

Marie-Amélie COUTANSAIS  
OrthophonisteMaster 2 recherche en sciences  
du langage  
ILPGA  
19 rue des Bernardins  
75005 ParisCentre référent des troubles des  
apprentissage  
Hôpital Raymond Poincaré  
104 avenue Raymond Poincaré  
92380 Garches

ma\_coutansais@yahoo.fr

**RÉSUMÉ :**

Ce travail est une analyse acoustique de babillages d'enfants de neuf mois, chinois et français (à l'aide du logiciel Praat), et une analyse perceptive de babillages d'enfants du même âge chinois, anglais et français. L'analyse acoustique consiste en des mesures de la fréquence fondamentale et de la durée des voyelles ainsi qu'en une analyse segmentale qualitative du type de productions. L'analyse perceptive consiste en un jury d'écoute, composé de juges français experts et naïfs, qui doivent identifier le bébé français lorsqu'il est apparié soit avec un anglais soit avec un chinois. Le corpus est composé d'items dissyllabiques, tirés des enregistrements de Levitt\* et ses collègues.

Les résultats de l'analyse acoustique montrent une différence significative au niveau de la fréquence fondamentale entre les babillages chinois et français. En effet, la fréquence fondamentale baisse plus sur la dernière voyelle pour les babillages chinois. La différence d'allongement de la deuxième voyelle par rapport à la première entre les babillages chinois et français est marginalement significative. La deuxième voyelle chez les bébés français est ainsi plus longue que chez les bébés chinois. Par ailleurs, l'étude qualitative du type de productions dans ces babillages chinois et français permet d'observer une complémentarité de la structure segmentale et des caractéristiques prosodiques.

Les résultats de l'analyse perceptive montrent que les juges naïfs sont significativement capables d'identifier le babillage français dans les paires français/anglais (59,7 %). Cependant, on observe une grande variabilité des réponses des juges dans cette tâche de perception.

Notre travail, réalisé dans le cadre d'un master 2 recherche en sciences du langage (spécialité phonétique), sous la direction de Pierre Hallé (2006-2007), a pour objectif de confronter nos résultats avec ceux de la littérature, notamment ceux des études de Levitt et ses collègues\* et ceux de De Boysson-Bardies et collègues\*\*.

**MOTS-CLÉS :**

Développement - Langage - Perception - Mesures - Jeune enfant - Babillage.

# ANALYSE ACOUSTIQUE ET PERCEPTIVE DE BABILLAGES FRANÇAIS, ANGLAIS ET CHINOIS D'ENFANTS DE 9 MOIS : UNE ÉTUDE DE CERTAINS PARAMÈTRES SUPRA-SEGMENTAUX

par Amélie COUTANSAIS

## SUMMARY : Acoustic and perceptive analysis of 9-month-old children babblings in French, English and Chinese: a study of some suprasegmental parameters

This study is divided into an acoustic analysis of 9-month-old children babblings in Chinese and French (with the software Praat) and a perceptive analysis of babblings of children of the same age in Chinese, English and French. The acoustic analysis consists of  $F_0$  measures and vowels durations, and also of a qualitative and segmental study of the babblings. For the perceptive analysis, French experts and naïve judges had to identify French babblings in French-English and French-Chinese pairs. The babblings we chose are bisyllabics and are coming from Levitt and colleagues data\*. The acoustic analysis results show a significant difference between the  $F_0$  measures in Chinese and in French babblings. Indeed, the  $F_0$  decreases more in the last vowel of the Chinese babblings. The second vowel lengthening difference is marginally significant between Chinese and French babblings. The second vowel is thus longer in French babblings. Moreover, the qualitative study of the babblings reveals complementarities between the segmental structure and the prosodic parameters.

The perceptive analysis results show that the naïve judges are in a significant way able to identify the French babblings among the French-Chinese pairs (59,7 %). However, we notice an important variability in the judges' answers.

Our investigation was made for a master 2 in language sciences (phonetics) directed by Pierre Hallé. Its aim was to compare our results with those described in the literature, in particular with the studies made by Levitt and colleagues\* and by de De Boysson-Bardies and colleagues\*\*.

**KEY-WORDS :**

Development - Language - Perception - Measures - Young child - Babbling.

L'acquisition du langage est un processus complexe qui passionne depuis de nombreuses années les chercheurs. Comment apprend-on sa langue maternelle ? Comment le bébé fait-il pour acquérir le langage ?

La littérature, abondante sur le sujet, rapporte de nombreuses expériences et observations visant à savoir quels processus cognitifs se mettent en place. Certaines études analysent l'apprentissage de la langue au niveau perceptif, c'est-à-dire comment le bébé gère l'input auditif de son environnement, comment il le traite et ce qu'il en fait pour le transformer en acquisition langagière. On peut parler ici rapidement de la maturation du système auditif, dès la vie intra-utérine, la perception des phonèmes d'abord universelle, puis spécifique à la langue acquise, du « motherese », etc.

Le développement perceptif, qui montre la grande attention de l'enfant pour son environnement sonore, précède celui de la production. La production sonore du bébé (productions réflexes, exploration des capacités vocales, pré-babillage, babillage, premiers mots) serait en quelque sorte le reflet des processus cognitifs en place chez le bébé. L'approche cognitive de l'acquisition du langage se fonde en effet sur les procédures qui guident cet apprentissage à partir de « connaissances innées ». Le bébé semble en effet avoir des prédispositions pour discriminer et catégoriser les aspects pertinents du langage. Puis des processus de sélection se mettent en place lors du babillage, qui permettent au bébé d'extraire les éléments pertinents dans sa langue maternelle, et enfin apparaissent des premiers mots, qui témoignent d'une association production/sens. Ensuite, l'enfant pourra acquérir les paramètres grammaticaux de sa langue maternelle. Ainsi, l'étude des productions du bébé nous apprend beaucoup sur l'évolution de ces acquisitions. L'étude du développement du langage chez l'enfant nous permet de savoir quel est le déroulement normal des acquisitions langagières, mais aussi de constater d'éventuelles pathologies du développement. En effet, bien que chaque enfant soit différent dans son rythme d'acquisition et de développement, certaines altérations ou absences alertent sur le développement de l'enfant.

Ce travail de recherche consiste en une étude des aspects suprasegmentaux du babillage d'enfants de neuf mois, issus de trois milieux linguistiques monolingues (français, anglais et chinois), et se divise en une analyse perceptive d'une part et une analyse acoustique d'autre part. L'hypothèse de ce travail est de montrer que les babillages sont différents à cet âge et que les différences sont perceptibles par des adultes. De plus, nous voulons voir quelles sont les corrélations acoustico-perceptives des babillages, selon la langue.

## LE DÉVELOPPEMENT DU LANGAGE CHEZ L'ENFANT

### LE DÉVELOPPEMENT PERCEPTIF DU LANGAGE

De nombreuses études apportent des éclairages sans cesse approfondis sur le développement perceptif du langage chez le nourrisson. Cette acquisition de la langue débute dès la vie intra-utérine\* : à partir de la 25<sup>e</sup> semaine d'aménorrhée, le fœtus possède un système auditif fonctionnel, qui lui permet dès lors d'entendre la voix de sa mère. En effet, la voix maternelle émerge bien des bruits internes : les propriétés spectrales et la structure prosodique de la langue maternelle sont conservées. Ainsi, des nourrissons après 12 heures de vie seulement préfèrent la voix de leur mère\*. L'acquisition de la langue maternelle semble commencer ainsi dès la naissance, et même avant. Des nourrissons de 3-4 jours de vie sont capables de discriminer la quasi-totalité des contrastes utilisés dans les langues du monde. Ensuite, les acquisitions apparaissent de façon rapide.

À deux mois, les bébés accordent un traitement tout particulier à la syllabe, perçue comme un tout\*. À quatre mois, ils montrent une perception catégorielle de syllabes consonne/voyel-

\*De Boysson-Bardies, 1996, 2005

\*De Casper et coll., 1980

le. À cinq mois, les bébés négligent les variations de locuteur et d'intonation pour une voyelle donnée, qu'ils reconnaissent comme une seule et même voyelle. Enfin, il a été montré que les bébés sont sensibles à l'exagération du rythme et des contours prosodiques propres au discours que l'entourage utilise en s'adressant à un enfant « motherese ».

Cette introduction au développement perceptif du langage est évidemment loin d'être exhaustive. Cependant, comme notre travail porte sur le versant productif du langage, nous allons plutôt nous attarder sur ce point.

## LE DÉVELOPPEMENT DES PRODUCTIONS LANGAGIÈRES

### 1 - Le développement vocal de l'enfant

Selon B. de Boysson-Bardies, dès trois mois\*, le voile du palais s'avance, s'abaisse et peut fermer l'arrivée d'air au nez. La langue s'allonge, sa musculature se développe et l'ouverture du pharynx permet à la langue de se mouvoir d'avant en arrière. De plus, un changement se produit dans le contrôle respiratoire : en effet, dès cinq mois, les bébés sont capables d'utiliser leur activité respiratoire et leur larynx à peu près comme les adultes. Cependant, la transformation du conduit vocal ne sera pas encore achevée à la fin de la première année. Ce n'est pas avant l'âge de cinq-six ans que le contrôle de l'ensemble des articulateurs deviendra possible. Les mouvements globaux sont maîtrisés avant les mouvements fins, le contrôle du bout de la langue et des lèvres sera le dernier à être acquis (5-6 ans).

Entre deux et cinq mois, les bébés ne vocalisent qu'en position couchée\*. Ainsi leurs productions incluent-elles presque uniquement des sons issus du larynx ou du velum. Le nourrisson ne maîtrise pas sa phonation : ce n'est que vers quatre ou cinq mois qu'il devient capable de moduler les variations de sa voix. Dès ce moment, le bébé manipule les traits prosodiques de hauteur de la voix, le niveau sonore, certains traits consonantiques (bruits de friction, de murmure nasal, bilabiales roulées, uvulaires). Il joue avec ses articulateurs. Les premières voyelles apparaissent à cette période.

Vers la fin du sixième mois\*, le bébé est capable de coordonner globalement des ajustements phonatoires et supraglottaux : il commence à pouvoir interrompre ses vocalisations volontairement. Il peut également imiter des schémas d'intonation simples à la suite d'exemples d'adultes.

Certaines des productions du babillage sont plus courtes (dès quatre et cinq mois) et incluent des séquences de type consonne-voyelle ressemblant à des syllabes. Mais ces pseudo-syllabes n'ont pas les caractéristiques de syllabes de langues parlées\*. Durant cette période de « prébabillage », l'enfant fait des gammes en manipulant surtout des sons vocaliques. Il joue à faire varier les intonations, les successions, les durées.

Selon certains auteurs, les tout premiers sons produits par le bébé seraient des sons plutôt laryngés. Des études récentes\* montrent que les premières vocalisations, dans les six premiers mois, seraient des réflexes phonétiques des mécanismes employés lors de la succion et de la déglutition. En effet, le mécanisme physiologique de la déglutition a les mêmes caractéristiques articulatoires que les sons pharyngaux ou certains types de phonation observés dans certaines langues du monde. Ce sont les mêmes gestes que l'on effectue. Selon l'hypothèse initiale d'Esling, lors du babillage, les bébés dont la langue environnémentale possède des registres de phonation ou des sons pharyngaux se servent de ces sons lors de la production des syllabes. Avec son équipe, Esling a étudié quatre langues : le français, l'anglais, l'arabe et le bai (dialecte chinois). L'arabe et le bai sont des langues qui possèdent des sons pharyngaux, à la différence du français et de l'anglais. À la naissance, la configuration laryngée est presque fermée (l'épiglotte et la luette sont en contact) afin de permettre au nourrisson de s'alimenter tout en respirant. Les premières constriction observées par Esling sont des coups d'épiglotte : toutes les articulations produites par le bébé auraient ainsi une constriction, il n'y aurait pas d'articulation ouverte. Un contrôle efficace du mécanisme laryngé primaire serait ainsi présent dès la naissance. Dans les premiers mois, le contrôle de la phonation s'établit et s'affine. Selon Esling, la

\*De Boysson-Bardies, 1996, 2005

\*De Boysson-Bardies, 1996, 2005

\*De Boysson-Bardies, 1996, 2005

\*Oller, 1980

\*Esling, 2006 ; Esling et coll., 2005

parole commence dans le larynx et le pharynx. Les premières vocalisations du bébé seraient toujours accompagnées d'une constriction du sphincter laryngé. Les modes articulatoires sont pratiqués au départ au niveau du pharynx, c'est-à-dire dans la partie supérieure du larynx, à l'endroit où le sphincter effectue la fermeture totale. Les enfants de toutes les langues utiliseraient de façon efficace ces mécanismes primaires pour explorer leurs capacités vocales et plus tard les caractéristiques phonétiques de leurs langues.

Esling et son équipe\* ont trouvé que la variation dans la production des différents modes et lieux d'articulation s'acquiert surtout par le moyen d'un amorçage laryngien. Les sons appris et pratiqués pendant le prébabillage déclenchent phonétiquement les sons du babillage. Selon eux, la base laryngienne serait innée dans le développement de l'enfant. Elle serait le point de repère de la production motrice. Leur hypothèse émergente est la suivante : les sons laryngaux (avec constriction) du prébabillage mènent aux sons à prédominance orale (sans constriction) du babillage. Les sons laryngaux et pharyngaux réapparaîtraient plus tard dans les langues qui les possèdent dans leur système phonologique. En effet, dans leurs travaux, Esling et ses collègues observent une forte utilisation des constriction laryngées dans les premiers mois de vie. Ceci évolue dans la deuxième moitié, selon la langue maternelle. Par exemple, dans leur étude, ils observent pour l'anglais beaucoup de pharyngales et glottales au début (sons qui n'existent pas en anglais en tant que phonèmes), puis beaucoup moins par la suite (vers 6 mois), avec une augmentation des consonnes alvéolaires, vélares et bilabiales (présentes en anglais). En ce qui concerne les registres de phonation, Esling et ses collègues observent au début une voix avec une forte constriction, puis des vocalisations sans constriction (voix modale, falsetto, voix breathy). Ces observations confirment selon les auteurs l'hypothèse selon laquelle les enfants utiliseraient des sons pour lesquels ils sont biologiquement prédisposés pour explorer leur voix, puis sélectionneraient les sons de leur langue.

\* 2005

## 2 - Le développement des productions langagières au niveau segmental (phonèmes)

Selon Hallé\*, « les productions précoces des enfants reflètent l'interaction entre tendances universelles et propriétés spécifiques de la langue particulière de l'environnement ». On observe dans les babillages des différentes langues du monde des grandes tendances universelles : dominance des occlusives et surtout des labiales, voyelles centrales, associations consonne/voyelle sur le lieu d'articulation. Cependant, des tendances spécifiques émergent assez tôt.

\*2004

Vihman et De Boysson-Bardies\* ont étudié la nature et l'origine des influences du langage environnant. Elles ont observé les effets globaux de ces influences durant la première année de vie et elles concluent à des effets manifestes concernant le développement des caractéristiques phonétiques auxquelles l'enfant est suffisamment exposé dans sa langue maternelle.

\*2000

Oller et ses collègues\* ont voulu comparer des enfants de milieu monolingue à des enfants de milieu bilingue, de quatre à dix-huit mois afin de voir s'il existe des différences. Selon les auteurs, il n'y a pas de différences : les enfants atteignent le stade du babillage canonique au même âge dans les deux groupes. Ils observent également des similarités dans la production des syllabes correctes et des sons vocaliques, tous milieux socio-économiques confondus, enfants nés à terme ou prématurés. Selon ces auteurs, le développement vocal dans la première année résiste aux conditions environnantes.

\*1997

Dans les productions des bébés, les voyelles semblent devenir spécifiques à la langue maternelle plus tôt (entre 8 et 10 mois) que les consonnes. Certains auteurs\* ont enregistré 20 enfants âgés de dix mois français, algériens, anglais britanniques et cantonnais et effectué une analyse spectrale des premier (F1) et deuxième (F2) formants des voyelles produites par les enfants. Ils trouvent des différences significatives entre les quatre langues, différences qui correspondent à celles retrouvées chez les adultes. Selon les auteurs, il existe une représentation de l'espace vocalique de la langue qui permet à l'enfant de réaliser des voyelles en fonction des caractéristiques des voyelles perçues.

\*De Boysson-Bardies et coll., 1989

\*Levitt, Utman, 1992

D'autres auteurs\* ont réalisé une étude longitudinale d'un enfant français et d'un enfant anglais et les ont enregistrés à 5, 8, 11 et 14 mois. Ils observent des similarités initiales : répertoire consonantique, contrôle des F1 et F2 des voyelles et changements développementaux dans le babillage (patterns universels). Ensuite, le babillage diffère progressivement pour rejoindre les patterns de la langue environnante. Les productions reflètent les fréquences des phonèmes adultes, l'anglais produisant par exemple plus de syllabes fermées (caractéristique de l'anglais).

\*Sussman et coll., 1999

D'autres auteurs encore\* ont effectué une analyse des productions consonne-voyelle (CV) d'un bébé de 7 à 40 mois (avec comme consonnes une labiale, une alvéolaire et une vélaire) et ont observé le F2 des voyelles en fonction du lieu d'articulation de l'occlusive. Ils ont trouvé une progression des productions CV avec les labiales, alvéolaires et vélaire vers les normes adultes de coarticulation.

\*Whalen et coll., 1995

On constate que les voyelles produites par les enfants se ressemblent au départ, correspondant à des patterns universels. Puis, ces productions vont s'affiner pour s'ajuster aux patterns de la langue maternelle (fréquences des formants, effets de coarticulation), pour ensuite ressembler aux cibles adultes. Cependant, il existe dans les productions adultes des paramètres que l'on retrouve dans toutes les langues. Par exemple, les voyelles /i/et/u/ont des fréquences fondamentales (Fo) « intrinsèques » plus hautes que /a/. Cette Fo intrinsèque a été retrouvée chez les enfants à différents stades du développement, mais n'avait pas été étudiée dans le babillage. Une équipe s'est donc penchée sur la question\* : elle a étudié les voyelles de six enfants français et six enfants anglais à 6, 9 et 12 mois. Les résultats montrent que la Fo intrinsèque apparaît aussi dans le babillage indépendamment de la langue acquise. Selon les auteurs, la Fo intrinsèque est une conséquence automatique (mécanique) de la production de ces voyelles.

\*Vilain et coll., 2000

En ce qui concerne l'acquisition des consonnes, certains auteurs\* ont utilisé un modèle anthropomorphique (réalisé à partir de données cinéradiographiques) dont les paramètres représentent les degrés de liberté du tractus vocal d'un locuteur français. Ils ont voulu voir si ce modèle peut décrire l'acquisition du langage, du babillage canonique au contrôle global des syllabes CV. Selon ces auteurs, les bébés commencent par contrôler des strictions locales, puis développent des stratégies de contrôle global. Par exemple, au départ, il y aurait selon eux *peu de labiales*, car le bébé doit apprendre la position de la langue pour la voyelle suivante indépendamment de la fermeture des lèvres. Ceci contredit d'autres données de la littérature selon lesquelles les labiales sont très majoritaires dans le babillage. En effet, les données basées sur des transcriptions confirment la prépondérance initiale des labiales\*. Ensuite, vers 14 mois, le nombre de labiales augmente ou diminue pour atteindre la proportion adulte.

\*De Boysson-Bardies, Vihman, 1991 ; Locke, 1983

\*Vilain et coll., 2000

Les auteurs précédents\* observent quant à eux beaucoup d'alvéolaires au début, dont le nombre décroît rapidement jusqu'à trois ans. Selon eux, le bébé est d'abord capable de produire des sons avec les deux parties de la langue nécessaires aux alvéolaires et aux voyelles, ce qui explique leur apparition plus précoce. Puis, le bébé développe le contrôle moteur des différents degrés de liberté de mouvement de la langue, guidé par l'input auditif de son environnement. Le nombre de vélaire reste relativement stable, dans la proportion adulte, dès la fin du premier mois. Ils relèvent aussi plus d'items CV avec la consonne et la voyelle qui partagent les mêmes points d'articulation. Ces auteurs pensent que même si le bébé a la capacité de produire des strictions consonantiques différenciées, il doit encore apprendre le contrôle global du tractus vocal. Dans ce sens, leur approche est compatible avec la théorie de MacNeilage « Frames, then Content »\*.

\*1998

\*McCune, Vihman, 2001

Une étude longitudinale réalisée sur vingt enfants de 9 à 16 mois, avec des critères de productivité longitudinaux\*, analyse les patterns vocaux produits en relation avec le lexique, les auteurs cherchant à démontrer la relation phonétique/lexique de l'acquisition du langage. Leurs résultats montrent que le nombre de consonnes spécifiques produites de façon consistante au fil des mois prédit l'utilisation lexicale de référence à 16 mois. Elles observent que la paire d'occlusives p/b est la première à être utilisée de façon conséquente (prépondérance des occlusives labiales, contrairement à ce que trouvent Vilain et coll.) et observent également un lien entre les consonnes et les mots pro-

duits par l'enfant dans les deux mois suivant l'observation : les mots contiennent surtout les consonnes produites précédemment.

De Boysson-Bardies et Vihman\* ont quant à elles fait l'étude des processus par lesquels l'enfant adapte ses capacités initiales aux productions phonétiques spécifiques à sa langue maternelle. Cette étude consiste en des analyses de la distribution de la place consonantique (lieu d'articulation) chez cinq enfants français, cinq japonais, cinq anglais et cinq suédois, de 10 à 17 mois, afin de déterminer :

\*1991

- quelles différences systématiques peut-on trouver dans les babillages et premiers mots en fonction de la langue maternelle,
- dans quelle mesure ces différences sont liées à la structure phonétique de la langue environnante.

Les résultats montrent que les enfants, tous confondus, utilisent beaucoup plus de labiales et de dentales que de vélares et plus d'occlusives que les autres modes d'articulation. On observe cependant une sélection phonétique réelle dans les babillages significativement différente selon la langue et correspondant aux patterns de la langue cible.

On peut ainsi voir que l'acquisition des consonnes suit une progression au fil du temps, sachant que les enfants ont au départ des répertoires consonantiques similaires qui vont se différencier pour laisser apparaître les spécificités de chaque langue. Cependant, il convient de préciser qu'il existe des différences individuelles. Le développement de chaque enfant lui est propre. Ainsi, Vihman et Greenlee\* ont étudié la persistance des différences individuelles dans le développement phonologique de dix enfants normaux de un à trois ans. Les préférences à un an pour des catégories phonologiques particulières (fricatives, liquides, consonnes finales) ne correspondent pas à la maîtrise de ces catégories à trois ans. La forte utilisation de productions contenant de « vraies » consonnes est, selon ces auteurs, prédictive d'une bonne avancée phonologique à trois ans.

\*1987

Les enfants diffèrent selon l'intelligibilité, les substitutions de segments, les réductions et la proportion d'erreurs prosodiques et segmentales.

Dans une étude récente, Whalen et ses collègues\* analysent les babillages d'enfants français et anglais de 9 et 12 mois dans le but de voir l'influence de la langue maternelle sur les caractéristiques de voisement (Voice Onset Time = VOT) des consonnes, et l'apparition des différences en fonction du point d'articulation. Les auteurs présentent les résultats suivants :

\*2007

- effet significatif de la langue sur les VOT positifs,
- effet significatif de l'âge sur les VOT positifs (augmentation entre 9 et 12 mois de environ 13 ms pour le français, d'environ 4 pour l'anglais),
- effet significatif de la place d'articulation de la consonne sur le VOT,
- différence de VOT significative entre alvéolaire et vélaire,
- différence de VOT marginalement significative entre labiale et alvéolaire,
- effet significatif de la langue sur la fréquence des items prévoisés dans le babillage, donc une pertinence de la différence du taux de prévoisement entre français et anglais.

En effet, 44 % des occlusives sont voisées\* en français et 91 % des productions adultes ont un prévoisement. Les valeurs du VOT chez l'adulte varient en fonction de la place d'articulation\* : les labiales ont un VOT plus court et les alvéolaires ont un VOT plus court ou identique (selon les données) à celui des vélares.

\*Delattre, 1965

\*Nearey, Rochet, 1994

Les auteurs retrouvent donc dans leur étude les changements du VOT en fonction de la place d'articulation (conséquences automatiques de l'articulation). Selon eux, les labiales auraient un VOT plus court car le relâchement prend moins de temps.

### 3 - Le développement des productions langagières au niveau suprasegmental (prosodie)

La durée des segments et des pauses est une caractéristique importante de la prosodie d'une langue. Afin d'acquérir sa langue maternelle, l'enfant doit apprendre à produire les phonèmes, mais doit également apprendre les paramètres rythmiques de cette langue.

Vihman et ses collègues\* ont réalisé une étude sur plusieurs langues pour montrer que

\*2002

les enfants utilisent leurs connaissances des patterns segmentaux et de la prosodie de la langue de leur environnement pour produire du babillage ou des mots conformes à ces patterns. Ils ont pour ce faire enregistré du babillage dissyllabique et des mots identifiables produits par cinq enfants américains, français et écossais, de dix à dix-huit mois. Ils ont aussi enregistré des productions dissyllabiques d'adultes (mots et non-mots), productions choisies pour leurs ressemblances segmentales et phonotactiques avec celles des enfants. Le but de ce travail était d'étudier les patterns rythmiques (durées) dans ces trois langues, qui ont différents systèmes d'accentuation. Les consonnes étudiées sont des occlusives.

Les résultats montrent des différences significatives dans les durées des productions voyelle-consonne-voyelle (VCV) adultes : en anglais, la première voyelle (V1) est plus longue, en français la seconde voyelle (V2) est plus longue et en écossais, la consonne (C) est plus longue. Un seul enfant anglais, deux écossais et trois français produisent des éléments avec des durées correspondant à la langue adulte. Ces différences sont cependant non significatives. Vihman et ses collègues retrouvent par ailleurs un changement entre les deux périodes repères (10 vs. 18 mois). À 18 mois, les patterns sont plus proches de ceux de l'adulte. En anglais, V1 devient plus longue et C plus courte ; en français, V1 devient plus courte et V2 plus longue ; et enfin en écossais, V1 devient plus courte et C ainsi que V2 deviennent plus longues. Les enfants français et écossais sont plus proches du modèle adulte plus tôt que les enfants anglais et la différence est significative. En conclusion, les auteurs montrent que les enfants peuvent être assignés à la bonne langue sur la seule base du pattern rythmique. Vers 10/14 mois, ils ont l'allongement final. Les français semblent plus homogènes dans leurs productions, mais la différence n'est pas significative. Au fil du temps, les trois groupes se rapprochent du modèle adulte.

\*1992

De la même façon, Levitt et Utman\* ont enregistré un enfant français et un enfant américain à 5, 8, 11 et 14 mois. Au début, les babillages sont similaires (patterns universels), puis ils diffèrent pour rejoindre les patterns de la langue environnante. L'anglais produit plus de syllabes fermées et le français présente un allongement final significatif.

\*Hallé et coll., 1991

Une autre étude\* a analysé certains aspects prosodiques des productions dissyllabiques (babillage et mots) de cinq enfants français et cinq enfants japonais d'environ dix-huit mois. Hallé et ses collègues ont étudié les durées des voyelles. Les résultats montrent que ces durées sont clairement spécifiques à la langue maternelle de l'enfant.

\*Nathani et coll., 2003

D'autres auteurs encore\* ont voulu vérifier le pattern d'allongement final, pour voir si c'est un effet biologique (comportement inné) ou acquis par l'enfant. Ils ont enregistré huit enfants normaux d'un an et huit enfants sourds de quatre ans à plusieurs étapes du développement vocal prélinguistique : précanonique, canonique, post-canonique. Bien qu'ils constatent une variabilité individuelle considérable, les auteurs concluent que la surdité peut affecter de façon significative les aspects temporels de la production des enfants, ce qui laisserait supposer que l'allongement final est bien acquis par l'enfant.

\*1991

La fréquence fondamentale (Fo) est un autre paramètre fondamental de la prosodie.

Dans l'étude de Hallé et ses collègues\* citée précédemment, les contours de Fo sont également analysés. Les auteurs retrouvent pour les enfants français des contours Fo montants et un allongement final, et pour les enfants japonais des contours Fo descendants et une absence d'allongement final. Ces résultats montrent que les productions reflètent les patterns de Fo des adultes. Les productions bisyllabiques produites par les enfants japonais reflètent les formes adultes non seulement en terme d'intonation globale, mais aussi en terme de tons et de durée.

\*2000

Davis et ses collègues\* ont conduit une étude consistant en la comparaison des séquences dissyllabiques de quatre enfants normaux entre 7 et 12 mois et cinq adultes américains, portant sur Fo, l'intensité et la durée des syllabes accentuées et non accentuées. Le but était d'étudier les propriétés acoustiques des mots dissyllabiques perçus comme différemment accentués dans le babillage et d'explorer les contraintes physiologiques dues à la position des syllabes. Les résultats ne montrent pas de différence significa-

tive entre les trois corrélats acoustiques (Fo, durée, intensité) pour les items perçus uniformément accentués. Par contre, on observe des différences significatives pour les items asymétriquement accentués, entre syllabes accentuées et non accentuées pour les trois corrélats. Pour les items adultes, les différences sont significatives. Chez les enfants, la durée plus longue dans la deuxième voyelle varie en fonction de l'accent, mais pas de façon significative. La Fo des voyelles varie aussi en fonction de la position et de l'accent : la Fo est plus haute dans les voyelles initiales et dans les syllabes accentuées, significativement pour l'accent, mais pas pour la position. On observe également ces mêmes variations pour l'intensité, et les deux effets (position et accent) sont significatifs. Bien que l'étude ait été menée sur peu d'enfants, ces résultats apportent des informations intéressantes.

## LES ÉTUDES QUI ONT SERVI DE POINT DE DÉPART À NOTRE TRAVAIL

### POUR L'ANALYSE ACOUSTIQUE

Les deux études de Levitt et ses collègues\* sont le prolongement l'une de l'autre. Les auteurs ont enregistré cinq enfants français monolingues et cinq anglais monolingues de 7 à 11 mois environ. Elles comportent des productions de babillage contenant deux ou trois syllabes jugées identiques, ceci pour éliminer les éventuelles variations de Fo entre deux voyelles différentes (Fo intrinsèque). Ces stimuli ont été présentés à des juges experts en phonétique, natifs des USA, dont la tâche, dans la première étude\*, était la classification des contours d'intonation en montant, descendant, montant-descendant, descendant-montant et plat. L'analyse acoustique portait sur Fo.

Les résultats montrent que les contours sont classifiés dans 65 % des cas comme montant ou descendant. Pour ces contours, ceux des enfants français étaient partagés également entre montant et descendant, tandis que pour les enfants anglais, 75 % des contours étaient jugés descendants. Les différences entre les deux langues sont significatives et correspondent aux attentes par rapport à la langue de l'environnement. L'analyse de Fo confirme l'étude perceptive : il y a une bonne corrélation entre données acoustiques et perceptives. Les auteurs notent un effet perceptif intéressant : un grand changement d'amplitude a parfois été perçu comme un changement de Fo, alors que l'analyse acoustique ne montrait pas de changement de Fo. Ils ont alors effectué une analyse acoustique, sans comparaison perceptive, consistant en :

- une comparaison des deux premières syllabes des items,
- une comparaison des deux dernières,
- une comparaison des résultats selon les items de deux ou trois syllabes.

Les enfants anglais ont une Fo plus haute sur la première syllabe ; les enfants français ont, en moyenne, une Fo presque identique sur les deux premières syllabes, ou sur la première par rapport à la dernière, selon le nombre de syllabes. Pour les items de deux syllabes, les enfants français ont plus de contours montants et moins de contours descendants que les anglais. Cela reflète, selon les auteurs, les contours intonatifs des langues maternelles. Les auteurs notent cependant qu'il est possible de faire des confusions entre leurs résultats et les changements de Fo dus à l'état émotionnel de l'enfant.

La seconde étude\* prolonge la première et consiste, pour les mêmes stimuli, en une analyse :

- de l'allongement final de l'item,
- de la durée des syllabes non finales,
- du nombre de syllabes par items.

Les auteurs effectuent ainsi trois mesures rythmiques sur le même corpus. Les syllabes non finales en français adulte sont en général de longueur équivalente, plus que les syllabes non finales en anglais. L'hypothèse est donc que les enfants français montrent une

\*Whalen et coll., 1991 ; Levitt, Wang, 1991

\*Whalen et coll., 1991

\*Levitt, Wang, 1991

meilleure régularité dans la durée des syllabes non finales que les enfants anglais (effet de l'accent en anglais). Le nombre de syllabes des items peut varier entre français et anglais, à cause des variations de pattern accentuel entre les deux langues : il y a plus de mots longs (56 % de quatre syllabes ou plus) en français. D'où l'hypothèse que les enfants français produisent plus de mots longs. L'allongement de la syllabe finale est mesuré par comparaison entre la finale et la pénultième. Un allongement est compté quand il y a plus de 20 ms de différence. Les résultats montrent que 61,4 % des français ont un allongement final contre 32,1 % des anglais : la différence est significative. Les résultats concernant l'isosyllabité (durées homogènes des syllabes) ne montrent pas de différence significative, mais une tendance des enfants français à avoir des durées plus régulières que celles des enfants anglais. Enfin, concernant le nombre de syllabes par production, les résultats montrent que les enfants français produisent 43,8 % de mots de quatre syllabes ou plus, contre 18,9 % pour les enfants anglais (différence significative).

En conclusion, Levitt et ses collègues concluent à un effet des propriétés rythmiques des langues sur le babillage des enfants. La sensibilité des enfants aux informations prosodiques de leur langue peut diminuer vers 9-10 mois, âge où les enfants se concentrent sur les aspects segmentaux.

## POUR L'ANALYSE PERCEPTIVE

De Boysson-Bardies et ses collègues\* ont enregistré des enfants de huit mois tunisiens, cantonnais et français et ont demandé à des adultes français de reconnaître le babillage des bébés français. Pour les babillages à huit mois, ils ont obtenu 70 % de réponses correctes de la part des adultes. Cela suggère, selon ces auteurs, que le babillage comprend des caractéristiques d'intonation et de qualité vocale spécifiques à la langue maternelle des bébés.

## MÉTHODOLOGIE

### L'ANALYSE ACOUSTIQUE

Nous avons récupéré sur support informatique tous les enregistrements effectués par Levitt et ses collègues\*. Nous avons sélectionné, pour notre étude, les enregistrements à 9 mois et uniquement les items bissyllabiques. Un tri avec critères de rejet (présence de falsetto (voix de fausset, registre de voix le plus aigu empêchant la contraction normale des cordes vocales), vocalisations non segmentables, syllabe produite en aspiration, impossibilité d'extraire la Fo par le programme) nous a permis de garder 70 items chinois et 55 français.

Ces enregistrements ont été faits par les parents des enfants, à domicile, par intervalles d'une semaine. Chaque session d'enregistrement durait 10 ou 20 minutes. Les parents avaient pour consigne de choisir un moment où l'enfant était calme, non endormi et content. Dans la mesure du possible, le micro était situé à 20 cm de l'enfant. Les parents avaient l'autorisation d'induire le babillage en interagissant avec l'enfant, en se taisant dès qu'il commençait à babiller.

Nous avons choisi d'orienter ce travail vers l'étude de certaines caractéristiques prosodiques des voyelles. Nous nous sommes basés sur plusieurs critères pour les localiser dans les productions de babillage :

- l'énergie,
- le F2 de la voyelle (si possible),
- la courbe de variabilité spectrale (si exploitable).

Voici ci-dessous un exemple, où l'on peut voir tout d'abord le signal, la courbe d'énergie, la dérivée spectrale et la courbe de Fo.

\*1984

\*1991

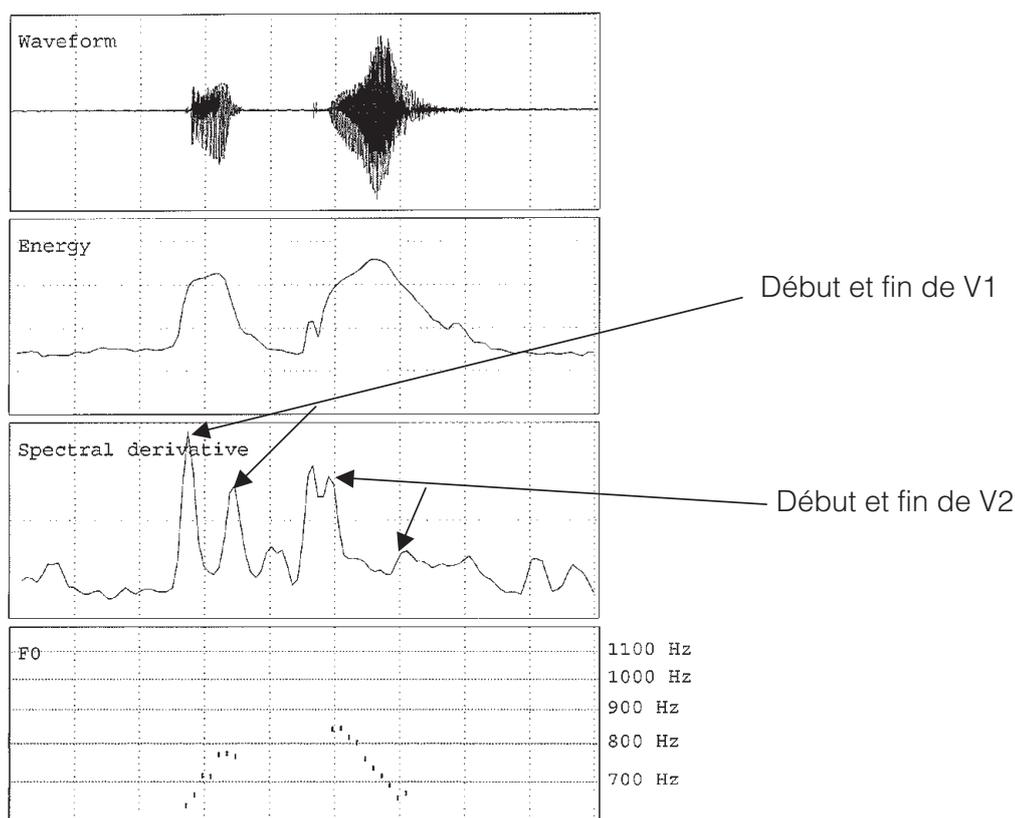


Figure 1 : Données acoustiques utilisées pour la sélection des énoncés.

Les mesures acoustiques ont été faites avec le programme Praat\* et un autre programme créé spécialement pour étudier les caractéristiques prosodiques d'énoncés\*.

Le script Praat reprend la segmentation manuelle, c'est-à-dire la segmentation en voyelles. Ce script calcule les données de Fo (Fo minimum, maximum, moyenne et déviations standard), la durée et l'intensité pour chaque voyelle.

Nous obtenons ainsi deux mesures prosodiques. Le **gradient** est calculé par la soustraction de la Fo de la deuxième voyelle et celle de la première, sur celle de la première. On peut l'écrire ainsi :  $(Fo(v2) - Fo(v1)) / Fo(v1) * 100$  (en %).

Le **ratio** des durées des voyelles est le rapport de la durée de la deuxième voyelle sur celle de la première. On peut l'écrire ainsi :  $lg(v2) / lg(v1)$ . Le script Praat fait donc les calculs sur deux voyelles distinctes.

Le risque possible de ces mesures concerne les petites différences de la Fo de l'énoncé, car l'on peut obtenir des contours globalement plats. C'est pourquoi nous avons voulu associer à ces mesures celles faites par le programme spécifique.

Le programme spécifique considère le contour de l'énoncé comme un tout et effectue un calcul de Fo (avec corrections manuelles éventuelles) moyenne, minimum et maximum, un calcul de « l'**excursion** »  $((Fo \text{ extremum } 2 - Fo \text{ extremum } 1) / Fo \text{ début}) * 100$  (en %) et un ajustement d'une courbe Fo continue. Une excursion positive montre une hausse de Fo dans le contour de l'énoncé et une négative indique une baisse.

Nous obtenons ici également deux mesures : un **gradient** calculé de la façon suivante :  $(\text{valeur Fo finale} - \text{valeur Fo initiale}) / \text{valeur initiale}$ . Le **ratio** est calculé de la même façon qu'avec Praat, mais ne peut être ici calculé que si deux sections voisées sont trouvées par le programme. Cette mesure est automatique et nous n'avons donc pas systématiquement un ratio par énoncé avec ce programme. Le gradient mesuré par ce programme n'est pas le même que celui calculé par Praat et permet de voir plus de mouvements de la Fo.

Les deux figures suivantes présentent les données obtenues grâce aux deux programmes. La première figure (programme spécifique) montre en première partie le spectrogramme d'un énoncé français, en deuxième partie le signal sonore et en troisième partie (poin-

\*Boersma, Weenink, 1992-2007

\*Hallé et coll., 1991

tillés) le contour  $F_0$ . Sur ce contour, on peut voir une courbe qui représente la ligne obtenue par lissage afin de faire le lien entre les deux sections.

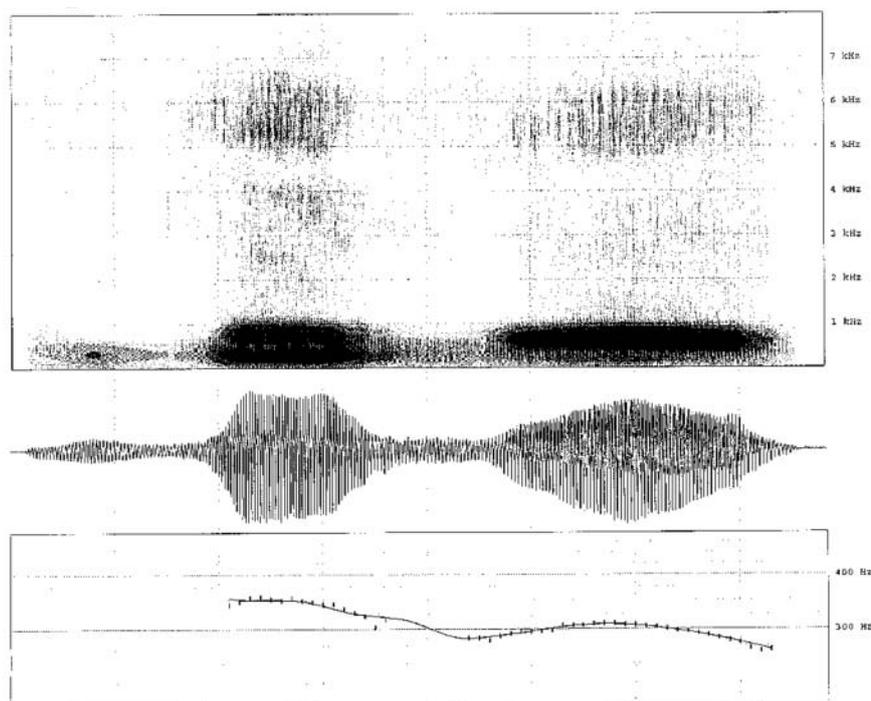


Figure 2 : Données acoustiques pour un babillage français.

La seconde figure (Praat) présente un énoncé français, avec en première partie le signal sonore et en deuxième partie le spectrogramme. Sous le spectrogramme se trouve tout d'abord la segmentation manuelle des syllabes et sous cette ligne se trouve la segmentation manuelle des voyelles. C'est sur cette segmentation manuelle que s'appuient les calculs effectués par Praat.

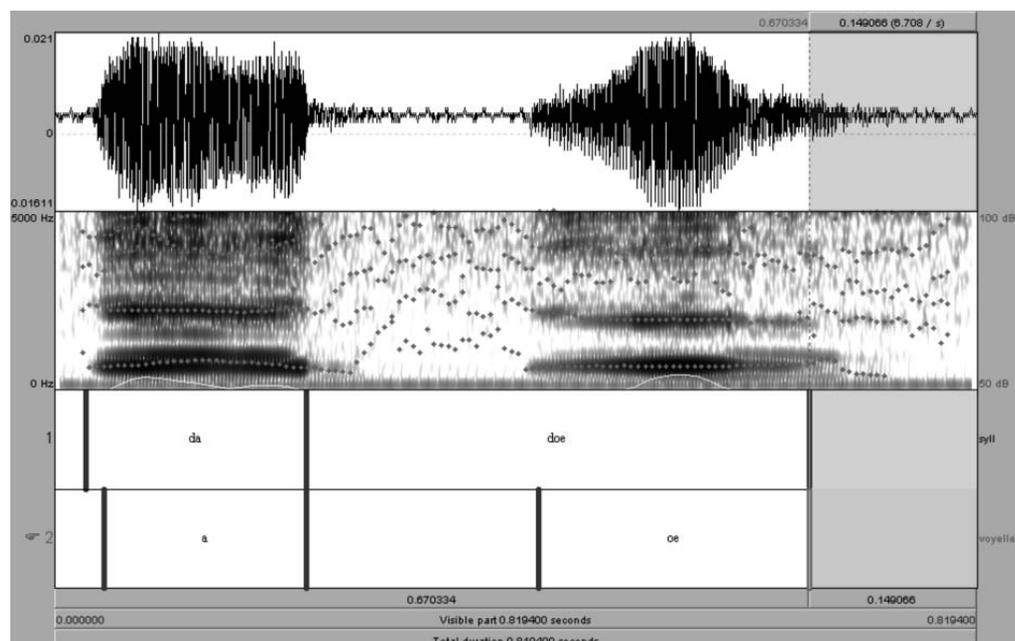


Figure 3 : Données acoustiques pour un babillage français avec Praat.

Après cette analyse acoustique, nous nous sommes posés la question de savoir de quel genre de productions était composé notre corpus de babillages. Nous avons donc effectué un comptage et un classement des productions selon leur « type segmental » (conson-

ne/voyelle, voyelle/consonne/voyelle, voyelle/voyelle, etc.). Lors de ce comptage, nous avons été confrontés pour les enfants chinois à la présence fréquente de /h/ aspirés ou de coups de glotte. La classification de ces phonèmes pouvant être litigieuse, nous avons choisi de ne pas en tenir compte. Nous avons ainsi pris en compte dans cette analyse les 55 items français et les 45 chinois.

## L'ANALYSE PERCEPTIVE

Pour monter notre test de perception, nous avons décidé de sélectionner dix énoncés (stimuli) de chaque langue. Nous avons donc retenu parmi les énoncés chinois et français triés pour l'analyse acoustique dix items caractéristiques (proches des moyennes mesurées dans l'analyse acoustique ; c'est-à-dire allongement final et baisse finale de  $F_0$  pour le français et contour  $F_0$  descendant pour le chinois). Pour les items anglais, nous avons sélectionné 10 items dissyllabiques clairement articulés parmi les items enregistrés par Levitt et collègues.

Nous avons ensuite construit des paires français/non-français avec soit anglais soit chinois. Pour chaque couple de langues, nous avons construit 30 paires de stimuli. Chaque stimulus français est associé à trois chinois ou à trois anglais et chaque chinois ou chaque anglais à trois français différents. Pour chaque paire, le stimulus français apparaît soit en premier soit en second de façon aléatoire, mais en sorte que la moitié des paires soit français/non français et l'autre moitié non français/français. Soixante paires tests ont été ainsi construites. Trois autres paires « d'entraînement » (qui ne réapparaissent pas dans le test) ont été construites, ce qui fait au total 63 items.

Nous avons, pour finir, créé deux ordres pseudo-aléatoires de passation : la contrainte était de ne jamais avoir deux paires de même composition (exemple : français/chinois) à la suite. Les trois premières paires d'entraînement sont les mêmes pour les deux ordres. Ces deux ordres ont été présentés aux juges de façon équivalente en alternance.

L'analyse perceptive se base sur le principe décrit par De Boysson-Bardies, Sagart et Durand\* : à chaque "essai", on présente aux juges français deux énoncés dont l'un est du babillage d'enfant français ; ils doivent indiquer lequel des deux est le babillage français. La tâche des sujets dans notre étude est la même : entre deux énoncés entendus, ils doivent choisir celui qui selon eux est le babillage français ; ils doivent de plus donner un degré de confiance de leur réponse allant de 1 (pas sûr) à 5 (très sûr).

La passation du test de perception s'effectue avec le programme Praat, de façon individuelle. Une pause est insérée toutes les 15 paires. Le juge peut alors se reposer comme il le souhaite. Chaque juge reçoit les mêmes consignes et explications. La passation du test dure environ 20 minutes, les juges n'ont pas le droit de réécouter les items.

Avant et après le test, nous avons fait remplir aux juges deux questionnaires.

Les juges sélectionnés sont des juges dits « experts » (orthophonistes) et des juges « naïfs » de langue française. Nous avons proposé aux sujets un questionnaire pré-test, comportant des indications, telles que l'âge, le niveau d'études, la profession, la langue maternelle, les langues étrangères parlées, la pratique musicale, etc. Nous avons également proposé un questionnaire post-test afin de recueillir leurs impressions.

\*1984

## LES RÉSULTATS

### RÉSULTATS DE L'ANALYSE ACOUSTIQUE.

Nous avons tout d'abord recueilli les données brutes grâce aux deux programmes cités précédemment. Ensuite, nous avons effectué les mesures de l'excursion, du gradient et du ratio pour les Chinois et les Français.

Tableau 1 : Résultats des mesures comparées pour les énoncés chinois.

CH	Fo min	Fo max	eng max	gradient	excursion	ratio
Praat	258	348	66,8	-6,9		1,12
pr. Spé	249	322	65,4	-8,6	-13,8	1,05

Tableau 2 : Résultats des mesures comparées pour les énoncés français.

FR	Fo min	Fo max	eng max	gradient	excursion	ratio
Praat	306	393	70,1	-1,9		1,56
pr. Spé	329	436	69,7	-1,6	-2,1	1,49

Afin d'illustrer la comparaison des ratios et gradients (calculés par Praat) entre les Français et les Chinois, voici deux graphiques.

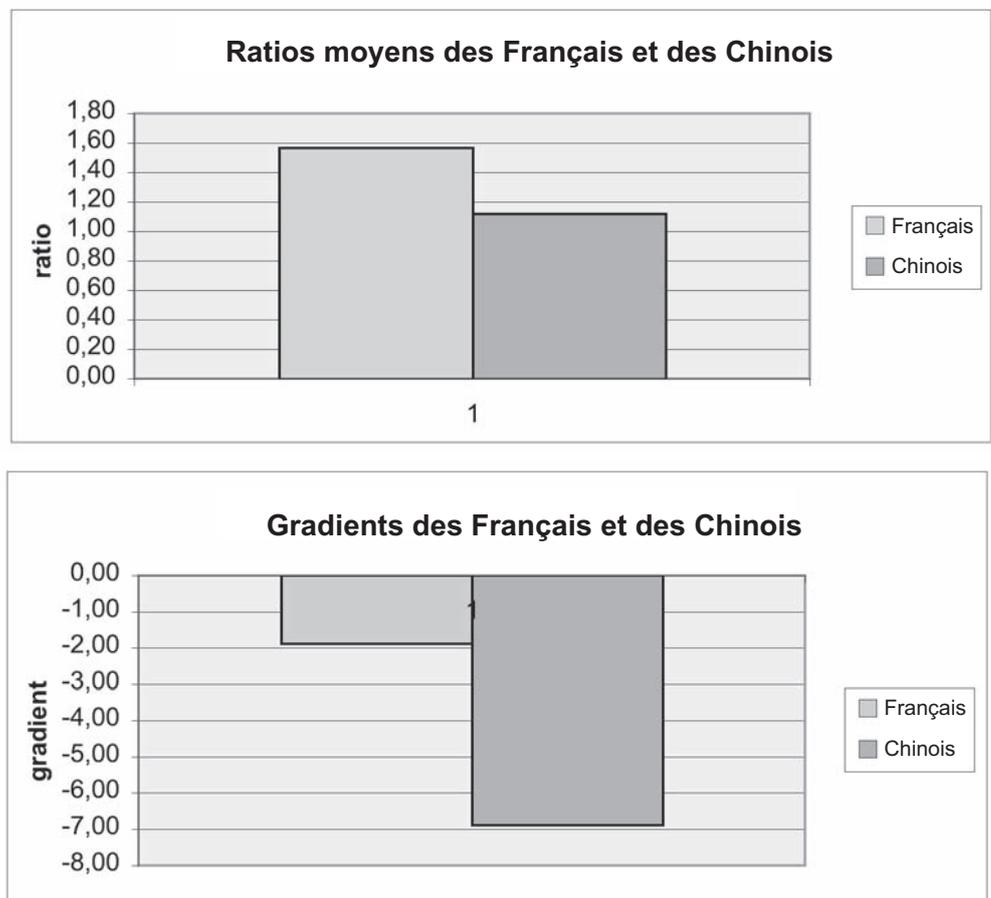


Figure 4: Ratios et gradients moyens des Français et des Chinois.

- Les analyses statistiques menées (test *t* de Student) montrent les résultats suivants :
- une différence significative entre les énoncés chinois et français en ce qui concerne l'excursion ( $p < .05$ ),
  - une différence non significative en ce qui concerne la différence d'énergie ( $p = .104$ ),
  - une différence non significative en ce qui concerne le ratio calculé par le programme spécifique (30 énoncés français sur 55 et 30 énoncés chinois sur 70) ( $p = .21$ ),
  - une différence marginalement significative concernant le ratio calculé par Praat ( $t(123) = 1,96$ ;  $p = .053$ ),

- une différence significative concernant le gradient des deux programmes ( $t(123) = 2,27; p < .05$ ).

L'ensemble de ces résultats montre une baisse de  $F_0$  significativement plus importante pour le chinois que pour le français, dont le contour semble plutôt plat (-6,9 pour le chinois contre -1,9 pour le français). L'excursion montre des contours significativement plus descendants pour les Chinois et le gradient montre que la  $F_0$  de la deuxième voyelle est significativement plus basse. La différence de  $F_0$  entre les énoncés chinois et français semble donc bien présente. Ces résultats montrent également une différence d'allongement de la deuxième voyelle marginalement significative entre le français (1,56) et le chinois (1,12), en faveur d'un allongement plus prononcé pour le français.

L'analyse segmentale qualitative nous a permis d'observer que la majorité des productions des bébés français est de type CVCV (47,3 %, contre 21,8 % de productions VV), alors que celle des chinois est de type voyelles seules (33,3 %, contre 15,5 % de productions CVCV). Les consonnes les plus fréquemment observées dans les productions CVCV des français sont des vélaires et des dentales (avec une prédominance de l'occlusive voisée /d/). Par ailleurs, les deux langues ont un pourcentage quasi-identique de productions VCV.

Les deux graphiques ci-dessous illustrent ces distributions selon le type de productions des Français puis des Chinois. On peut ainsi constater les différences, dans chaque type de productions, mais également au niveau de la quantité de types de productions présents.

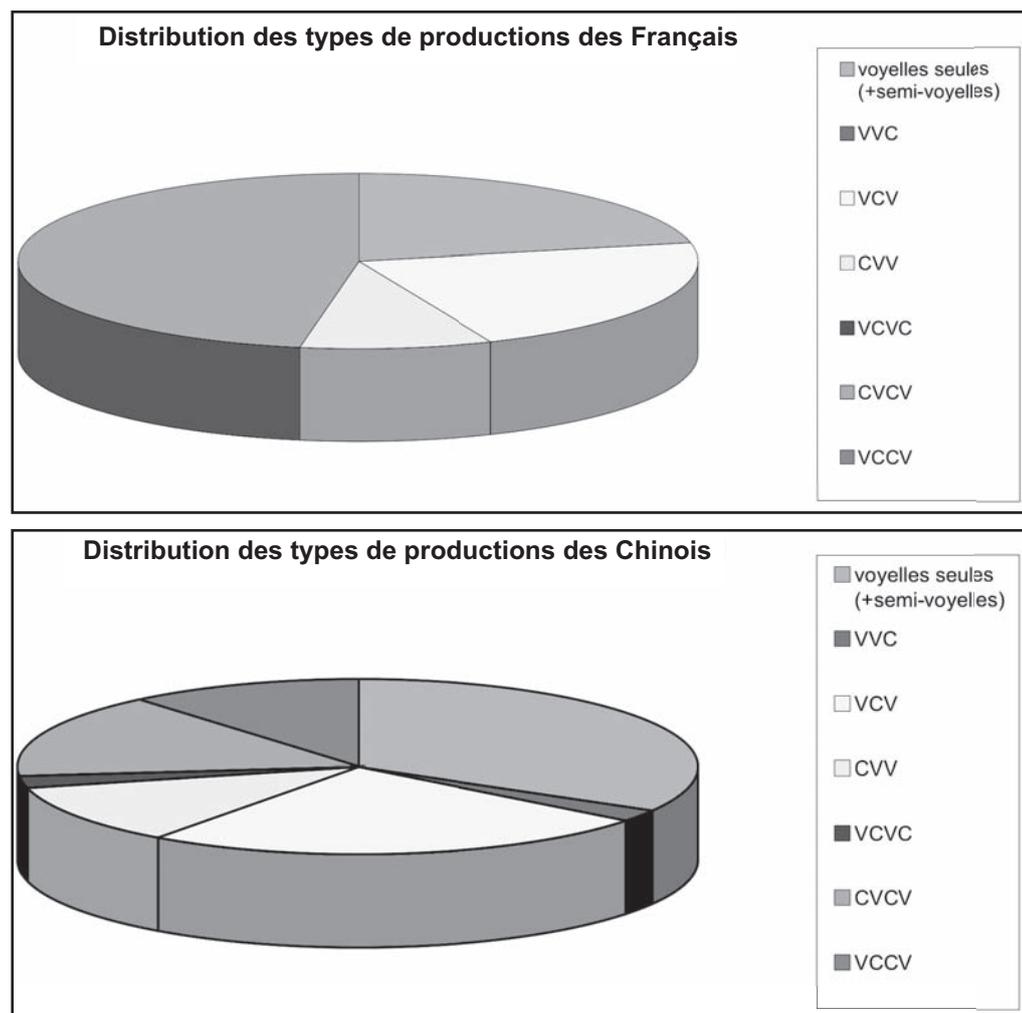


Figure 5 : Distribution des types de productions des Français et des Chinois.

Ces observations peuvent également suggérer une attention des bébés chinois et français différente aux différents paramètres de leur langue maternelle. Il est possible que les chinois soient plus sensibles aux paramètres suprasegmentaux et les français plus sensibles aux paramètres segmentaux. Nous avons donc regardé de plus près les caractéristiques des productions CVCV et des productions vocaliques de chaque langue.

Nous faisons les observations suivantes : les productions CVCV des Français ont une importante baisse moyenne de Fo (gradient moyen -5 %), avec un allongement final moyen modéré (1,08), tandis que les productions vocaliques n'ont pas de baisse de Fo (gradient moyen 1 %) et un allongement final modéré (ratio moyen de 1,1).

Nous avons également observé que les moyennes pour les productions vocaliques des Chinois sont très différentes de celles des Français, notamment pour la durée de la deuxième voyelle par rapport à la première (ratio moyen 0,2). Par ailleurs, les Chinois ont une baisse de Fo plus importante que les Français (gradient moyen -8,55 %). En ce qui concerne les productions CVCV des Chinois, nous observons là aussi une baisse importante de Fo (gradient moyen -9,8 %) et un allongement modéré (ratio moyen de 1,1).

Les analyses statistiques (test *t* de Student) ne sont pas significatives, ni pour la comparaison CVCV/VV dans les productions françaises, ni dans les productions chinoises.

Les graphiques suivants illustrent les tendances cependant observées. En français : les CVCV ont un contour tombant, pas les VV ; les ratios de durée sont analogues pour CVCV et VV. Quant au chinois, les contours sont également tombants pour CVCV et VV mais les ratios de durée sont différents (raccourcissement pour les VV). Il est possible que, d'une façon générale, les VV soient plus caractéristiques de la prosodie du modèle adulte en chinois.

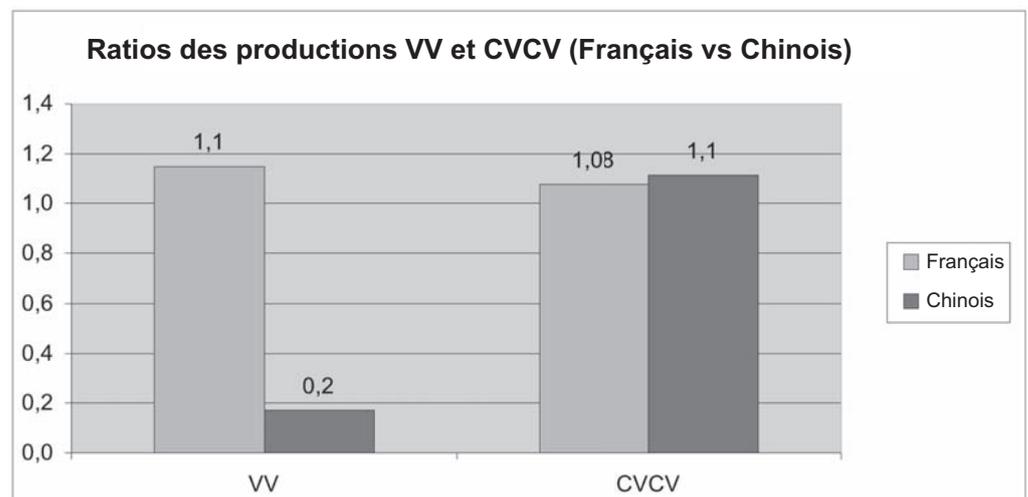
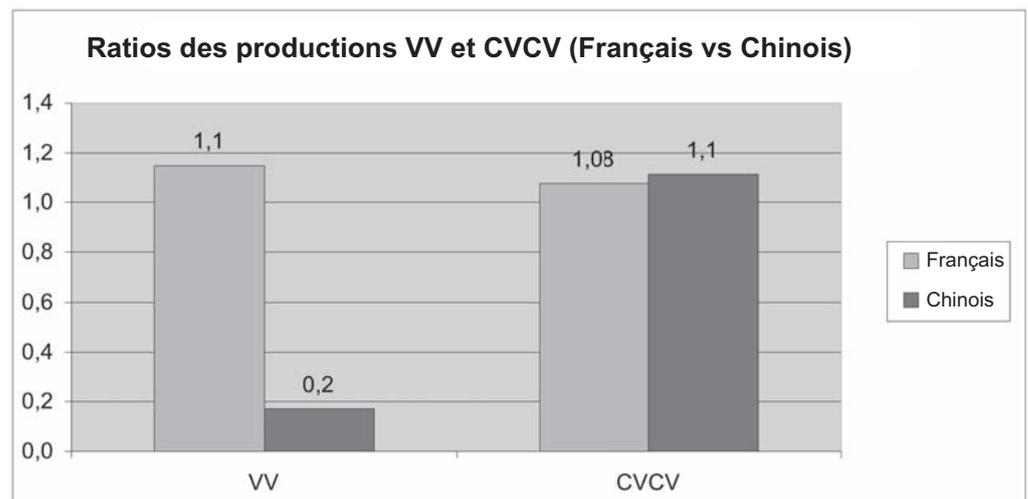


Figure 6 : Ratios et gradients moyens des productions VV et CVCV des Français et des Chinois.

## RÉSULTATS DE L'ANALYSE PERCEPTIVE

Les données brutes par sujet (moyennées à travers les paires tests) sont les pourcentages de réponses correctes, les notes moyennes de confiance pour les réponses correctes et celles pour les réponses incorrectes. Un pourcentage de réponses correctes de 50 % correspond au hasard. Le tableau et le graphique ci-dessous illustrent les données de pourcentage de réponses correctes et les notes moyennes de confiance pour les bonnes réponses et pour les réponses fausses.

Tableau 3 : Pourcentages de réussite et degrés de confiance selon les paires français/anglais ou français/chinois.

	FR / CH			FR / ANG		
	% correct	note OK moyenne	note fausse moyenne	% correct	note OK moyenne	note fausse moyenne
<b>naïfs</b>	52,4	2,53	2,63	59,7	2,79	2,47
<b>experts</b>	51,2	2,28	2,31	54,0	2,48	2,30

On observe que les naïfs semblent meilleurs que les experts dans l'identification du français et les résultats semblent meilleurs pour les paires français/anglais. La note moyenne de confiance est plus grande pour les bonnes réponses dans les paires fr/ang. La note moyenne est plus haute dans les réponses fausses pour les paires fr/ch.

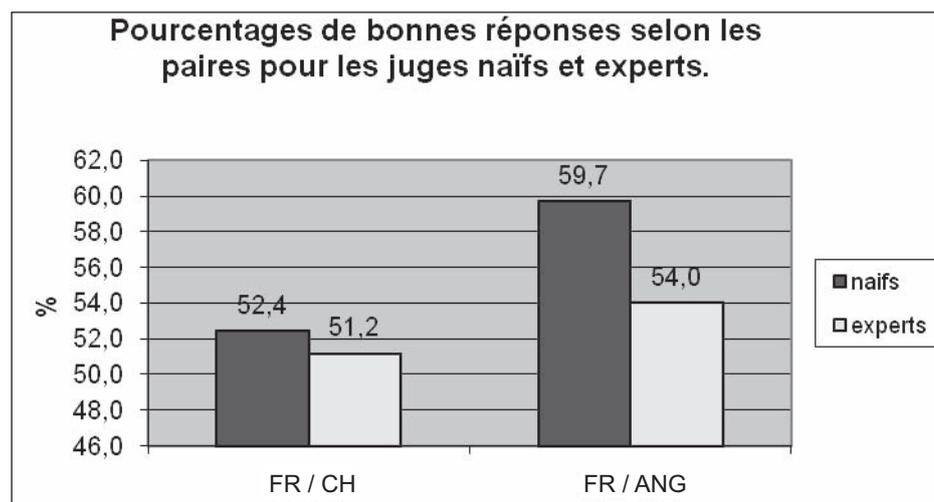


Figure 7 : Pourcentages de réussite selon les juges et les paires.

Nous avons conduit une analyse statistique par sujet. L'analyse de variance par sujet ne donne aucun effet significatif.

Le test *t* de Student donne un résultat marginalement significatif ( $p = 0,0507$ ) chez les naïfs pour les paires fr/ang.

Nous avons ensuite effectué une analyse statistique par item. Tout d'abord, nous avons réalisé une analyse de variance, qui a donné les résultats suivants :

- pas d'effet significatif du type de paire (fr/ch ou fr/ang),
- pas d'effet significatif du groupe (naïf ou expert),
- effet marginalement significatif ( $p = 0,054$ ) pour fr/ang chez les naïfs.

Puis, nous avons effectué un test *t* de Student par item, qui a donné un résultat significatif ( $p < 05$ ) pour les paires fr/ang chez les naïfs.

En conclusion de cette analyse perceptive, on peut voir que les juges naïfs semblent significativement capables d'identifier le bébé français dans une paire fr/ang.

L'ensemble des résultats nous permet d'observer la même tendance que dans notre pre-

mière étude (2005-2006) : il semble plus facile d'identifier le français dans les paires français/anglais, pour les experts comme pour les naïfs.

Une analyse qualitative des énoncés français les mieux identifiés révèle la présence d'un allongement de la deuxième voyelle par rapport à la première ou d'une baisse de Fo sur la deuxième voyelle. D'autre part, on retrouve dans ces énoncés la structure CVCV majoritairement représentée dans l'ensemble des énoncés français. Il semblerait que la structure de l'énoncé joue également, probablement en complément des caractéristiques prosodiques, dans l'identification d'un énoncé français par des juges.

Cependant, nous avons remarqué une grande variabilité des réponses, ce qui semble indiquer une difficulté à distinguer des productions bisyllabiques. Ces productions étant plutôt courtes, on peut imaginer que les sujets n'avaient pas toujours assez d'informations pour identifier le bébé français, d'autant plus qu'ils n'avaient pas la possibilité de réécouter.

Il serait intéressant de monter un test avec une tâche moins difficile, par exemple en permettant plusieurs écoutes ou en sélectionnant des énoncés plus longs.

Ces deux analyses, perceptive et acoustique, nous permettent d'effectuer certaines corrélations. Nous avons obtenu grâce au test de perception des résultats significatifs lors de l'analyse de variance par item pour les paires français/anglais chez les naïfs, au test *t* pour les paires français/anglais, toujours chez les naïfs, ainsi qu'un résultat marginalement significatif au test *t* (dans l'analyse par sujet) pour les paires français/anglais, chez les naïfs. En regardant de plus près les cinq énoncés les mieux identifiés par les juges, nous avons observé la présence d'un allongement final et/ou d'une baisse de Fo, ainsi que la structure CVCV présente dans 47,3 % des énoncés français.

Par ailleurs, l'analyse acoustique révèle des résultats qui suivent la même tendance : on observe une baisse de Fo significativement plus importante pour le chinois que pour le français, dont le contour semble plutôt plat (-6,9 pour le chinois contre -1,9 pour le français), ainsi qu'un allongement de la deuxième voyelle qui semble plus important (marginale-ment significatif) pour le français (1,56 fois plus longue que la première) que pour le chinois (1,12 fois plus longue).

## DISCUSSION

Nos études acoustique et perceptive nous ont permis de faire un certain nombre de constatations. Voici un récapitulatif de nos observations :

- une baisse de Fo significativement plus importante pour le chinois que pour le français, dont le contour semble plutôt plat (-6,9 pour le chinois contre -1,9 pour le français) ;
- un allongement de la deuxième voyelle qui semble plus important (résultat marginalement significatif) pour le français (1,56 fois plus longue que la première) que pour le chinois (1,12 fois plus longue) ;
- concernant le test de perception, les naïfs semblent meilleurs que les experts dans l'identification du français et les résultats semblent meilleurs pour les paires français/anglais ;
- l'analyse qualitative des items français les mieux identifiés révèle la présence d'un allongement de la deuxième voyelle par rapport à la première ou d'une baisse de Fo sur la deuxième voyelle, voire de la combinaison des deux ;
- la majorité des productions des bébés français (47,30 %) est de type CVCV, alors que celle des Chinois est de type voyelles seules ;
- les productions CVCV des Français ont une importante baisse moyenne de Fo (avec une moyenne de -5), avec un allongement final moindre (1,08), alors que les productions VV ont une baisse quasi inexistante (gradient moyen de 1) et un allongement également moindre (ratio moyen de 1,1). Ces différences ne sont pas statistiquement significatives ;
- les caractéristiques des productions VV des Chinois sont très différentes de celles des Français, notamment pour la durée de la deuxième voyelle par rapport à la première ;

- les productions VV des Chinois ont une baisse de Fo encore plus importante que les Français (-8,55), avec un allongement inexistant (0,2). Les productions CVCV montrent également une baisse de Fo (-9,8) et peu d'allongement final (ratio moyen de 1,1). Ces différences ne sont pas statistiquement significatives.

En reprenant les résultats décrits dans la littérature, on note que Levitt et ses collègues\* trouvent pour les Français un allongement de 61,4 % plus important que pour les Anglais (32,1 %).

Selon Levitt et ses collègues, ces résultats montrent que le babillage reflète les propriétés rythmiques de la langue maternelle. Nos résultats sont plus nuancés et semblent aller dans le sens d'une complémentarité baisse de Fo/allongement final/structure CVCV.

En ce qui concerne les résultats de Fo, Levitt et ses collègues trouvent dans leur étude que les enfants français ont plus de contours hauts et moins de contours bas que les Anglais dans les items de deux syllabes. Cela reflète selon eux les contours intonatifs majeurs de groupe prosodique, en particulier l'intonation de continuation\* des deux langues cibles. Nos résultats rejoignent leurs conclusions sur ce point. Notre comparaison acoustique des mesures de Fo entre français et chinois montre que la baisse finale de Fo est significativement plus importante chez les Chinois. Il semblerait donc, selon nos résultats, que les contours Fo soient plus marqués par une dominance des contours tombants en chinois qu'en français, ce qui rejoint l'idée de Levitt et ses collègues selon laquelle le babillage reflète les contours intonatifs des langues.

Par ailleurs, nous ne retrouvons pas une majorité de contours montants pour les Français, comme Levitt et ses collègues. Nos résultats semblent montrer des contours plutôt plats en moyenne. Il est important de noter cependant que notre étude a porté uniquement sur des items bisyllabiques, tandis que les études de Levitt et collègues ont porté sur des énoncés bi- et trisyllabiques. Cette différence pourrait expliquer qu'ils trouvent des contours plutôt montants, les variations de Fo étant plus importantes dans des énoncés plus longs.

Nos résultats concernant les mesures de Fo suggèrent ainsi que le babillage reflète les contours intonatifs de la langue maternelle. Dans la littérature, les résultats sont similaires : Hallé et ses collègues\* trouvent des contours Fo clairement montants pour le français.

Nos résultats acoustiques sont moins concluants pour ce qui est de l'allongement final de la deuxième syllabe en français. Vihman et ses collègues\* retrouvent dans leur étude que les enfants français ont la deuxième voyelle significativement plus longue dans des énoncés bisyllabiques et que les productions des enfants français sont plus homogènes.

Levitt et Utman\* retrouvent également l'allongement final chez les enfants français, allongement qui apparaît après une période où les babillages, quelle que soit la langue, présentent des patterns universels. En effet, tous les enfants ont au départ cet allongement final, mais il devient peu à peu plus marqué en français. L'étude de Nathani et ses collègues\* montre que ce pattern d'allongement final est bien acquis par l'enfant car les enfants sourds ne le produisent pas.

Nos résultats perceptifs ne nous permettent pas de conclure de façon tranchée sur les capacités d'identification des juges adultes. Nous retrouvons des pourcentages de réussite guère supérieurs à 50 % pour les paires fr/ch (52,4 % pour les naïfs et 51,2 % pour les experts) et un peu plus hauts pour les paires fr/ang (59,7 % pour les naïfs et 54 % pour les experts). Les résultats ne sont significatifs ou marginalement significatifs que pour les paires fr/ang chez les juges naïfs. Nos résultats montrent également un meilleur score moyen chez les juges naïfs. En regardant dans le détail et avec les réponses aux questionnaires, on remarque que les résultats des juges experts sont hétérogènes et que le meilleur juge est un juge expert. Il est intéressant de remarquer que parmi les sept meilleurs juges (ceux qui ont eu les meilleurs scores de bonnes réponses) :

- 6 ont une pratique musicale,
- 6 ont fréquenté durant une période de leur vie des personnes de langue étrangère,
- 1 est de langue maternelle autre que français (mais fait partie des juges experts),
- 2 présentent une baisse auditive connue,

\*1991

\*cf. aussi Hallé et coll., 1991

\*1991

\*2002

\*1992

\*2003

- 5 ont un niveau d'études supérieur ou égal à Bac +3,
- tous sont droitiers.

Nos résultats sont ainsi moins tranchés que ceux de De Boysson-Bardies et ses collègues\*, qui trouvaient 70 % de bonnes réponses.

Il semblerait ainsi que l'association avec une langue différente au niveau prosodique mais relativement familière à des Français, comme l'anglais, occasionne moins d'erreurs qu'une langue moins familière, comme le chinois. Il serait intéressant de confronter l'anglais et le chinois dans un test de perception de même principe. L'association du français avec le chinois, langue très différente au niveau prosodique, ne permet pas aux juges de faire la différence de façon significative, comme initialement supposé. Peut-être qu'une étude comparant français et chinois avec des babillages d'enfants plus âgés changerait les résultats.

De plus, une étude acoustique plus approfondie des énoncés anglais permettrait d'analyser les différences et similitudes supra-segmentales présentes dans les babillages. Les questionnaires post-test remplis par les juges ont révélé une impression de grande difficulté à répondre, avec des énoncés trop courts. Le fait de ne pas réécouter a beaucoup gêné les juges. Un test plus simple et plus souple permettrait sans doute d'obtenir des résultats plus proches de ceux décrits dans la littérature.

Cependant, l'analyse des énoncés français ayant le plus fort pourcentage de réussite montre que ceux-ci présentent les caractéristiques acoustiques attendues. Une explication à ces résultats serait l'âge des sujets : les enfants choisis pour cette étude avaient neuf mois, c'est-à-dire qu'ils seraient au début de la phase de babillage canonique décrite dans la littérature. En effet, chaque enfant a son propre rythme de développement, et il se peut que les enfants sélectionnés soient entrés de façon variable dans le babillage canonique. Ceci expliquerait que les propriétés supra-segmentales de la langue maternelle ne soient pas à cet âge suffisamment prégnantes pour que les juges les identifient. Cependant, si l'on considère les productions des bébés français, on constate une prédominance du babillage canonique (majorité observée de CVCV) dans les productions à 9 mois. Ceci semble conforter les données de la littérature, selon lesquelles le babillage canonique apparaît vers cet âge. Les pourcentages de productions vocaliques et de VCV uniquement peuvent suggérer qu'à cet âge les bébés sont dans la transition entre prébabillage et babillage.

L'analyse qualitative des types de productions chez les Français et les Chinois montre d'importantes différences, tant dans le pourcentage de productions les plus fréquentes que dans la variabilité des types de productions. En effet, les Français ont moins de variabilité dans leurs productions, et les productions CVCV sont majoritairement représentées. A l'inverse, les Chinois présentent une grande variabilité dans leurs productions, avec une prédominance des productions VV.

Su Tzu Ting\* a montré dans son travail de thèse une réduction et une disparition plus importante des consonnes dans la parole spontanée en chinois mandarin de Tai Wan qu'en français, ainsi qu'une réduction et disparition des voyelles moins importantes en chinois qu'en français. Selon l'auteur, cela indique une plus grande vulnérabilité des consonnes en chinois mandarin, probablement due au fait que c'est une langue à tons, monosyllabique, dans laquelle la voyelle, porteuse de ton et souvent unique noyau syllabique du mot, est plus protégée.

Il serait intéressant de conduire une étude sur un plus grand nombre d'énoncés et à plusieurs âges (6, 9 et 12 mois), afin de voir la progression de l'apparition du babillage dit canonique.

Enfin, en conclusion de ce travail, je tiens à remercier particulièrement Pierre Hallé, sans qui cela n'aurait pas été réalisable.

- BOERSMA, P., WEENINK, D. (1992-2007). Praat: a system for doing phonetics by computer (version 5.0.38). [Computer program]. <http://www.praat.org>.
- DAVIS, B.L., MACNEILAGE, P.F., MATYEAR, C.L., POWELL, J.K. (2000). Prosodic correlates of stress in babbling: an acoustical study. *Child Development*, 71 (5), 1258-1270.
- DE BOYSSON-BARDIES, B. (1996, Édition révisée en 2005). *Comment la parole vient aux enfants ?* Paris : Éditions Odile Jacob.
- DE BOYSSON-BARDIES, B., VIHMAN, M.M. (1991). Adaptation to language: evidence from babbling and first words in four languages. *Language*, 67, 297-319.
- DE BOYSSON-BARDIES, B., HALLÉ, P., SAGART, L., DURAND, C. (1989). A cross-linguistic investigation of vowel formants in babbling. *Journal of Child Language*, 16, 1-17.
- DE BOYSSON-BARDIES, B., SAGART, L., DURAND, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of Child Language*, 11, 1-15.
- DECASPER, A.J., FIFER, W.P. (1980). Of human bonding: newborns prefer their mothers' voices. *Science*, 208 (4448), 1174-1176.
- DELATTRE, P. (1965). *Comparing the Phonetics Features of English, German, Spanish and French: an interim report*. New York, Chilton, Heidelberg: Julius Groos Verlag.
- ESLING, J. (2006, Mai). La production de la parole chez le nouveau-né. *Journée du pré-babillage au babillage: aspects physiologiques, phonétiques, prosodiques*. Paris.
- ESLING, J.H., BENNER, A., BETTANY L., ZEROUAL, C. (2004). Le contrôle articulatoire phonétique dans le pré-babillage. *Actes des XXVes Journées d'étude sur la parole, 19 avril, Fez, Maroc*. Aix en Provence: association francophone de la communication parlée.
- ESLING, J.H., BETTANY, L.D., BENNER, A., ROSE, S. (2005). *Phonetic structure and acquisition of laryngeal and pharyngeal articulations: Text-Analysis considerations*. CHWP A. 31 août: editors of Computing Humanities Working Papers 2005. <http://www.chass.utoronto.ca/epe/chwp/>
- HALLÉ, P. (2004). Acquisition du langage: spécialisation des enfants dans leur langue maternelle. *Actes du colloque MIDL, 29-30 novembre*. Paris. [http://www.limsi.fr/MIDL/actes/conferenceinviteeIII/Halle\\_MIDL2004.pdf](http://www.limsi.fr/MIDL/actes/conferenceinviteeIII/Halle_MIDL2004.pdf)
- HALLÉ, P.A., DE BOYSSON-BARDIES, B., VIHMAN, M. (1991). Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants. *Language and Speech*, 34, 299-318.
- LEVITT, A.G., WANG, Q. (1991). Evidence for language specific rhythmic influences in the reduplicative babbling of French- and English-learning infants. *Language and Speech*, 34 (3), 235-249.
- LEVITT, A.G., UTMAN, J.G. (1992). From babbling towards the sound systems of English and French: a longitudinal two-case study. *Journal of Child Language*, 19, 19-49.
- LOCKE, J.L. (1983). *Phonological acquisition and change*. New York: Academic Press.
- MCCUNE, L., VIHMAN, M.M. (2001). Early phonetic and lexical development: a productivity approach. *Journal of speech, language and hearing research*, 44, 670-84.
- MACNEILAGE, P.F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21, 499-511. <ftp://ftp.princeton.edu/pub/harnad/BBS/WWW/bbs.macneilage.html>
- NATHANI, S., OLLER, K., COBO-LEWIS, A.B. (2003). Final syllable lengthening (FSL) in infant vocalizations. *Journal of child language*, 30, 3-25.
- NEARY, T.M., ROCHET, B.L. (1994). Effects of place of articulation and vowel context on VOT production and perception in French and English stops. *Journal of the international phonetic association*, 24, 1-19.
- OLLER, K., BUDER, E.H. (2005). Evolutionary implications of vocal development in the first half-year of life. *Colloque ELA (Emergence of Linguistic Abilities), 8-20 décembre*. Lyon: CNRS et Université Lyon2, Laboratoire dynamique du langage.
- OLLER, K., EILERS, R.E., URBANO, R., COBO-LEWIS, A.B. (1997). Development of precursors to speech in infants exposed to two languages. *Journal of child language*, 24 (2), 407-425.
- OLLER, D.K. (1980). The emergence of the sounds of speech in infancy; Chapter 6. In YENI-KOMSHIAN, G., KAVANAGH, J.F., FERGUSON, C.A. (eds), *Child Phonology: Vol.1. Production*, 93-112. New York: Academic Press.
- SU, T.T. (2003). *Etude sur la disparition des phonèmes dans la parole spontanée du français et du mandarin de Taó Wan*. Doctorat de phonétique: Université Paris III à Sorbonne Nouvelle, UFR de Phonétique, 318 p.
- SUSSMAN, H.M., DUDER, C., DALSTON, E., CACCIATORE, A. (1999). An acoustic analysis of the development of CV coarticulation: a case study. *Journal of speech, language and hearing research*, 42 (5), 1080-1096.
- VIHMAN, M.M., DE BOYSSON-BARDIES, B. (2000). The nature and origins of ambient language influence on infant vocal production and early words. *Phonetica*, 57, 255-266.
- VIHMAN, M.M., DEPAOLIS, R., DAVIS, B. (1998). Is there a « trochaic bias » in early word learning? Evidence from infant production in English and French. *Child development*, 69 (4), 935-949.
- VIHMAN, M.M., GREENLEE, M. (1987). Individual differences in phonological development: ages one and three years. *Journal of speech hearing and research*, 30, 503-21.
- VIHMAN, M.M., NAKAI, S., DEPAOLIS, R. (2002). Getting the rhythm right: a cross-linguistic study of

- segmental duration in babbling and first words. In GOLDSTEIN, L.M., WHALEN, D.H., BEST, C.T. (eds), *Laboratory Phonology VIII: Varieties of Phonological Competence*, 341-366. New York, Berlin: Mouton de Gruyter. <http://www.bangor.ac.uk/library/>
- VILAIN, A., ABRY, C., BADIN, P. (2000). Coproduction strategies in French VCVS: confronting Öhman's model with adult and developmental articulatory data. *Proc. of 5th Seminar on Speech Production: Models and data*. Munich, Kloster Seeon, Bavaria, Germany.
  - WHALEN, D.H., LEVITT, A.G., HSIAO, P.L., SMORODINSKY, I. (1995). Intrinsic F0 of vowels in the babbling of 6-, 9-, and 12-month-old French- and English-learning infants. *Journal of the acoustical society of America*, 97 (4), 2533-2539.
  - WHALEN, D.H., LEVITT, A.G., WANG, Q. (1991). Intonational differences between the reduplicative babbling of French- and English-learning infants. *Journal of child language*, 18 (3), 501-516.
  - WHALEN, D.H., LEVITT, A.G., GOLDSTEIN, M.L. (2007). VOT in the babbling of French- and English-learning infants. *Journal of phonetics*, 35 (3), 341-352. doi: 10.1016/j.2006.10.001