maternelle [4], [5]. Ils marquent d'ailleurs une forte préférence pour le langage qui leur est adressé (« motherese » ou L.A.E. ; [6], [7]). Ce langage aux caractéristiques prosodiques fortement exagérées [8], [6] se caractérise notamment, par une mise en saillance des frontières entre les constituants

de différentes tailles (pauses, allongements de syllabes, montées et chutes de Fo, variations d'intensité) [9] [10]. Ainsi, des bébés de 2 à 4 mois sont capables de discriminer des syllabes au contenu segmental identique, mais affectées de contours d'intonation différents [11], [12] et des bébés anglophones de 2 mois, de discriminer les contrastes prosodiques de l'anglais, mais seulement lorsque la parole leur est adressée [13]. Ils sont donc capables de différencier les contours prototypiques de leur langue maternelle et de répondre de manière adéquate aux messages qu'ils véhiculent, tels que l'incitation au réveil, à l'attention et au contact visuel [14]. En ce qui concerne le lexique, dès le premier mois, ils se montrent sensibles aux mots fortement accentués et à la terminaison des phrases [15] et vers 2-3 mois, ils ne font d'ailleurs pas attention aux mots prononcés sans exagération de la prosodie [16]. Parallèlement à cette sensibilité prosodique, la compréhension du sens des mots se développe, puisque les bébés sont capables vers 10-12 mois de comprendre une trentaine de mots en contexte. La prosodie pourrait donc contribuer à fixer le format des différentes unités, « *en offrant les premiers indices pour la segmentation et le groupement des syllabes en unités-mots* » [17 : 104]. Il en résulte que, au niveau de la production, les propriétés rythmiques caractérisant la langue maternelle apparaissent longtemps avant les premiers mots. Ainsi, le « protolangage » de l'enfant contient déjà vers 6-8 mois les caractéristiques rythmiques et intonatives de sa langue maternelle [18], [19], ce qui le rend reconnaissable par des auditeurs de la même communauté linguistique [20]. On observe ainsi une récurrence de patrons mélodiques affectés de traits contrastifs constants et précis, qui s'opposent entre eux. L'enfant fait une utilisation linguistique des schémas mélodiques, la mélodie du babillage étant devenue intonation. A la fin du dixième mois, l'enfant dispose déjà de quelques schémas prosodiques oppositifs, mais vides. Ces schémas forment un cadre prosodique qui va progressivement se remplir et recevoir le niveau segmental [21]. Au stade du premier mot, les schémas intonatifs et la rythmique spécifique de la langue maternelle jouent un rôle essentiel en fournissant le format prosodique des premières productions. L'enfant produit de vrais énoncés, structurés au niveau intonatif et rythmique, dans lesquels apparaissent vers 12 mois, les premières formes lexicales reconnaissables par l'adulte (le premier mot ou stade holophrastique). Ces émissions vocales ont des significations complexes et sont interprétées par

l'entourage comme de véritables phrases révélatrices d'une attitude intentionnelle, dont la fonction est de désigner, d'exprimer et d'ordonner. C'est à ces schèmes prosodiques que s'intéresse notre étude et notre objectif est d'observer leur apparition et leur fixation progressive au sein même de l'interaction avec l'adulte. Il est probable en effet que ces schèmes représentent une étape charnière pour l'enfant dans l'acquisition de sa langue maternelle, si ce n'est la condition préalable à la production d'énoncés plus complexes. Nous posons donc l'hypothèse que non seulement la prosodie va permettre la structuration progressive des énoncés enfantins, mais qu'elle pourrait également constituer un marqueur précoce de l'émergence linguistique. Plus spécifiquement, nous pensons que : 1°) la prosodie va permettre l'émergence et la stabilisation de formes prosodiques spécifiques (répertoire de base de contours prototypiques), 2°) le contexte communicationnel va engendrer/contraindre l'utilisation de ces formes qui vont s'ancrer à un contexte spécifique au sein d'unités premières d'acquisition que sont les actes conversationnels.

## 2. MÉTHODOLOGIE

### Corpus / Participants

Les productions spontanées de deux enfants francophones, Madeleine (MAD) et Théophile (THE) ont été analysées. Ces derniers sont enregistrés de l'âge de 10 à 12 mois et à leur domicile, lors des interactions avec leur entourage. Ces enregistrements font partie du corpus Léonard<sup>1</sup>, qui rassemble les données longitudinales filmées de 10 enfants monolingues et bilingues en situation naturelle avec leur entourage.

**Table 1 :** Vidéos du corpus Léonard<sup>1</sup> utilisées pour les analyses.

Vidéos	Durée	Durée exploitée	Nbre énoncés	Nbre énoncés analysés
MAD 10 mois	1.22.34	0.32.30	191	160
MAD 11 mois	0.45.01	0.45.01	230	207
MAD 12 mois	1.19.00	1.19.00	--	--
THE 10 mois	0.56.46	0.28.30	129	49
THE 11 mois	0.58.17	0.58.17	224	118
THE 12 mois	1.41.00	0.56.05	284	104
<b>TOTAL</b>			<b>1058</b>	<b>638</b>

### Transcriptions, codage et analyses acoustiques

Toutes les prises de parole des sujets ont été transcrites phonétiquement, grâce au logiciel CLAN (format Chat, CHILDES). Cependant, seules 638 prises de parole enfantines ont fait l'objet d'une analyse détaillée, les autres étant écartées pour diverses raisons (énoncés non linguistiques, chevauchements, environnement bruités...). Par ailleurs, la mélodie de chaque énoncé a été codée à l'oreille (ligne %pro) par un musicien entraîné. En l'occurrence, cette analyse a consisté à catégoriser la configuration du contour global comme « Plat », « Montant », « Descendant », « Montant-descendant »,

« Cloche » ou « Complexe ». Afin de tester la validité de ce codage, des analyses acoustiques ont été réalisées sur chaque énoncé avec le logiciel Praat. Les phrases ont été échantillonnées à 44kHz, 16 bits en mono. Le contour de Fo a été extrait (100-1000 Hz), puis post-traité (suppression des sauts d'octave, lissage, interpolation). En outre, chaque énoncé a également été identifié comme unité « monologuée » ou « dialoguée ». Les unités monologuées (mono) sont produites lorsque l'enfant joue seul et manipule des objets sans regarder l'adulte, qui ne répond d'ailleurs pas (productions vocales autocentrées). Les unités dialoguées (dialo) autrui, lors d'une activité conjointe avec un adulte par exemple. Enfin, une troisième catégorie a été créée, la catégorie « indéterminée », afin de classer les unités ambiguës (les productions sont espacées par des pauses longues, l'enfant sort du champ de la vidéo, la production est en réalité un son végétatif...). Les informations concernant la situation (ligne %sit), c'est-à-dire les activités des locuteurs ont aussi été mentionnées. Et enfin, la qualité de la voix était également codée lorsqu'elle était affectée (voix craquée, murmurée... : ligne %qua).

## 3. RÉSULTATS

### 3.1. Résultats et discussion

#### 3.1.1. Répartition des types d'interaction selon l'âge

Chez MAD, la majorité des productions est réalisée en situation dialoguée (66,25% à 10 mois, soit 106 énoncés sur 160 et 62,31% à 11 mois, soit 129 énoncés sur 207 ; cf. Table 2). A 10 mois, 33,75% des productions sont produites en situation monologuée, dont 23,75% en situation de jeu solitaire ou jasis.

**Table 2 :** Répartition du nombre des énoncés en fonction des situations d'interaction.

	MAD10	MAD11	THE10	THE11	THE12
<b>MONO</b>	16	129	48	64	46
<b>DIALO</b>	106	78	75	54	58
<b>JASIS</b>	38	0	0	0	0
<b>TOTAL</b>	160	207	123	118	104

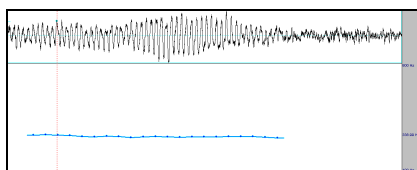
Cette proportion est un peu plus importante à 11 mois (37,65%) en raison de la situation : la mère ménage beaucoup de temps de jeux pour sa fille. Chez THE, on relève, à 10 mois, une plus grande proportion d'énoncés en situation monologuée (26,19 %, soit 22 énoncés sur 126), par rapport à la situation dialoguée (12,69%, soit 16 énoncés sur 126). Cependant, la majorité des productions n'a pu être clairement catégorisée (61,11%, soit 77 énoncés sur 126). A 11 mois, la catégorisation devient plus aisée puisque 57,24% des énoncés de THE sont monologués (soit 79 sur 138), le reste étant dialogué (42,75%, soit 59 énoncés sur 138). Cette proportion s'inverse à 12 mois puisque 60,97% des énoncés sont dialogués (soit 75 énoncés sur 123), le reste étant monologué (39,02%, soit 48 énoncés sur 123).

#### 3.1.2. Répartition des contours en situation monologuée

Chez MAD, les contours sont majoritairement plats et descendants en situation monologuée. Ainsi, à 10 mois, 50

<sup>1</sup> ANR « Acquisition du Langage et Grammaticalisation », n°JC0547273 (<http://anr-leonard.ens-lsh.fr>) – Porteuse de projet : Aliyah Morgenstern, ENS-LSH Lyon.

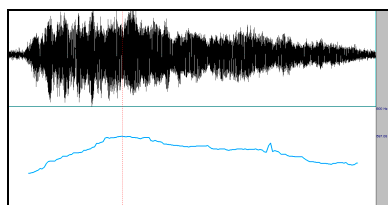
% des productions (soit 8 énoncés sur 16) sont supportées par un contour plat et 31,25% (soit 5 énoncés sur 16) par un contour descendant en situation monologuée. La proportion de contours plats atteint 78,26% à 11 mois (101 énoncés sur 207), contre 13,95% d'énoncés descendants (soit 18 énoncés sur 129). Chez THE, on relève également une prédominance pour les contours plats (51,51 %, soit 17 énoncés sur 33) à 10 mois, 11 mois (37,5%, soit 24 énoncés sur 64) et 12 mois (58,69%, soit 27 énoncés sur 46). Ainsi, les productions réalisées par les deux enfants en situation monologuée sont caractérisées par une forte proportion de contours plats. Il semble par conséquent que ce contexte particulier dans lequel les jeunes locuteurs ne cherchent pas à transmettre un contenu, entraîne un minimum de modulation, de nuances. Cette proportion semble même se renforcer avec l'âge chez THE.



**Figure 1 :** Contour plat produit par THE à 10 mois (338 Hz), en situation monologuée.

### 3.1.3. Répartition des contours en situation dialoguée

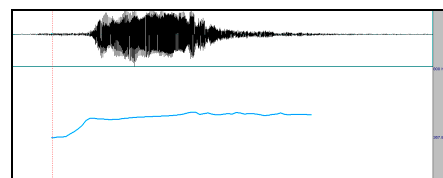
Si les énoncés produits étaient relativement semblables en situation monologuée chez les deux enfants, il en va tout autrement pour les productions réalisées en situation dialoguée. Les profils deviennent complètement différents. Ainsi, chez MAD, on relève à 10 mois un répertoire de courbes mélodiques très varié (cf. Table n°3).



**Figure 2 :** Contour en cloche produit par MAD à 10 mois (344/597/400 Hz), en situation dialoguée.

Les contours en cloche sont utilisés dans l'interaction pour initier et/ou maintenir le contact avec l'adulte (16,03% des contours dialogués ; cf. figure n°2), ce qui donne lieu à de véritables rencontres vocales par un jeu d'imitation mutuelle (concordance de hauteur à 570 Hz). Stern et al. [22] relèvent d'ailleurs que les mères utilisent ce type de contours pour établir et maintenir le contact oculaire avec leur bébé. On trouve également des contours descendants (23,58%), des contours montant-descendant (24,52%), des contours plats (16,98%) et des contours complexes (8,49%) mais sans pattern fortement prédominant. A 11 mois en revanche, la complexité semble se réduire et MAD semble montrer une prédilection pour les contours montants (47,43 %), utilisés pour solliciter l'attention voire une réponse de la part des adultes. Parmi ces contours montants, on trouve un contour spécifique lié à l'appel, à chaque fois clairement identifié par l'entourage (cf. figure n°3). Enfin, on relève des contours plats (23,07%) et dans une proportion

moindre, des contours descendants (14,10%), des contours complexes (8,97%) et des contours montant-descendants (6,41%). Contrairement à MAD, la proportion de contours plats domine fortement en situation dialoguée à 10 mois (68,75%), même si elle se réduit par la suite (46,29% à 11 mois et 39,65% à 12 mois). Ce phénomène est notamment lié au fait que les productions de THE sont plutôt brèves. Toutefois, THE produit également un autre contour de façon privilégiée à 11 et 12 mois, le contour complexe (enchaînement de plusieurs contours simples).



**Figure 3 :** Contour lié à l'appel produit par MAD à 11 mois (367-507 Hz), en situation dialoguée.

Les contours en cloche, encore peu nombreux à 10 mois (6,25%), servent à formuler le désir qu'une action intéressante se poursuive. Leur proportion augmente à 11 mois (11,11%) et à 12 mois (20,68%) et supportent notamment une nouvelle production [ai], utilisée par THE pour prolonger l'interaction.

**Table 3 :** Répartition des énoncés en % en fonction des types de contours utilisés en situation dialoguée.

	Montant	Descendant	Plat	Cloche	Montant-descendant	Complexe
MAD10	10,37%	23,58%	16,98%	16,03%	24,52%	8,49%
MAD11	47,43%	14,10%	23,07%	0%	6,41%	8,97%
THE10	12,5%	0%	68,75%	6,25%	0%	12,5%
THE11	0%	12,96%	46,29%	11,11%	0%	5,55%
THE12	13,79%	15,51%	39,65%	20,68%	1,72%	8,62%

### 3.1.4. Interprétation et discussion

Il ressort de ces résultats que les deux enfants produisent des énoncés sensiblement différents en fonction de la situation d'interaction. En situation monologuée, MAD et THE produisent une majorité de contours plats, ainsi que des contours descendants, et ceci, à tous les âges étudiés. Ces productions sont produites en situation de jeu et s'accompagnent le plus souvent d'une activité exploratoire motrice ou visuelle. L'enfant manipule des objets pour lui-même et ne semble pas solliciter l'attention de l'adulte, qui d'ailleurs ne lui répond pas (productions « autocentrées », [23]). L'utilisation de contours plats ou descendants est le plus souvent associée à une qualité de la voix altérée (voix faible voire chuchotée, craquements vocaux...) et à une séquence phonétique simple (V ou CV). En revanche, en situation dialoguée (interaction avec l'adulte lors d'une activité conjointe, ou interaction pure), les enfants utilisent une voix plus claire et plus forte lorsqu'ils s'adressent à un adulte et les contours mélodiques sont beaucoup plus variés. Cependant, on relève de grandes différences entre les deux enfants. Ainsi à 10 mois, le répertoire de MAD est apparemment plus riche que celui de THE. Quand elle a l'occasion de s'adresser à sa mère pour initier, appeler, répondre ou imiter, elle utilise des contours mélodiques plus

diversifiés. Par exemple, à 10 mois, elle utilise la mélodie pour provoquer de véritables rencontres vocales avec sa mère dans lesquelles l'une et l'autre s'imitent mutuellement (concordances de Fo, cf. 3.1.3). A 11 mois cependant, cette diversité se restreint et MAD produit une majorité de contours ascendants. THE de son côté, manifeste un comportement très différent en situation dialoguée. Il produit une majorité de contours plats, bien que cette proportion tende à diminuer à 12 mois, au profit d'autres contours spécifiques à la situation d'interaction, tels que les contours en cloche et les contours montants. Il existe donc un décalage entre les deux enfants, MAD manifestant une utilisation plus riche et plus précoce de la prosodie en situation d'interaction. Ces différences, outre la variabilité intra-individuelle caractérisant les enfants à cet âge, peuvent être également liées au fait que les adultes eux-mêmes s'adressent différemment à leurs enfants. Ainsi, la mère de MAD réserve du temps au sein des différents tours de parole afin que sa fille puisse interagir avec elle. En réponse, MAD sollicite davantage l'attention de sa mère. De leur côté, les parents de THE produisent de nombreuses onomatopées auxquelles leur fils fait « écho » dans ses productions, constituant ainsi de véritables « séquences-écho » [24] (reprise du rythme, du débit, de la mélodie ou des phonèmes). Il semble donc que l'émergence et la stabilisation progressive des formats prosodiques soient étroitement liées à leur contexte d'apparition au sein de l'interaction mère-enfant [19].

#### 4. CONCLUSION

Malgré une certaine variabilité, les productions des deux enfants semblent se différencier nettement en fonction de la situation d'interaction, entre l'âge de 10 et 12 mois. Alors que les productions « monologuées » se caractérisent essentiellement par des contours mélodiques finals plats ou descendants et une qualité de la voix affectée, les productions dialoguées se caractérisent par une plus grande diversité de contours, associés à une situation d'interaction spécifique (appeler, raconter, etc.). Ces formats prosodiques stables semblent constituer des unités précoces permettant aux enfants d'interagir pleinement avec leur entourage bien avant l'apparition des premiers mots. La plupart du temps, les intentions des deux enfants étudiés sont d'ailleurs correctement interprétées par leur entourage. Nous pensons que ces formats prosodiques pourraient constituer les unités les plus précoces permettant la transition entre les niveaux pré-linguistique et linguistique. La question se pose de savoir quel rôle ces formats prosodiques vont jouer lors des stades ultérieurs (émergence des premières unités linguistique, émergence de la première syntaxe).

#### BIBLIOGRAPHIE

- [1] Armitage, S.E., Baldwin, B.A., Vince, M.A. (1980). The foetal sound environment of sheep. *Science*, 208, 1173-1174.
- [2] Lecanuet, J.P. (1995). L'expérience auditive prénatale. In Delième, I., Sloboda, J. A. *Naissance et Développement du Sens Musical*. Paris : P.U.F., 7-38.
- [3] Querleu, O., Renard, X. Versyp, F., Paris-Delrue, L., Crépin, G. (1988). Fetal hearing. *European Journal of Obstetrics and Reproductive Biology*, 29, 191-212.
- [4] Mehler, J., Bertoncini, J., Barrière, M., Jassik-Gershenfeld, D. (1978). Infant recognition of Mother's voice. In *Perception*, 7, 491-497.
- [5] Jusczyk, P. W. (1998). "Dividing and conquering linguistic input", *CLS34: The Panels*: 293-310.
- [6] Fernald, A. (1989). Intonation and communication intent in mother's speech to infants: is the melody the message? *Child Development*, 60, 1497-1510.
- [7] Cooper, R. P., Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, 66, 1584-1595.
- [8] Papousek, M., Papousek, H. (1981). Musical elements in the infant's vocalizations: Their significance for communication, cognition and creativity. Lipsitt, Lewis P. and Rovee-Collier, Carolyn K. *Advances in Infancy Research*. Vol I: XXII. Norwood: Ablex, 163-224.
- [9] Christophe, A., Dupoux E., Bertoncini, J., Mehler, A. (1994). Do infants perceive word boundaries ? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America*, 95, 1570-1580.
- [10] Millotte, S. (2005). *Le rôle de la prosodie dans le traitement syntaxique adulte et l'acquisition de la syntaxe*, Thèse de Doctorat, Laboratoire de Sciences Cognitives et Psycholinguistique.
- [11] Morse, P.A. (1972). "The discrimination of speech and non-speech stimuli in early infancy". *Journal of Experimental Child Psychology*, 14, 477-492.
- [12] Kuhl, P. K., Miller, J. D. (1982). Discrimination of auditory target dimensions in the presence or absence of variation in a second dimension by infants. *Perception and Psychophysics*, 31, 279-292.
- [13] Best, C. T. (1993). Emergence of language-specific constraints in perception of non-native speech a window on early phonological development. In Boysson-Bardies, Schonen, Jusczyk, MacNeilage and Morton, *Developmental Neurocognition: speech and face processing in the first year of life*. Boston: Kluwer Academic Publishers, 289-304.
- [14] Papousek, M., Bornstein, M. H., Nuzzo, C., Papousek, H. Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behavior and Development*, 13, 539-545.
- [15] Menyuk, P. (1977). *Language and Maturation*. Cambridge: MIT Press, 180 p.
- [16] Kagan, J., Lewis, M. (1965). Studies of attention in the human infant. In *Merill-Palmer Quarterly*, 11, 95-122.
- [17] Bertoncini, J., de Boysson-Bardies, B. (2000). La perception et la production de la parole avant deux ans. Kail, Fayol, *L'Acquisition du Langage : le Langage en Emergence de la Naissance à Trois ans*, Vol. I. Paris : PUF, 95-136.
- [18] Konopczynski, G. (1990). *Le Langage Emergent : Caractéristiques Rythmiques*. Hamburg : Buske Verlag, 362 p.
- [19] Konopczynski, G. (1991). *Le Langage Emergent II : Aspects Vocaux et Mélodiques*. Hamburg : Buske Verlag, 425 p.
- [20] Boysson-Bardies, B., Sagart, L., Durand, C. (1984). Discernible differences in the babbling of infant according to target language. *Journal of Child Language*, 11, 1-15
- [21] Konopczynski, G., Tessier, S. (1994). Structuration intonative du langage émergent. Dans Halford et Pilch, 157-192.
- [22] Stern, D.N., Spiker, S., MacKain, K. (1982). Intonation as signals in maternal speech to pre-linguistic infants. *Developmental Psychology*, 18, 727-735.
- [23] Piaget, J. (1950). Introduction à l'Epistémologie Génétique. Tome III : La Pensée Biologique, la Pensée Psychologique et la Pensée Sociale, P.U.F.
- [24] Deshaies, D., Paradis, C. (1998). Fonctions de l'énoncé-écho : le rôle de la prosodie. *Orage 1998 : Oralité et Gestualité. Communication Multimodale et Interaction*. Paris : L'Harmattan, 435-440.