

Segmentation précoce de la parole continue en mots : évaluation inter-linguistique de l'hypothèse d'initialisation rythmique

Thierry Nazzi*

Université Paris Descartes

Laboratoire Psychologie de la Perception, (CNRS UMR 8158)

RÉSUMÉ

Ce papier propose une synthèse des études récentes portant sur la segmentation précoce de la parole continue en mots. Après avoir exposé l'importance de cette question d'un point de vue développemental, nous présentons les recherches ayant été conduites avec des enfants américains anglophones. Ces recherches montrent que les capacités de segmentation émergent vers 8 mois, se développent dans les mois qui suivent, et s'appuient sur le traitement de différents indices dont le poids relatif évolue au cours du développement. Étant donné que ces études montrent que les enfants utilisent des indices dont la réalisation diffère selon les langues, nous soulignons que le développement de ces capacités devrait varier suivant la langue en acquisition. Nous faisons alors l'hypothèse, concernant l'utilisation de certains indices prosodiques, que l'unité de segmentation utilisée précocement (unité de stress, syllabe) va varier suivant le type de rythme de la langue maternelle (langue à stress, langue syllabique). Après avoir exposé les éléments sur lesquels repose cette hypothèse, nous présentons des données récentes obtenues avec des enfants français qui apportent une première validation de cette hypothèse.

**Early segmentation of fluent speech into words :
Cross-linguistic evaluation of the rhythmic bootstrapping hypothesis**

ABSTRACT

The present paper reviews recent studies on the early segmentation of fluent speech into words. After having emphasized the importance of this issue from a developmental perspective, we present studies conducted on English-speaking American infants. These studies

*Adresse de correspondance : Thierry Nazzi, Laboratoire Psychologie de la Perception, Institut Pluridisciplinaire des Saints Pères, 45 rue des Saints-Pères, 75006 Paris. Email : thierry.nazzi@univ-paris5.fr

Remerciements : l'écriture de ce papier et les recherches effectuées dans ce cadre ont été financées en partie par une bourse de la European Science Foundation (programme EUROCORES "The Origin of Man, Language and Languages") et une ACI du Ministère de la Recherche. Je voudrais remercier les enfants et leurs familles, ainsi que tous mes collaborateurs, en particulier Josiane Bertoncini, Peter Jusczyk et Galina Iakimova.

show that segmentation abilities emerge around 8 months of age, develop in the following months, and rely on the processing of various cues the relative weight of which changes during development. Given that these studies show that infants use cues that are instantiated differently across languages, we highlight the fact that the development of these abilities might vary according to the language in acquisition. We then propose, with respect to the use of prosodic cues, that the unit of segmentation used by young infants (the trochaic unit, the syllable) will vary according to the rhythmic type of the native language (stress- vs. syllable-based language). After having presented the elements on which this hypothesis is based, we present recent data from French-learning infants that bring the first piece of evidence in favor of this hypothesis.

En quoi consiste, du point de vue développemental, la formation du lexique ? Un mot correspond à l'appariement spécifique entre la représentation mentale d'une forme sonore et un concept, qui constitue le sens associé à la forme sonore. La construction du lexique va résulter du développement de trois grands ensembles de capacités : la capacité à élaborer des représentations mentales de la forme sonore des mots, la capacité à construire des représentations abstraites (concepts) des objets et événements dans le monde et, finalement, la capacité à apparier les formes sonores acquises et les concepts formés.¹ L'acquisition du lexique fait depuis longtemps l'objet d'un très grand nombre de recherches. Dans ce qui suit, nous faisons une revue des recherches portant sur une étape spécifique de cette acquisition, à savoir celle de l'extraction, ou segmentation, des formes sonores de mots de la parole continue.

La segmentation de la parole continue en mots constitue un aspect crucial du traitement de la parole. Cette étape du traitement permet aux enfants comme aux adultes de déterminer la séquence d'unités lexicales constituant les énoncés qu'ils entendent. De plus, dans la mesure où les enfants entendent essentiellement des phrases et pas des mots isolés, la segmentation de la parole pourrait jouer un rôle central dans l'acquisition du langage. Cette capacité à extraire et représenter les formes sonores des mots individuels est bien évidemment un pré-requis pour l'acquisition du lexique. Mais il s'agit en fait aussi d'un pré-requis pour l'acquisition de la syntaxe, dans la mesure où toutes les théories d'acquisition de la syntaxe font l'hypothèse que les enfants ont accès à la séquence segmentée des mots formant les énoncés qu'ils entendent.

Ainsi donc, il est crucial pour les enfants d'avoir la capacité d'accéder aux formes sonores des mots qu'ils entendent. Accéder aux formes de mots ne serait pas un problème si les frontières de mots étaient clai-

¹On notera que l'extraction de formes sonores, comme la formation de concepts, commence avant le début de l'acquisition des mots vers l'âge de 10-12 mois. Le très jeune enfant constituerait ainsi un répertoire de formes sonores et de concepts qui seraient ultérieurement appariés pour former des mots. À partir de la deuxième année, ces acquisitions pourraient se faire de façon simultanée.

rement marquées acoustiquement, ou si les mots étaient (souvent) présentés en isolation. Suite à des études montrant que ce n'est pas le cas dans le discours entre adultes (Cole & Jakimik, 1978, 1980 ; Klatt, 1979, 1989), un tel marquage a été évalué pour la parole entendue par les enfants. En particulier, quelques études ont évalué la question de la présence de mots isolés dans l'input reçu par des enfants acquérant l'anglais (Aslin, 1993 ; Brent & Siskind, 2001) ou l'input reçu par un enfant bilingue néerlandais/allemand (van de Weijer, 1998).

Les résultats de ces études suggèrent que la plupart de la parole adressée aux enfants consiste en des énoncés composés de plusieurs mots, les mots prononcés en isolation formant moins de 10 % de tous les mots présents dans les données analysées. Ces formes prononcées en isolation pourraient aider leur acquisition par les enfants, comme le suggère le résultat montrant que le fait qu'un mot soit produit par un enfant est mieux prédit par la fréquence de présentation en isolation de ce mot quelques mois plus tôt que par la fréquence de présentation totale de ce mot au même moment (Brent & Siskind, 2001). Toutefois, la nécessité de procédures de segmentation de la parole continue persiste étant donné que tous les types de mots n'apparaissent pas en isolation (dont notamment les mots grammaticaux), et que beaucoup des mots apparaissant en isolation correspondent à des « fillers » (*oui, hmm...*), des vocatifs (« prénom de l'enfant »...) et des expressions sociales (*coucou!*...), comme l'a montré van de Weijer (1998). De plus, l'étude d'Aslin (1993) a fait apparaître que, dans une situation dans laquelle il est spécifiquement demandé à des mères d'apprendre de nouveaux mots à leurs jeunes enfants, les mots nouveaux n'étaient présentés en isolation qu'environ 25 % du temps, avec des différences entre dyades mères-enfants allant de 0 % à 70 %. On notera toutefois que, parallèlement à la mise en évidence de cette limitation dans l'apparition de mots nouveaux en isolation, cette étude a fait apparaître deux facteurs qui pourraient aider les enfants à extraire les mots de leur contexte phrastique : la grande variabilité de mots précédant les mots à apprendre d'une part, la présentation des mots à apprendre en position finale de phrase d'autre part.

INDICES DE FRONTIÈRES DE MOTS ET SENSIBILITÉ DES ENFANTS À CES INDICES

Toutefois, bien qu'il n'y ait pas de pauses entre mots consécutifs dans le signal de parole, il existe néanmoins un certain nombre d'indices linguistiques plus subtils signalant les frontières entre les mots, ou la cohésion

entre deux sons appartenant au même mot. Il s'agit premièrement d'indices prosodiques, c'est-à-dire relatifs à la façon dont accentuation et variations intonatives sont affectées par la position à l'intérieur des mots. Par exemple, en anglais, on observe qu'une majorité de mots est accentuée en position initiale (Cassidy & Kelly, 1991 ; Cutler & Carter, 1987 ; Kelly & Bock, 1988), alors qu'en français, on observe des allongements de syllabes en fin de mots ou plutôt d'unités prosodiques (Delattre, 1966 ; Fletcher, 1991). Il s'agit deuxièmement d'indices allophoniques, c'est-à-dire relatifs au fait que la réalisation acoustique de certains phonèmes varie suivant qu'ils sont en bordure ou à l'intérieur d'un mot. Ainsi en anglais, la réalisation des phonèmes /t/ et /r/ diffèrent dans le mot *nitrate* et dans la séquence *night rate* (Hohne & Jusczyk, 1994.). Troisièmement, des indices phonotactiques, c'est-à-dire relatifs aux contraintes sur les séquences phonétiques autorisées au niveau lexical, peuvent indiquer des frontières de mots. Ainsi, la séquence /zt/ pour l'anglais, ou les séquences /kf/ ou /vg/ pour le français, ne se retrouvent jamais à l'intérieur d'un même mot. Leur présence dans le signal de parole impliquerait donc nécessairement la présence d'une frontière lexicale. Quatrièmement, certaines informations statistiques/distributionnelles ayant trait à l'agencement des phonèmes et des syllabes consécutives au sein des mots peuvent fournir des indications sur la probabilité que deux sons appartiennent ou non au même mot. Ainsi, certaines transitions phonémiques sont plus fréquentes à l'intérieur des mots que d'autres (Hockema, 2006 ; Jusczyk, Luce & Charles-Luce, 1994). Et, de façon similaire, il a été montré que les probabilités de transition entre deux syllabes sont plus élevées à l'intérieur des mots qu'aux frontières lexicales (Curtin, Mintz & Christiansen, 2005).

La présence de ces indices linguistiques appelle les remarques suivantes. D'une part, aucun de ces indices linguistiques ne marque les frontières de mots de façon systématique. Mais d'autre part, il est possible que leur utilisation conjointe fournisse suffisamment d'information pour permettre une segmentation correcte du signal de parole (Christiansen, Allen & Seidenberg, 1998). De plus, la façon dont les indices précédents vont fournir des indications relatives aux frontières de mots va varier selon les langues ; les enfants devront acquérir la façon dont ces indices marquent les frontières de mots dans la langue parlée dans leur environnement. Cette remarque n'est pas sans importance dans la mesure où se pose ici un problème de « bootstrapping » (initialisation de l'acquisition) : afin de pouvoir utiliser ces indices pour segmenter la parole continue, donc pour trouver les mots dans les phrases, les enfants vont devoir acquérir comment ces indices s'alignent sur les frontières de mots. Nous discuterons un peu plus loin une solution à ce problème, au moins

pour ce qui est de certains indices prosodiques. Mais avant cela, nous présentons un certain nombre d'études ayant porté sur la sensibilité des enfants aux marqueurs linguistiques de frontières de mots dans leur langue maternelle, études qui montrent que les jeunes enfants (grandissant dans un environnement monolingue anglophone, sauf précision contraire) y sont sensibles très précocement.

Ainsi, il est apparu que les nouveau-nés français distinguent entre une même séquence bisyllabique (*mati*) suivant qu'une frontière de mots/d'unité prosodique est présente (*panorama typique*) ou non (*mathématicien*) entre les deux syllabes de cette séquence (Christophe, Dupoux, Bertoncini & Mehler, 1994). Une sensibilité aux différences allophoniques a aussi été trouvée chez les enfants dès l'âge de 2 mois (Hohne & Jusczyk, 1994), comme l'atteste leur capacité à discriminer entre paires telles *nitrate* et *night rate*. De plus, les enfants deviennent sensibles aux propriétés phonotactiques de leur langue maternelle entre les âges de 6 et 9 mois (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993 ; Jusczyk, Luce & Charles-Luce, 1994 ; Mattys, Jusczyk, Luce, & Morgan, 1999 ; voir aussi Friederici & Wessels, 1993, pour le néerlandais ; Sebastián-Gallés & Bosch, 2002, pour le catalan). Ceci est montré par l'émergence d'une préférence pour des séquences de phonèmes légales ou fréquentes dans la langue maternelle comparées à des séquences illégales ou rares (par exemple, en anglais, *chun* est un non-mot à phonotactique fréquente alors que *yush* est un non-mot à phonotactique rare). Finalement, une préférence pour les mots anglais ayant le pattern d'accentuation prédominant trochaïque (accentuation initiale ou « fort-faible », comme *PORTer*) sur les mots anglais ayant le pattern d'accentuation plus rare iambique (accentuation finale ou « faible-fort », comme *rePORT*) émerge également entre les âges de 6 et 9 mois chez les enfants acquérant l'anglais (Jusczyk, Cutler, & Redanz, 1993 ; Turk, Jusczyk, & Cutler, 1995).

Faisant suite à ces études montrant l'émergence, au cours de la deuxième partie de la première année, de la sensibilité des enfants à certaines propriétés de leur langue maternelle qui marquent partiellement les frontières de mots, d'autres études ont commencé à explorer si, et quand, les enfants utilisent ces informations pour segmenter les mots de la parole continue. Ces études sont présentées dans la section suivante. Nous rapportons d'abord les études ayant porté sur des enfants acquérant l'anglais, car c'est la langue sur laquelle portent la majorité des études, et la seule langue sur laquelle l'utilisation de plusieurs indices a été explorée à ce jour.

ÉMERGENCE ET DÉVELOPPEMENT DES CAPACITÉS DE SEGMENTATION CHEZ LES ENFANTS ACQUÉRANT L'ANGLAIS

Ces études sur l'utilisation par les enfants d'indices de frontières de mots pour segmenter la parole reposent sur la démonstration princeps du fait que les enfants commencent vers 8 mois à extraire les mots de la parole continue (Jusczyk & Aslin, 1995). Dans la première expérience de cette étude reposant sur une adaptation de la méthode de regard préférentiel, des enfants de 7.5 mois ont été familiarisés avec deux mots monosyllabiques (*cup* et *dog*, ou *bike* et *feet*) puis ont entendu quatre passages, chaque passage étant construit autour de l'un des quatre mots cibles (chacun de ces mots étant répété 6 fois dans le passage en question). Les résultats ont fait apparaître une préférence pour les passages correspondant aux mots présentés pendant la familiarisation, suggérant que les enfants ont reconnu dans les passages les mots familiarisés, reconnaissance impliquant que les enfants ont dû segmenter ces mots des passages.¹ Jusczyk et Aslin (1995) ont ensuite cherché à répliquer ce résultat avec des enfants de 6 mois, mais sans succès, d'où leur conclusion que les capacités de segmentation doivent apparaître entre 6 et 7.5 mois.

Deux autres expériences ont ensuite été réalisées pour préciser les capacités de segmentation des enfants de 7.5 mois. D'une part, Jusczyk et Aslin (1995) ont montré que les formes segmentées semblent être phonétiquement spécifiées, ces enfants ne montrant pas d'effet de segmentation lorsqu'ils sont familiarisés avec le pseudomot *zeet* alors que le mot cible contenu dans les passages est le mot *feet*, résultat ultérieurement étendu à la consonne finale (*feek*) par Tincoff et Jusczyk (1996). Par contre, la question de la spécification de la voyelle n'a pas été explorée (voir Nazzi, 2005, sur cette question). D'autre part, Jusczyk et Aslin (1995) ont montré que si l'on inverse l'ordre de présentation des stimuli, à savoir que l'on commence par familiariser les enfants avec deux passages (chacun comprenant un mot cible), les enfants montrent une préférence pour l'écoute des deux mots cibles comparée à l'écoute de deux mots contrôles. Cette dernière expérience, qui correspond plus à la situation dans laquelle se trouvent les jeunes enfants, qui entendent des phrases

¹Comme nous allons le voir plus loin, cette interprétation de la reconnaissance en terme de segmentation est confirmée par l'étude de Jusczyk *et al.* (1999), dont, par exemple, le fait que des enfants du même âge ne reconnaissent pas la seule syllabe accentuée d'un mot bisyllabique trochaïque (comme *DOCTOR*) alors qu'ils reconnaissent le mot entier.

sans connaître (beaucoup) de mots, rend encore plus crédible l'hypothèse que la segmentation continue de la parole en mots pourrait permettre aux enfants d'accéder à la forme sonore des mots de leur langue maternelle.

L'importance de l'étude de Jusczyk et Aslin (1995) réside donc dans le fait (1) d'avoir mis au point une méthode expérimentale adaptée à l'étude des capacités de segmentation de la parole en mots, et (2) d'avoir mis en évidence pour la première fois de telles capacités chez de jeunes enfants. De nombreuses questions restent toutefois posées, notamment la question de savoir comment les enfants arrivent à segmenter ces mots. Cette question a été reprise dans de nombreuses études ultérieures, lesquelles ont exploré l'utilisation par les enfants de divers indices linguistiques de segmentation.

Des études sur les enfants acquérant l'anglais émerge le scénario développemental suivant. Vers l'âge de 7-8 mois, les enfants utiliseraient des informations rythmiques afin de segmenter la parole continue en séquences d'une ou plusieurs syllabes commençant par une syllabe accentuée, à savoir donc, en unités trochaïques (Jusczyk, Houston & Newsome, 1999b ; pour des données allant dans le même sens, voir aussi Curtin, Mintz & Christiansen, 2005 ; Echols, Crowhurst, & Childers, 1997 ; Houston, Santelmann & Jusczyk, 2004 ; Johnson & Jusczyk, 2001 ; Morgan & Saffran, 1995 ; Nazzi, Dilley, Jusczyk, Shattuck-Hufnagel & Jusczyk, 2005). Étant donné que la plupart des mots bisyllabiques anglais ont un pattern accentuel trochaïque (Cassidy & Kelly, 1991 ; Cutler & Carter, 1987 ; Kelly & Bock, 1988), cette procédure de segmentation rythmique (similaire à la stratégie de segmentation métrique utilisée par les adultes, voir Cutler, Mehler, Norris, & Segui, 1986 ; Cutler & Norris, 1988 ; McQueen, Norris, & Cutler, 1994) permettrait aux enfants acquérant l'anglais, dès un très jeune âge, de segmenter de façon appropriée la plupart des mots bisyllabiques. Cette hypothèse est soutenue par les résultats de Jusczyk *et al.* (1999b) montrant que des enfants acquérant l'anglais segmentent des noms trochaïques (fort-faible, tels *DOCTOR*) dès l'âge de 7.5 mois, en même temps qu'ils font des erreurs de segmentation avec des noms iambiques (faible-fort, tels *guiTAR*), plaçant une frontière entre la syllabe initiale/faible et la syllabe finale/forte (par exemple, *gui / TAR*, voir Jusczyk *et al.*, 1999b). Un décalage similaire à l'avantage des mots trochaïques a également été trouvé pour les verbes anglais, alors même que la plupart des verbes bisyllabiques anglais ont un pattern d'accentuation non-prédominant iambique, suggérant que le biais pourrait s'appliquer à toutes les catégories lexicales de cette langue (Nazzi *et al.*, 2005). Finalement, on notera que cette proposition de procédure de segmentation

rythmique est compatible avec les données sur la segmentation de mots monosyllabiques (Jusczyk & Aslin, 1995) étant donné que ces mots correspondaient à des syllabes fortes (donc des unités trochaïques « monosyllabiques »).

Il a aussi été montré que les régularités distributionnelles dans l'ordre des syllabes du signal de parole (que l'on nommera par la suite de la façon suivante : informations distributionnelles) constitue un indice crucial de segmentation précoce. Par exemple, il est apparu que des enfants de 7.5 mois testés avec des passages contenant des mots trochaïques tels *DOCTOR* ont un effet de segmentation si familiarisés avec les mots entiers, mais pas s'ils sont familiarisés avec la syllabe initiale de ces mots, tels *DOC* (Jusczyk *et al.*, 1999b). De plus, utilisant un paradigme de langage artificiel dans lequel les enfants sont présentés avec une séquence continue formée de la répétition en ordre aléatoire de 4 pseudomots trisyllabiques, il a été montré que des enfants de 8 mois groupent ensemble des syllabes contiguës dans le signal sur la base des probabilités transitionnelles entre syllabes consécutives (Saffran, Aslin, & Newport, 1996 ; mais voir Perruchet & Vinter, 1998, pour une interprétation alternative de ces résultats, et Brent & Cartwright, 1996, Dahan & Brent, 1999, pour un modèle alternatif).

De façon importante, les données obtenues en utilisant le paradigme de Jusczyk et Aslin (1995) suggèrent que les enfants acquérant l'anglais utilisent d'abord l'information rythmique pour segmenter la parole, les informations distributionnelles étant initialement utilisées seulement au sein des unités rythmiquement définies (Jusczyk *et al.*, 1999b). En effet, la frontière prosodique/rythmique placée entre les deux syllabes d'un mot iambique (tel *guiTAR*) semble bloquer l'utilisation par les enfants de 7.5 mois des informations distributionnelles (le fait que *gui* et *TAR* apparaissent toujours de façon consécutive), ce qui conduit à la segmentation isolée de la syllabe forte. De façon similaire, si un mot iambique est toujours suivi de la même syllabe faible (par exemple, *guiTAR_is*), les enfants de 7.5 mois placent toujours une frontière de mots entre les deux premières syllabes (du fait de l'utilisation d'information rythmique) et groupent ensemble les deux dernières syllabes (du fait de l'utilisation d'informations distributionnelles). Ceci résulte en une segmentation erronée (par exemple, *gui / TARis*). Ces résultats suggèrent que les enfants de 7.5 mois utilisent l'information rythmique pour effectuer une segmentation première de la parole continue en unités plus petites, lesquelles servent de base à l'analyse ultérieure du signal (voir toutefois le débat entre Johnson & Jusczyk, 2001, et Thiessen & Saffran, 2003, concernant

les données sur cette question obtenues avec le paradigme de langage artificiel¹).

Plus tard au cours du développement, il apparaît des changements dans le rôle des informations rythmiques et/ou la relation entre l'utilisation des indices rythmiques et des autres indices de segmentation. Cette évolution est attestée par des données montrant que les enfants de 10.5 mois (contrairement à ce que l'on avait vu pour les enfants de 7.5 mois) peuvent segmenter correctement les mots iambiques, et ne font plus d'erreur de segmentation lors de la présentation de mots iambiques qui sont toujours suivis de la même syllabe faible (Juszyk *et al.*, 1999b). Ceci pourrait refléter le fait que ces enfants utilisent maintenant les informations distributionnelles même lorsqu'elles traversent une frontière rythmiquement définie, ce qui suggérerait qu'à l'âge de 10.5 mois, les enfants accordent plus de poids aux informations distributionnelles qu'aux informations rythmiques.

De façon complémentaire, ceci pourrait également refléter le fait que les enfants de 10.5 mois utilisent d'autres indices de frontières lexicales, tels les indices allophoniques (Juszyk, Hohne, & Bauman, 1999a), phonotactiques (Mattys & Juszyk, 2001a). Ils utilisent aussi des indices prosodiques autres que l'indice rythmique : les frontières de syntagmes (Gout, Christophe & Morgan, 2004 ; Nazzi *et al.*, 2005) à 10.5 mois,² les informations d'accent de pitch entre 10 et 13 mois (Nazzi *et al.*, 2005). Il a de plus été montré que la segmentation est influencée par la coarticulation à 8 mois (Johnson & Juszyk, 2001). Enfin, les mots commençant par une consonne semblent plus faciles à segmenter que ceux commençant par une voyelle à 8-13 mois en anglais (Mattys & Juszyk, 2001b ; Nazzi *et al.*, 2005). On notera que cette question de la segmentation des mots commençant par une voyelle se pose aussi en français, en lien avec celle de l'acquisition de la liaison (Chevrot, Dugua, & Fayol, 2005 ; Wauquier-Gravelines, 2002).

¹Thiessen et Saffran (2003) ont répliqué le résultat de Johnson et Juszyk (2001) montrant que les enfants de 9 mois s'appuient plus sur les informations rythmiques que sur les informations distributionnelles pour segmenter les mots dans un paradigme de langage artificiel. Toutefois, ils ont trouvé un pattern inverse à 7 mois, ce qui les a amené à la conclusion que les informations distributionnelles étaient utilisées plus tôt que les informations rythmiques. Bien que cela puisse être le cas dans ce contexte expérimental, il est important de garder à l'esprit que les langages artificiels qui doivent être appris dans ces expériences sont très simples, puisque constitués de seulement 4 mots ayant le même nombre de syllabes (3 syllabes). De fait, bien que les résultats de Thiessen et Saffran (2003) montrent que les enfants de 7 mois peuvent exploiter les informations distributionnelles et les utiliser dans les minutes suivant l'exposition à un langage très simple, il est loin d'être montré qu'ils pourraient exploiter ces informations à un si jeune âge dans le contexte d'une langue naturelle constituée de milliers de mots de durée syllabique variée.

²L'utilisation de ces frontières prosodiques pourraient permettre aux enfants de découper les énoncés en séquences plus courtes, clauses ou syntagmes, dans lesquelles s'appliqueraient les autres procédures de segmentation « lexicales » (c.f. aussi les études de Hirsh-Pasek, Kemler Nelson, Juszyk, Wright Cassidy, Druss & Kennedy, 1987 ; Juszyk, Hirsh-Pasek, Kemler Nelson, Kennedy, Woodward & Piwoz, 1992 ; Nazzi, Kemler Nelson, Juszyk & Juszyk, 2000 ; Soderstrom, Seidl, Kemler Nelson & Juszyk, 2003).

En résumé, les études décrites ci-dessus appuient la proposition d'une acquisition précoce des formes sonores de mots selon un scénario *bottom-up* de « bootstrapping » phonologique. En effet, ces études ont commencé à tracer l'esquisse du développement au cours de la première année des capacités de segmentation de la parole continue en formes sonores de mots, lesquelles reposent sur l'exploitation de divers indices de frontières de mots. Les informations rythmiques et distributionnelles sont utilisées dès l'âge de 8 mois, et il apparaît que les enfants utilisent une procédure de segmentation rythmique lors de leur découpage initial du signal (en unités trochaïques), utilisant d'abord les informations distributionnelles seulement à l'intérieur des unités rythmiquement définies. À 10.5 mois, le poids relatif de ces deux indices a changé, et les enfants ont commencé à utiliser d'autres indices (par exemple, phonotactiques, allophoniques).

À ce stade, on notera que la discussion précédente porte sur l'utilisation de procédures de segmentation *bottom-up* dans une perspective de « bootstrapping ». Toutefois, on a vu que quelques mots sont prononcés de façon isolée. Ces mots pourraient être stockés et utilisés pour effectuer un autre type de segmentation : une segmentation *top-down*. Le modèle Incdrop (Brent & Cartwright, 1996) se situe dans cette perspective. En effet, les enfants mémoriseraient les énoncés entendus de façon entière (ex : *mariboidelo*), à moins que ces énoncés contiennent une séquence qui a été antérieurement mémorisée (*boi*) ; dans ce cas de figure, l'énoncé nouveau est segmenté de la séquence préalablement mémorisée, et les séquences complémentaires sont mémorisées (*mari* et *delo*). Des données ont pu être obtenues en appui de ce modèle, à la fois d'un point de vue de la modélisation (Brent & Cartwright, 1996) ainsi que d'un point de vue de l'acquisition par les adultes d'un langage artificiel (Dahan & Brent, 1999). Des données récentes suggèrent que des mots connus peuvent également faciliter la segmentation chez les enfants. Ainsi, il a été trouvé que des enfants acquérant l'anglais peuvent segmenter des mots non-familiers de la parole continue dès l'âge de 6 mois (comparé aux 7.5 mois trouvés dans l'étude de Jusczyk & Aslin, 1995) si ces mots sont précédés de mots très familiers, tels le prénom de l'enfant ou le mot *mommy* (Bortfeld, Morgan, Golinkoff & Rathbun, 2005). Si ces effets jouent probablement un rôle marginal au début de l'acquisition lexicale quand les enfants n'ont mémorisé qu'un nombre limité de mots, leur rôle augmente probablement en parallèle à l'augmentation de la taille du vocabulaire des enfants.

Pour finir, on notera que l'utilisation par les enfants de ces différents indices est conforme aux prédictions développementales du modèle hiérarchique d'utilisation des indices de segmentation proposé par Mattys, White et Melhorn (2005) afin de rendre compte de la segmentation de la

parole par les adultes (à l'exception du fait que ce modèle ne discute pas de l'utilisation d'informations distributionnelles). Dans ce modèle, il existe trois niveaux d'indices : les indices sub-lexicaux suprasegmentaux, à savoir la prosodie (niveau 1), les indices sub-lexicaux segmentaux (phonotactique, allophonie, coarticulation ; niveau 2), et enfin les indices lexicaux (sémantique, mais aussi syntaxe et pragmatique ; niveau 3). Dans de bonnes conditions d'écoute, les adultes vont s'appuyer sur le niveau lexical ; mais si l'information entendue est dégradée, ils vont avoir recourt au niveau segmental, voire au niveau prosodique dans les conditions les plus dégradées. Ce modèle prédit aussi, ce qui est partiellement confirmé par les données présentées ci-dessus, qu'au cours du développement, les enfants vont d'abord s'appuyer sur le niveau prosodique, puis sur le niveau segmental avant de commencer à s'appuyer sur le niveau lexical.

DIFFÉRENCES INTERLANGUES : HYPOTHÈSE D'UNE INITIALISATION RYTHMIQUE DE LA SEGMENTATION

Le pattern développemental décrit dans la section précédente a été établi pour des enfants acquérant l'anglais. Étant donné que la plupart des indices utilisés pour segmenter la parole sont réalisés de façon spécifique dans une langue donnée (indices prosodiques, phonotactiques, allophoniques), donc que les frontières de mots sont signalées différemment dans différentes langues, il est possible que le pattern d'émergence des capacités de segmentation va différer selon les langues, au moins dans le poids relatif accordé à chaque indice à différents âges. Par exemple, il est possible que la force du lien entre contraintes phonotactiques et frontières de mots (donc la saillance relative de ces indices dans le signal) diffère selon les langues, ce qui pourrait entraîner des décalages dans le timing de début d'utilisation de ces indices et des différences dans leur poids relatif selon les langues. Pour évaluer cette possibilité, il est crucial de déterminer les patterns d'acquisition de ces capacités pour différentes langues.

Mais peut-être de façon encore plus cruciale, l'approche interlinguistique devrait nous permettre de préciser comment les enfants commencent à utiliser ces différents indices de segmentation. En effet, quand bien même les recherches sur l'anglais ont permis d'identifier certains des indices utilisés pour segmenter la parole, ces recherches laissent sans réponse, à l'heure actuelle, la question cruciale suivante : comment

est-il possible que les enfants puissent s'appuyer, pour commencer à segmenter les (formes sonores de) mots, sur l'utilisation de propriétés dont la plupart sont définies au niveau lexical (pattern de stress/unité trochaïque, contraintes phonotactiques et variations allophoniques) ? Ou, en d'autres termes : quelles sont les origines développementales des procédures de segmentation utilisées par les enfants ?

Différentes réponses à cette question ont été proposées. D'une part, les enfants pourraient apprendre ces propriétés en analysant les quelques mots qu'ils entendent en isolation, et pourraient ensuite utiliser leur connaissance de ces propriétés pour segmenter les énoncés multisyllabiques. Cette proposition est intéressante du fait de sa simplicité. Toutefois, il semble contradictoire d'utiliser d'une part le fait que très peu de mots apparaissent en isolation dans le signal de parole afin de proposer l'hypothèse d'une implication cruciale des procédures de segmentation de la parole continue dans l'acquisition précoce des premiers mots et, d'autre part, de soutenir que ces quelques mots entendus en isolation pourraient permettre l'acquisition des propriétés lexicales qui seront utilisées pour segmenter la parole.

Une deuxième solution à ce problème est que les enfants commenceraient en fait par effectuer des analyses distributionnelles des signaux de parole entendus, ce qui leur permettrait d'analyser et de spécifier les régularités rythmiques, allophoniques et phonotactiques des énoncés entendus. Dans la mesure où les processus d'analyses distributionnelles pourraient être domaine-général (c'est-à-dire, aussi utilisables pour traiter des signaux autre que linguistiques, comme de la musique ou des signaux visuels), cette proposition semble au premier abord apporter une solution au problème de l'acquisition phonologique préalable. Cependant, il est important de noter que ces mécanismes devraient opérer sur des représentations phonologiques spécifiques à la langue en acquisition (si ce n'est pour résoudre la question de la normalisation des réalisations acoustiques des phonèmes), d'où la nécessité d'une acquisition préalable. De plus, la position d'une utilisation d'informations distributionnelles antérieure à celle d'informations d'autre nature, bien que supportée par une partie des recherches sur l'acquisition de langages artificiels (Thiessen & Saffran, 2003 ; mais voir Johnson & Jusczyk, 2001), ne semble pas tenir au regard d'études portant sur le traitement de stimuli de parole plus variés et complexes que ceux utilisés dans les études sur les langues artificielles (Jusczyk *et al.*, 1999b). Dans ces conditions expérimentales, les données suggèrent plutôt que l'information rythmique est utilisée avant les informations distributionnelles (et allophoniques et phonotactiques).

C'est d'ailleurs sur la base de ces résultats suggérant une antériorité de l'utilisation des informations rythmiques sur tous les autres indices de segmentation qu'une troisième hypothèse a été proposée. Cette hypothèse précise comment les enfants pourraient utiliser, au moment où se mettent en place les procédures de segmentation, des indices qui semblent être définis au niveau lexical (puisque'il est généralement, et de façon paradoxale, fait un lien entre segmentation rythmique et le fait que le pattern prédominant des mots anglais est le pattern trochaïque). Cette hypothèse, dite d'**initialisation rythmique de la segmentation**, peut se résumer de la façon suivante.

La segmentation mise en évidence par Jusczyk *et al.* (1999b) correspondrait en fait à une segmentation basée sur l'extraction par les enfants anglophones de portion du signal de parole se conformant à l'unité trochaïque (à savoir, une syllabe accentuée éventuellement suivie d'une ou plusieurs syllabes non-accentuées). Notre hypothèse est donc que la segmentation rythmique précoce s'appuie en fait sur l'unité rythmique de la langue maternelle. Cette unité varierait suivant les langues. Toutefois, il n'existerait qu'un nombre limité d'unités rythmiques (définissant un nombre limité de classes rythmiques) comme le montre un certain nombre d'études en linguistique et en psycholinguistique adulte. L'acquisition par les enfants de la procédure de segmentation rythmique appropriée à leur langue maternelle passerait donc par la spécification de l'unité rythmique sous-jacente : l'unité trochaïque pour l'anglais, la syllabe pour le français. Selon notre hypothèse, qui se place dans le cadre de la théorie du « bootstrapping » (ou : initialisation de l'acquisition) phonologique, cette spécification serait possible du fait d'une sensibilité des jeunes enfants au marquage acoustique à un niveau suprasegmental des différents types de rythme, comme le montrent les études sur la sensibilité des jeunes enfants au rythme global de leur langue maternelle (Mehler, Dupoux, Nazzi & Dehaene-Lambertz, 1996 ; Nazzi, Bertoncini & Mehler, 1998 ; Nazzi, Jusczyk & Johnson, 2000 ; Nazzi & Ramus, 2003).¹

Ainsi, si les informations rythmiques jouent un rôle crucial au moment de l'émergence des capacités de segmentation, c'est parce que la procédure de segmentation prosodique pourrait être apprise indépendamment du niveau lexical. De ce fait, cette acquisition permettrait aux enfants de commencer à segmenter les premières formes sonores de mots et, partant de là, de spécifier progressivement les autres indices phonologiques (allophonie,

¹On notera que cette proposition rejoint un certain nombre d'autres propositions récentes concernant la sensibilité précoce des enfants aux informations prosodiques, et l'importance de la prosodie dans l'acquisition langagière précoce (Fernald & Kuhl, 1987 ; Jusczyk *et al.*, 1993a, 1993b ; Karzon & Nicholas, 1989 ; Kuhl & Miller, 1982 ; Nazzi *et al.*, 1998a, 1998b ; Spring & Dale, 1977).

phonotactique...) marquant les frontières de mots. Cette hypothèse d'initialisation rythmique de la segmentation offre donc une solution à la question de la mise en place des procédures de segmentation, quelle que soit la langue en acquisition. Mais comme on le voit, elle fait des prédictions différentes pour des langues de rythmes différents (utilisation de différentes unités rythmiques pour segmenter), et pour être évaluée, nécessite donc une approche interlinguistique. Avant de présenter les études sur la segmentation de la parole en français permettant d'évaluer cette hypothèse, nous allons exposer en plus de détails les données sur le rythme linguistique et son influence sur le traitement de la parole par les adultes et les jeunes enfants sur lesquelles s'appuie notre hypothèse d'initialisation rythmique.

DU RYTHME LINGUISTIQUE ET DE SON INFLUENCE SUR LE TRAITEMENT DE LA PAROLE

Comme nous l'avons mentionné ci-dessus, l'hypothèse d'initialisation rythmique de la segmentation repose sur plusieurs séries de faits suggérant (a) l'existence d'indices rythmiques, présents dans le signal de parole. Ces indices semblent (b) influencer le traitement de parole par les adultes, et (c) être perçus par les enfants dès la naissance. Dans ce qui suit, nous présentons une revue des données concernant ces trois points, et discutons du fait qu'ils pourraient permettre l'acquisition des procédures de segmentation spécifiques à la langue maternelle.

Classes de rythme

L'idée qu'il existe plusieurs classes rythmiques remonte à plusieurs décennies (Abercrombie, 1967 ; Pike, 1945). Dans la littérature, trois classes rythmiques de langues sont généralement considérées : les langues à stress (comme l'anglais, le néerlandais...), les langues syllabiques (comme le français, l'italien...) et les langues moraïques¹ (comme le japo-

¹La more est une unité rythmique qui peut être soit syllabique soit sub-syllabique. En anglais, une more correspond globalement à une syllabe CV avec une voyelle courte (comme *the* par opposition à *thee*, qui a une voyelle longue). En japonais, les syllabes CV avec une voyelle longue, ou les syllabes avec une consonne finale nasale (comme la première syllabe de *Honda*) ou géminée (comme la première syllabe de *Nissan*) ont deux mores.

nais, le tamoul). À chacune de ces classes correspondrait une unité de rythme sous-jacente donnée : respectivement, l'unité de stress, la syllabe et la more. On notera que ces trois unités rythmiques sont dans une relation hiérarchique au niveau phonologique : les unités de stress sont constituées de syllabes qui sont elles-même constituées de mores.

La définition initiale de ces classes rythmiques supposait une isochronie dans la réalisation des unités de rythme sous-jacentes à une classe de rythme donnée. Soit, par exemple pour les langues à stress, une durée relativement égale (ou une tendance à l'égalité de durée dans des versions plus modérées de l'hypothèse d'isochronie) de toutes les unités de stress indépendamment de leur nombre de syllabes, de leur position dans la phrase... Si cette définition s'est révélée incorrecte, des études plus récentes, cherchant une signature acoustique plus subtile de ces différentes classes, ont commencé à apporter les premiers éléments supportant d'un point de vue acoustique l'existence de ces classes de rythme (Arvaniti, 1994 ; den Os, 1988 ; Fant, Kruckenberg & Nord, 1991 ; Ramus, Nespor & Mehler, 1999 ; Shafer, Shucard, & Jaeger, 1999). Ainsi, les analyses de Ramus *et al.* (1999) sur des énoncés provenant de huit langues différentes ont fait apparaître que deux mesures acoustiques (proportion de durée des intervalles vocaux ; déviation standard des durées d'intervalles consonantiques moyens par phrase) permettaient de définir un espace acoustique bidimensionnel dans lequel ces langues se répartissaient en fonction des trois classes rythmiques à stress (anglais, néerlandais, polonais), syllabiques (français, italien, espagnol, catalan) et moraique (japonais).

Différences inter-linguistiques de traitement de la parole par des sujets adultes et classes rythmiques

Le recours à ces classes rythmiques est apparu utile pour expliquer les différences inter-linguistiques dans la façon dont les adultes segmentent la parole et accèdent au lexique, en cela qu'il paraît y avoir un lien entre le comportement des adultes dans certaines tâches expérimentales et les propriétés rythmiques globales de leur langue maternelle. En effet, plusieurs études portant sur la détection de fragments de mots par les adultes indiquent que la segmentation de la parole est influencée par le système rythmique de la langue maternelle, de telle façon que des adultes parlant français, anglais et japonais utilisent des procédures de segmentation prosodique/rythmique différentes. Il est également apparu que chaque procédure prosodique est basée sur l'unité métrique caractéristique de la langue considérée.

Ainsi, la syllabe est apparue comme l'unité de segmentation des adultes parlant français (Mehler, Dommergues, Frauenfelder, & Segui, 1981 ; Peretz, Lussier & Béland, 1998 ; mais voir Content, Meunier, Kearns, & Frauenfelder, 2001¹), espagnol et catalan (Sebastián-Gallés, Dupoux, Segui, & Mehler, 1992). Par contre, la procédure de segmentation utilisée en anglais (Cutler *et al.*, 1986 ; Cutler & Norris, 1988 ; McQueen *et al.*, 1994) et néerlandais (Vroomen, van Zon, & de Gelder, 1996) repose apparemment sur le pattern d'accentuation typique des mots de ces langues, ou unité de stress trochaïque. Un troisième pattern a été trouvé pour les adultes japonais, montrant qu'ils s'appuient sur la more (Otake *et al.*, 1993). Dans la mesure où il existe une relation hiérarchique entre ces trois unités rythmiques, l'acquisition de la procédure maternelle pourrait découler de la spécification du niveau d'unité rythmique le plus approprié à la langue en acquisition.

Finalement, les procédures de segmentation des adultes paraissent être profondément ancrées dans leurs capacités spécifiques à la langue maternelle, et acquises à un très jeune âge. Cette affirmation repose sur le fait que la procédure qu'ils utilisent est déterminée par leur langue maternelle et non par le rythme de la langue qu'ils sont en train d'écouter : une fois qu'ils ont acquis la procédure appropriée à une langue donnée, les adultes s'appuient sur cette procédure même lorsqu'ils entendent des stimuli prononcés dans une autre langue (Cutler *et al.*, 1986 ; Otake *et al.*, 1993). De plus, il a été montré que même des bilingues très performants sont dominants dans l'une de leur langue, et ont développé les procédures de segmentation prosodiques dans seulement l'une de leurs langues (Cutler *et al.*, 1992).

L'interprétation qui a été faite de ces données est qu'il existe différents types de procédures de segmentation rythmique/prosodique, chaque procédure étant adaptée de façon optimale au traitement d'une classe rythmique de langues donnée (Cutler & Mehler, 1993 ; Otake *et al.*, 1993 ; see also Sebastián-Gallés *et al.*, 1992 ; Vroomen *et al.*, 1996), même si des différences mineures de traitement peuvent être trouvées au sein de chaque classe.

Comme nous l'avons vu plus tôt, l'hypothèse selon laquelle les propriétés rythmiques des signaux de parole déterminent les procédures de traitement des adultes a influencé notre conception de la façon dont les

¹Cette dernière étude a fait apparaître des limitations dans l'obtention d'effets syllabiques en français, lesquels semblent être particulièrement robustes pour les seuls cas où une consonne liquide se trouve en position inter-syllabique pivotale. On notera toutefois que malgré les limitations relatives à l'effet syllabique en français, l'existence de différences inter-linguistiques dans ces tâches n'est pas totalement remise en question puisque des effets syllabiques ne sont pas obtenus en anglais y compris avec des consonnes inter-syllabiques pivotales liquides.

enfants acquièrent les procédures de traitement appropriées à la segmentation des énoncés de leur langue maternelle. Ainsi, Mehler, Dupoux, Nazzi et Dehaene-Lambertz (1996) ont proposé que l'émergence des procédures de segmentation prosodique reposerait sur une sensibilité précoce à la prosodie, et plus précisément ici, au rythme linguistique spécifié à un niveau non-lexical. Dans ce qui suit, nous présentons des données expérimentales relatives à cette proposition.

Sensibilité des enfants aux classes de rythme et acquisition des propriétés rythmiques : études sur la discrimination précoce des langues

L'hypothèse d'une sensibilité précoce aux propriétés/classes rythmiques qui permettrait d'acquérir les procédures de segmentation prosodiques/rythmiques appropriées à la langue maternelle prend appui sur un certain nombre de données explorant la capacité des jeunes enfants à discriminer entre des énoncés de différentes langues. Des premières études explorant cette question, il avait été conclu que les capacités de discrimination des langues par les nouveau-nés reposaient sur la reconnaissance de la langue maternelle (Mehler, Jusczyk, Lambertz, Halsted, Bertoincini & Amiel-Tison, 1988 ; Moon, Panneton-Cooper & Fifer, 1993). Toutefois, des études ultérieures ont exploré si ces discriminations précoces n'étaient pas en fait dues à la sensibilité des nouveau-nés aux caractéristiques rythmiques des énoncés, sensibilité permettant de séparer les énoncés (les langues) entre un nombre limité de classes rythmiques.

Ainsi, Nazzi, Bertoincini et Mehler (1998) ont présenté à des nouveau-nés différentes combinaisons de langues qui leur étaient non-familiales. Les stimuli étaient filtrés passe-bas afin de dégrader les informations phonétiques tout en préservant les informations prosodiques. La distance rythmique entre ces langues était manipulée de façon systématique, les langues contrastées appartenant soit à deux classes rythmiques différentes (anglais à stress et japonais moraïque), soit à la même classe (anglais et néerlandais à stress). Les résultats ont fait apparaître une discrimination seulement dans le cas des langues de classes rythmiques différentes. De plus, Nazzi *et al.* (1998a) ont montré que si l'on familiarise des nouveau-nés avec des énoncés de deux langues, puis qu'on les teste avec des énoncés de deux autres langues, les enfants discriminent si les langues sont arrangées de façon congruente à la partition en classes rythmiques (anglais et néerlandais à stress, *versus* italien et espagnol syllabiques), mais

pas si elles sont arrangées de façon orthogonale aux classes rythmiques (par exemple : anglais à stress et italien syllabique *versus* néerlandais à stress et espagnol syllabique). Ces données sont en accord avec l'hypothèse de discriminations précoces entre langues basées sur le rythme et les classes rythmiques.

Une étude ultérieure de discrimination néerlandais/japonais, utilisant des stimuli resynthétisés (dans lesquels les différences phonétiques sont neutralisées), a confirmé cette interprétation en terme de discrimination basée sur le rythme dès la naissance (Ramus, Hauser, Miller, Morris & Mehler, 2000). Ramus (2002) a aussi montré que si le contour de F0 des phrases initiales néerlandaises et japonaises est neutralisé, un résultat de discrimination est encore obtenu, suggérant que les différences d'intonation ne peuvent à elles seules expliquer la capacité des nouveau-nés à discriminer, ce qui confirme que les nouveau-nés peuvent s'appuyer sur des différences rythmiques pour discriminer. Toutefois, l'effet obtenu est plus faible, suggérant une implication de l'intonation (ce qui n'est pas étonnant, rythme et intonation n'étant pas entièrement indépendants).

Une sensibilité précoce au rythme étant établie, des études ultérieures ont cherché à préciser l'évolution des capacités de discrimination des langues au cours du développement, afin de déterminer si ces capacités précoces pouvaient permettre l'acquisition par les enfants des propriétés prosodiques/rythmiques de leur langue maternelle, acquisition qui leur permettrait ensuite de spécifier l'unité rythmique de leur langue maternelle, puis la procédure de segmentation prosodique qui y est associée. Les données obtenues par Nazzi, Jusczyk et Johnson (2000) avec des enfants américains anglophones de 5 mois suggèrent que la sensibilité des nouveau-nés aux classes rythmiques permet l'acquisition des propriétés rythmiques de la langue maternelle (quelles soient ou non partagées avec d'autres langues de sa classe) plutôt que les propriétés communes à l'ensemble de la classe maternelle, correspondant à une meilleure capacité à discriminer la langue maternelle d'autres langues de sa classe rythmique. En effet, il apparaît un rôle critique à la fois de la classe rythmique et de la langue maternelle. Ces enfants peuvent discriminer deux langues de même classe rythmique (contrairement aux nouveau-nés), mais seulement si une variante de leur langue maternelle est présentée : anglais britannique *versus* néerlandais, et anglais britannique *versus* anglais américain (voir aussi Bosch & Sebastián-Gallés, 1997, pour des résultats analogues avec des enfants espagnols et catalans âgés de 4 mois). Par contre, si la langue maternelle n'est pas présente, les enfants échouent à discriminer des langues de même classe rythmique, qu'elles fassent partie

de la classe rythmique maternelle (néerlandais/allemand) ou pas (italien/espagnol).

On notera que ces derniers résultats sont compatibles avec l'hypothèse d'une acquisition par les enfants de l'unité de rythme de leur langue maternelle sur la base de leur sensibilité initiale aux différents types de rythme linguistique, et confèrent à cette hypothèse une certaine plausibilité. Si la démonstration expérimentale de ce lien reste encore à faire, on notera toutefois que les résultats de recherches actuelles de modélisation s'appuyant sur les modèles dynamiques adaptatifs suggèrent que les enfants pourraient inférer l'unité de rythme de leur langue maternelle à partir de propriétés acoustiques du signal de parole continue (McLennan, 2005).

ÉVALUATION DE L'HYPOTHÈSE D'INITIALISATION RYTHMIQUE DE LA SEGMENTATION

Dans la partie précédente, nous avons présenté un certain nombre d'études suggérant que, sur la base d'une sensibilité précoce au rythme, les enfants pourraient spécifier l'unité rythmique sur laquelle repose le rythme de leur langue, et utiliser cette unité pour commencer à segmenter les énoncés de leur langue maternelle. Comme nous l'avons dit précédemment, l'unité trochaïque de stress devrait être l'unité de segmentation rythmique précoce des langues à stress, et la syllabe devrait être l'unité de segmentation rythmique précoce des langues syllabiques, donc du français. Ainsi, nous prédisons qu'un mot bisyllabique français (par exemple, *toucan*), sera initialement segmenté comme deux unités syllabiques (par opposition à des unités moraiques plus petites ou des unités de stress plus grandes) indépendantes. En d'autres termes, nous prédisons que la segmentation rythmique, observable au moment où les procédures de segmentation se mettent en place, entraînera l'extraction de séquences multisyllabiques commençant par une syllabe accentuée pour les langues à stress comme l'anglais, et celle de syllabes isolées pour les langues syllabiques comme le français. Dans ce qui suit, nous évaluons d'abord l'état des connaissances, au moment où nous avons entrepris notre projet de recherche, permettant d'évaluer notre hypothèse développementale d'une segmentation rythmique précoce reposant sur l'unité rythmique de la langue maternelle. Dans un deuxième temps, nous présenterons les travaux que nous avons accomplis dans ce domaine sur le français, ainsi que ceux que nous entendons réaliser dans le futur.

Évaluation de la littérature sur la segmentation précoce à la lumière de l'hypothèse d'initialisation rythmique

Peu de données sont disponibles pour permettre d'évaluer notre hypothèse rythmique. En effet, les données disponibles qui sont compatibles avec nos hypothèses se limitent aux données mettant en évidence une segmentation trochaïque précoce chez les enfants acquérant l'anglais (Juszyk *et al.*, 1999b). L'absence de données autres s'explique en fait assez aisément. En effet, les quelques études ayant exploré la segmentation précoce dans des langues autres que l'anglais se sont jusqu'ici essentiellement focalisées sur l'établissement de l'âge d'émergence de ces capacités de segmentation, sans considération systématique des mécanismes impliqués et des types de mots ainsi segmentés correctement.

Ainsi, concernant les langues à stress, il a été trouvé que les enfants acquérant le néerlandais commencent à segmenter des mots trochaïques entre les âges de 7.5 et 9 mois (Houston *et al.*, 2000 ; Kooijman, Hagoort & Cutler, 2005 ; Kuijpers, Coolen, Houston & Cutler, 1998) ; toutefois, on ne sait pas quand les enfants commencent à segmenter les mots iambiques, ce qui devrait être ultérieurement selon nos hypothèses. D'autre part, on sait que les enfants acquérant l'allemand commencent à segmenter les mots monosyllabiques entre les âges de 6 et 8 mois (Höhle & Weissenborn, 2003), un résultat qui, à nouveau, ne permet pas d'évaluer nos hypothèses. Toutefois, des données préliminaires ont rapporté l'existence probable d'un biais de segmentation trochaïque en allemand, similaire à celui mis en évidence pour l'anglais (Höhle & Weissenborn, 2005).

Concernant les langues syllabiques, des données ne sont pour le moment disponibles que pour le français. Au moment où nous avons commencé nos recherches, deux études seulement avaient exploré la segmentation de la parole par des enfants acquérant le français, une avec des enfants parisiens (Gout, 2001) et une avec des enfants québécois (Polka & Sundara, 2003). Ces deux études, qui s'intéressaient essentiellement à la détermination du pattern d'émergence des capacités de segmentation en français et n'étaient pas destinées à tester nos hypothèses de segmentation rythmique précoce sur la base de la syllabe, offrent des résultats très contrastés.

Les données sur les enfants québécois font apparaître des capacités de segmentation de mots bisyllabiques dès l'âge de 8 mois (Polka & Sundara, 2003). Dans cette étude, les enfants étaient familiarisés avec deux mots bisyllabiques, puis ils entendaient deux passages contenant six répétitions d'un des mots familiarisés (un passage pour chaque mot) et deux passages contenant des répétitions de mots non-familiarisés. Les stimuli avaient été prononcés dans un registre de parlé-bébé très intonné, et enregistrés soit

par une locutrice québécoise, soit par une locutrice parisienne. Des résultats de segmentation similaires ont été obtenus pour les deux dialectes, à savoir une préférence pour les passages contenant les mots familiarisés. Cette étude suggère donc que les enfants québécois segmentent des mots bisyllabiques entiers dès l'âge de 8 mois (on notera toutefois, et nous y reviendrons, que l'on ne sait pas comment les enfants auraient réagi si on les avait familiarisés avec seulement l'une des syllabes des mots cibles plutôt que les mots entiers).

Ces résultats avec les enfants québécois sont en opposition avec les données obtenues avec les enfants acquérant le français parisien. En effet, Gout (2001) n'a pas pu établir que les enfants parisiens segmentent des mots bisyllabiques entre les âges de 7.5 et 11 mois, alors même que ces enfants semblaient segmenter des mots monosyllabiques dès l'âge de 7.5 mois (effet toutefois non-répliqué à 11 mois). Ceci pourrait s'expliquer par notre hypothèse rythmique de segmentation syllabique précoce dans les langues syllabiques, qui prédit que la segmentation de mots bisyllabiques comme unités ne se retrouve que dans une deuxième étape, les syllabes d'un même mot étant initialement segmentées de façon indépendante. Afin de tester plus spécifiquement cette hypothèse, l'étude suivante a été réalisée par Nazi, Iakimova, Bertoncini, Frédonie & Alcantara (2006a).

Syllabe et segmentation précoce en français

Pour tester la prédiction que la syllabe est l'unité de segmentation rythmique précoce en français, Nazi *et al.* (2006a) ont étudié la façon dont sont segmentés des mots bisyllabiques insérés dans des passages de parole continue (suivant la procédure précédemment utilisée pour l'anglais notamment, voir Jusczyk & Aslin, 1995 ; Jusczyk *et al.*, 1999b). Leur prédiction était qu'au moment de l'émergence des capacités de segmentation, la procédure de segmentation rythmique placerait des frontières entre chaque syllabe consécutive, et qu'aucune autre information (dont, par exemple, les informations distributionnelles) ne serait utilisée afin de regrouper les deux syllabes consécutives des mots bisyllabiques (ce qui serait compatible avec ce qui a été observé pour les mots bisyllabiques iambiques chez les enfants de 7.5 mois acquérant l'anglais, voir Jusczyk *et al.*, 1999b). En conséquence, les enfants acquérant le français ne segmenteraient pas initialement des mots bisyllabiques « entiers » de la parole continue ; plutôt, ils segmenteraient les deux syllabes de mots bisyllabiques comme des unités rythmiques individuelles et indépendantes. Quelques mois plus tard au cours du développement, le poids de l'utilisation des informations rythmiques pour segmenter la parole

continue aurait diminué en comparaison des autres indices, et la segmentation de mots bisyllabiques en unité à part entière serait observée comme le résultat de l'utilisation par les enfants d'autres indices (par exemple : informations distributionnelles ; indices phonotactiques).

Afin de tester ce scénario développemental, des enfants parisiens acquérant le français ont été testés avec quatre passages contenant un mot bisyllabique cible (chaque passage était constitué de 6 phrases, chaque phrase contenant une occurrence du mot cible donné). Les mots-cibles utilisés étaient *putois*, *toucan*, *bandeau* et *guidon*. Deux des passages correspondaient aux items présentés préalablement dans une phase de familiarisation. Les items de familiarisation différaient selon les expériences réalisées. Les enfants ont été familiarisés soit avec deux mots bisyllabiques entiers (par exemple, *putois* et *toucan*), soit avec leurs syllabes finales (par exemple, *tois* et *can*) ou leurs syllabes initiales (par exemple, *pu* et *tou*). Les enfants ont été testés à 8, 12 et 16 mois dans les deux premières conditions (mots entiers, syllabes finales), et à seulement 12 mois dans la dernière condition (syllabes initiales).

Tout d'abord, on notera qu'à 8 mois, aucun effet de segmentation n'est obtenu pour les deux types de cibles testées : mots entiers (Fig. 1, barres de gauche) et syllabes finales (Fig. 2, barres de gauche), un résultat qui suggère un délai possible dans l'émergence des capacités de segmentation pour le français comparé à l'anglais, une question sur laquelle nous revenons plus loin.

Les autres résultats sont conformes à notre hypothèse d'une segmentation rythmique initiale reposant sur la syllabe. En effet, à 12 mois, un effet de segmentation a été trouvé pour les syllabes finales (Fig. 2, barres centrales), mais pas pour les mots entiers (Fig. 1, barres centrales). À 16 mois par contre, des effets opposés sont obtenus : il n'y a plus d'effet de segmentation pour la syllabe finale (Fig. 2, barres de droite), mais un effet de segmentation pour les mots entiers (Fig. 1, barres de droite). Prises ensemble, ces données suggèrent qu'à 12 mois, les enfants acquérant le français segmentent indépendamment les deux syllabes d'un même mot bisyllabique, en dépit du fait que ces deux syllabes apparaissent toujours de façon consécutive dans le signal de parole. Ces informations distributionnelles (ainsi probablement que d'autres indices de segmentation, question qui reste à explorer pour le français) semblent par contre être prises en compte à l'âge de 16 mois, où les enfants montrent des effets de reconnaissance du mot entier mais plus de la syllabe finale isolée. On notera que ce changement de pattern de segmentation est similaire à ce qui a été trouvé pour la segmentation des mots iambiques en anglais entre les âges de 7.5 et 10.5 mois (Jusczyk *et al.*, 1999b).

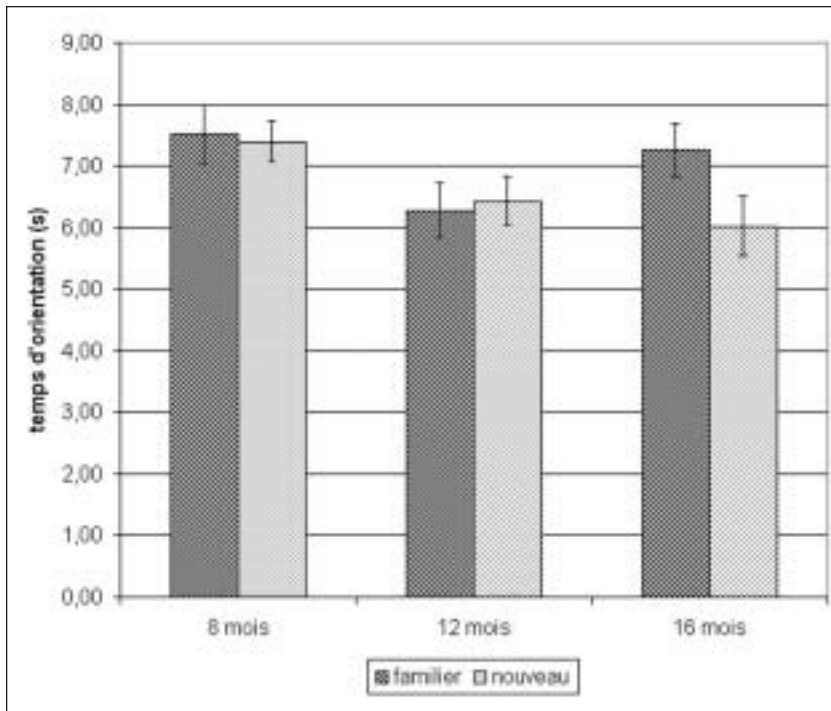


Figure 1. Résultats pour la condition de familiarisation « mots entiers ». Temps d’orientation moyens (et erreurs standards) pour les passages contenant les mots bisyllabiques familiarisés ou nouveaux, séparés selon l’âge des enfants (8, 12, 16 mois). [Extrait de Nazi *et al.*, 2006a, Expérience 1]

Figure 1. Results for the “whole words” familiarization condition. Mean orientation times (and standard errors) for the passages containing the familiarized vs. new bisyllabic words, broken down by age (8, 12, 16 months). [from Nazi *et al.*, 2006a, Experiment 1]

Concernant maintenant la syllabe initiale, notre hypothèse rythmique prédit que les enfants français (contrairement à ce que l’on trouverait en anglais, bien que cette prédiction n’ait pas été testée directement dans cette langue) devraient également extraire la syllabe initiale de mots multisyllabiques de façon isolée, au même âge qu’ils extraient de façon isolée les syllabes finales. Étant donné les résultats précédents sur les syllabes finales, montrant un effet de segmentation syllabique seulement à 12 mois, nous n’avons testé les enfants avec les syllabes initiales qu’à 12 mois. Dans la première expérience réalisée, les enfants étaient familiarisés avec des

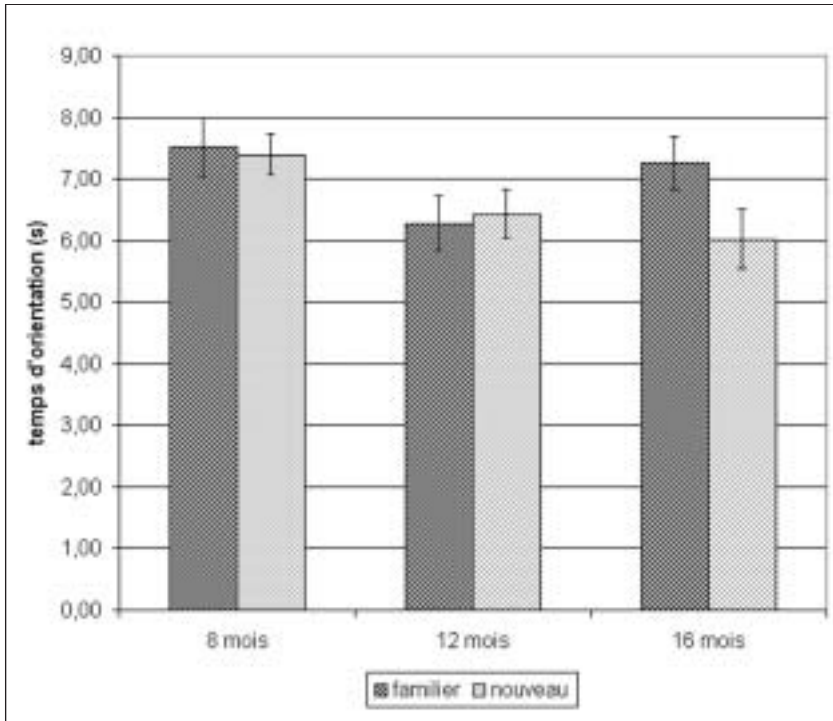


Figure 2. Figure 2. Résultats pour la condition de familiarisation « syllabes finales ». Temps d'orientation moyens (et erreurs standards) pour les passages contenant les mots bisyllabiques correspondant aux syllabes familiarisées ou les mots nouveaux, séparés selon l'âge des enfants (8, 12, 16 mois). [Extrait de Nazzi *et al.*, 2006a, Expérience 2]

Figure 2. Results for the “final syllables” familiarization condition. Mean orientation times (and standard errors) for the passages corresponding to the familiarized syllables vs. the new words, broken down by age (8, 12, 16 months). [from Nazzi *et al.*, 2006a, Experiment 2]

exemplaires de syllabes initiales enregistrées en isolation, comme cela avait été fait avec les autres types d'items de familiarisation. Dans cette condition, les enfants ont tendance à préférer écouter les passages correspondant aux items de familiarisation, mais cet effet, qui signale une segmentation/reconnaissance, n'est pas significatif, ce qui a priori est contraire à nos prédictions (voir Fig. 3, barres de gauche).

Toutefois, une analyse des caractéristiques acoustiques des items de familiarisation et de test utilisés a fait apparaître deux éléments pouvant

expliquer les effets réduits pour les syllabes initiales : (1) ces syllabes sont moins saillantes dans les passages que les syllabes finales, et (2) la distance acoustique entre les items utilisés en familiarisation (enregistrés en isolation) et ceux utilisés dans la phase de test était plus grande pour les syllabes initiales que pour les syllabes finales. Afin de déterminer si ces facteurs, et notamment le deuxième, peuvent avoir influencé les résultats obtenus, une deuxième expérience a été réalisée.

Dans cette dernière expérience, les syllabes initiales utilisées en familiarisation ont été extraites des passages test, supprimant ainsi la distance acoustique entre les items de familiarisation et leur réalisation dans les passages. Dans cette condition, un effet significatif de segmentation est obtenu (voir Fig. 3, barres de droite). Plusieurs conclusions peuvent être tirées de ce résultat. Tout d'abord, il apparaît, en accord avec nos prédictions de segmentation rythmique précoce (basée donc sur la syllabe en français), que les enfants français segmentent dans un premier temps, à 12 mois, les différentes syllabes de mots bisyllabiques de façon isolée ; puis, dans un deuxième temps, à 16 mois, ils segmentent les mots bisyllabiques comme des entités à part entière. Ces résultats suggèrent aussi (1) que les enfants sont sensibles à la distance acoustique entre items de familiarisation et de test (sur la base des résultats et des analyses acoustiques), et (2) que les jeunes enfants acquérant le français sont d'une certaine façon sensibles aux niveaux d'accentuation lexicale.

En résumé, l'étude de *Nazzi et al.* (2006a) apporte les premières données validant l'hypothèse, proposée en partie sur la base de nos travaux précédents sur la sensibilité précoce au rythme linguistique (*Nazzi et al.*, 1998a, 2000), que les enfants acquérant le français s'appuient initialement sur l'unité rythmique du français (la syllabe) pour segmenter le signal de parole, de façon similaire au fait que les enfants acquérant l'anglais s'appuient sur l'unité rythmique de l'anglais (l'unité trochaïque, correspondant à une séquence d'une ou plusieurs syllabes dont seule la première est accentuée). En effet, de même que les enfants anglophones commencent à segmenter les unités trochaïques avant les unités iam-biques (7.5 contre 10.5 mois), les enfants francophones segmentent isolément les syllabes avant de segmenter les mots bisyllabiques en unités cohérentes (12 contre 16 mois).

Finalement, on notera que si les données obtenues avec les enfants québécois de 8 mois semblent en contradiction avec nos données avec les enfants parisiens (*Nazzi et al.*, 2006a), la segmentation syllabique n'a en fait pas été testée par *Polka et Sundara* (2003). Il reste donc possible que les enfants québécois reconnaissent les syllabes individuelles des mots bisyllabiques à 8 mois, ou à un âge antérieur, une possibilité à tester dans le futur.

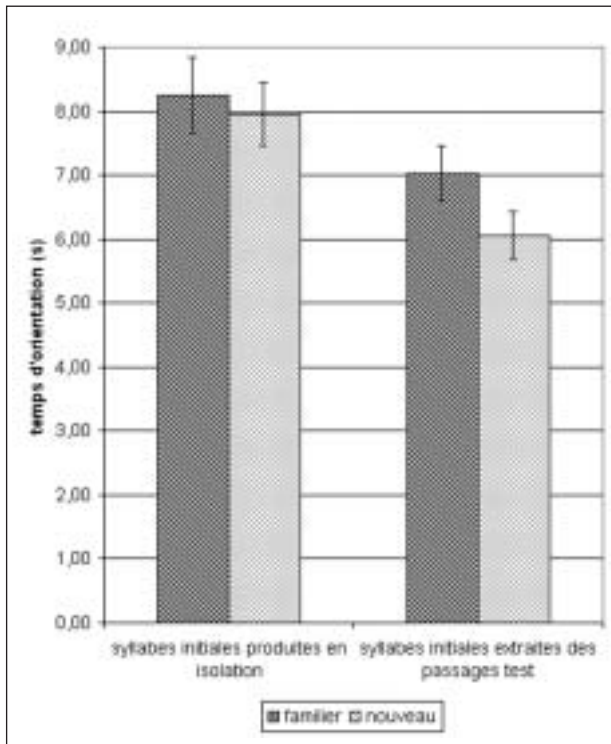


Figure 3. Figure 3. Résultats pour la condition de familiarisation « syllabes initiales ». Temps d’orientation moyens (et erreurs standards) pour les passages contenant les mots bisyllabiques correspondant aux syllabes familiarisées ou les mots nouveaux, séparés selon que les syllabes initiales familiarisées ont été enregistrées en isolation ou extraites des passages de test. [Extrait de Nazzi *et al.*, 2006a, Expériences 3-4].

Figure 3. Results for the “initial syllables” familiarization condition. Mean orientation times (and standard errors) for the passages corresponding to the familiarized syllables vs. the new words, broken down according to whether the familiarized syllables had been recorded in isolation or extracted from the test passages. [from Nazzi *et al.*, 2006a, Experiments 3-4]

CONCLUSIONS ET PERSPECTIVES

Les études présentées ci-dessus suggèrent que les enfants acquérant le français commencent à segmenter en utilisant l'unité rythmique de leur langue maternelle (la syllabe), de même que les enfants anglais commencent avec l'unité rythme de leur langue (l'unité trochaïque). Cette antériorité de l'utilisation de l'unité rythmique pour segmenter fait à la fois apparaître des similarités (même utilisation du rythme) et des différences (utilisation d'unités rythmiques différentes) dans le pattern d'émergence des capacités de segmentation entre langues. Le fait que différentes unités rythmiques soient utilisées pour segmenter dès le moment de l'émergence des capacités de segmentation suggère deux choses : (1) l'acquisition de l'unité rythmique se fait avant l'émergence des capacités de segmentation, donc probablement de façon indépendante du niveau lexical ; et (2) cette acquisition de l'unité rythmique pourrait être le facteur permettant l'émergence des capacités de segmentation. Ces données sont donc en accord avec l'hypothèse que nous avons proposée d'une initialisation rythmique des procédures de segmentation.

Avant de terminer, soulignons que de nombreuses questions concernant la mise en place des procédures de segmentation en français restent posées. Deux de ces questions sont brièvement discutés ci-dessous.

La première question porte sur l'élucidation des différences de résultats obtenus dans les études de Nazi *et al.* (2006a) et Polka et Sundara (2003). En effet, un effet de segmentation de mots bisyllabiques n'émerge qu'entre 12 et 16 mois chez les enfants parisiens, alors qu'un effet compatible à une segmentation de mots bisyllabiques est trouvé dès l'âge de 8 pour les enfants québécois. Comme nous venons de le mentionner, l'étude avec les enfants québécois n'a pas exploré la possibilité que les enfants utilisent aussi une procédure syllabique, ni à 8 mois, ni à un âge antérieur. De ce fait, il ne peut encore être exclu que ces enfants commencent aussi par segmenter sur la base de la syllabe. Dans cette perspective, les différences observées entre les deux populations d'enfants francophones pourraient résulter du fait que les enfants québécois auraient pu utiliser plus précocement d'autres informations pour regrouper les deux syllabes d'un même mot bisyllabique en une unité cohérente. La question qui se pose alors est de déterminer si cet avantage des enfants québécois est dû à des différences dans les stimuli utilisés dans les deux études (les stimuli utilisés par Polka et Sundara, 2003, étant globalement plus intonnés, plus parlé-bébé, ce qui aurait pu faciliter la segmentation, vois aussi Thiessen, Hill & Saffran, 2005), ou s'il existe des différences prosodiques/phonologiques entre les dialectes québécois et

parisiens ayant des répercussions sur le développement des capacités de segmentation dans ces deux dialectes. À ce propos, l'existence d'une plus grande variabilité d'intonation en québécois qu'en français parisien (Ménard, Ouellon & Dolbec, 1999) et les données suggérant que les variations de pitch affectent la segmentation (Nazzi *et al.*, 2005) sont compatibles avec une segmentation plus précoce par les enfants québécois. Des études sont actuellement en cours pour évaluer ces questions (Nazzi, Iakimova, Bertocchini, Polka, Sundara & Girard, 2006b).

Une autre question a trait au fait que les études sur le français n'ont à l'heure actuelle porté que sur l'utilisation de la prosodie pour segmenter. Or les études sur l'émergence de la segmentation en anglais ont fait apparaître l'utilisation par les enfants, de façon parallèle, de plusieurs indices de segmentation (prosodie, informations distributionnelles, allophonie, phonotactique, mots connus...). Ces études ont également fait apparaître que le poids relatif accordé à ces différents indices évoluait au cours du développement. Étant donné que le décours développemental d'utilisation de ces indices devrait varier suivant les langues, en fonction entre autre de leur saillance relative, il est impératif d'étudier l'utilisation par les enfants acquérant le français de ces différents indices. Ceci est d'autant plus crucial que nous avons fait l'hypothèse que les changements observés par Nazzi *et al.* (2006) entre 12 et 16 mois correspondraient à un plus grand poids accordé à des indices non prosodiques. Ces études permettront donc de continuer à tracer le pattern d'émergence des capacités de segmentation en français, et de continuer à tester l'hypothèse de Mattys, White et Melhorn (2005) selon laquelle les enfants commencent par utiliser des indices sub-lexicaux suprasegmentaux (prosodie), puis des indices sub-lexicaux segmentaux (phonotactique, allophonie, coarticulation), et enfin des indices lexicaux (sémantique, mais aussi syntaxe et pragmatique).

BIBLIOGRAPHIE

- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: University of Edinburgh Press.
- Arvaniti, A. (1994). Acoustic features of Greek rhythmic structure. *Journal of Phonetics*, 22, 239-268.
- Aslin, R.N. (1993). Segmentation of fluent speech into words: learning models and the role of maternal input. In B. de Boysson-Bardies *et al.* (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 305-315). Kluwer.
- Bortfeld, H., Morgan, J.L., Golinkoff, R.M. & Rathbun, K. (2005). Mommy and me: familiar names help launch babies into

- speech-stream segmentation. *Psychological Science*, 16, 298-304.
- Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, 65, 33-69.
- Brent, M.R., & Cartwright, T.A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, 61, 93-125.
- Brent, M.R., & Siskind, J.M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, 81, B33-B44.
- Cassidy, K.W., & Kelly, M.H. (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language*, 30, 348-369.
- Chevrot, J. P., Dugua, C., & Fayol, M. (2005). Liaison et formation des mots en français : un scénario développemental. *Langages*, 158, 38-52.
- Christiansen, M.H., Allen, J., & Seidenberg, M.S. (1998). Learning to segment speech using multiple cues : A connectionist model. *Language and Cognitive Processes*, 13, 221-268.
- Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (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.
- Cole, R.A. & Jakimik, J. (1978). Understanding speech : How words are heard. In G. Underwood (Ed.), *Strategies of information processing* (pp.). New York, NY : Academic Press.
- Cole, R.A., & Jakimik, J. (1980). How are syllables used to recognize words ? *Journal of the Acoustical Society of America*, 67, 965-970.
- Content, A., Meunier, C., Kearns, R.K., & Frauenfelder, U.H. (2001). Sequence detection in pseudowords in French : Where is the syllable effect ? *Language and Cognitive Processes*, 16, 609-636.
- Curtin, S., Mintz, T.H., & Christiansen, M.H. (2005). Stress changes the representational landscape : evidence from word segmentation. *Cognition*, 96, 233-262.
- Cutler, A. & Carter, D.M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, 2, 133-142.
- Cutler, A., Mehler, J., Norris, D. G., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, 25, 385-400.
- Cutler, A., Mehler, J., Norris, D. G., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, 24, 381-410.
- Cutler, A. & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology : Human Perception and Performance*, 14, 113-121.
- Dahan, D., & Brent, M. R. (1999). On the discovery of novel wordlike units from utterances : An artificial-language study with implications for native-language acquisition. *Journal of Experimental Psychology : General*, 128 (2), 165-185.
- Delattre, P.C. (1966). A comparison of syllable length conditioning among languages. *International Journal of Applied Linguistics*, 4, 182-198.
- den Os, E. (1988). *Rhythm and tempo of Dutch and Italian : a contrastive study*. Utrecht, Holland : Drukkerij Elinkwijk BV.
- Echols, C.H., Crowhurst, M.J. & Childers, J.B. (1997). The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language*, 36, 202-225.
- Fant, G., Kruckenberg, A., & Nord, L. (1991). Durational correlates of stress in Swedish, French and English. *Journal of Phonetics*, 19, 351-365.
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10, 279-283.
- Fletcher, J. (1991). Rhythm and final lengthening in French. *Journal of Phonetics*, 19, 193-212.

- Friederici, A.D. & Wessels, J.M.I. (1993). Phonotactic knowledge and its use in infant speech perception. *Perception & Psychophysics*, 54, 287-295.
- Gout, A., (2001). *Étapes précoces de l'acquisition du lexique*. Unpublished Dissertation, Ecole des Hautes Etudes en Sciences Sociales, Paris, France.
- Gout, A., Christophe, A., & Morgan, J.L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language*, 51, 548-567.
- Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, 26, 269-286.
- Hockema, S.A. (2006). Finding Words in Speech : An Investigation of American English. *Language Learning and Development*, 2, 119-146.
- Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, 6, 122-127.
- Höhle, B., & Weissenborn, J. (2005). Word segmentation in German-learning infants. Paper presented at the International workshop "Early word segmentation : a crosslinguistic approach taking advantage of Europe's linguistic diversity," Paris, France, February 25-26.
- Hohne, E.A., & Jusczyk, P.W. (1994). Two-month-old infants' sensitivity to allophonic differences. *Perception & Psychophysics*, 56, 613-623.
- Houston, D.M., Jusczyk, P.W., Kuijpers, C., Coolen, R., Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomics Bulletin & Review*, 7, 504-509.
- Houston, D.M., Santelmann, L.M., & Jusczyk, P.W. (2004). English-learning infants' segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes*, 19, 97-136.
- Johnson, E.K. & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44, 1-20.
- Jusczyk, P.W. & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29 : 1-23.
- Jusczyk, P.W., Cutler, A. & Redanz, N. (1993a). Preference for the predominant stress patterns of English words. *Child Development*, 64, 675-687.
- Jusczyk, P.W., Friederici, A.D., Wessels, J., Svenkerud, V.Y. & Jusczyk, A.M. (1993b). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, 32, 402-420.
- Jusczyk, P.W., Hohne, E.A., & Bauman, A. (1999a). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics*, 62, 1465-1476.
- Jusczyk, P.W., Houston, D.M., & Newsome, M. (1999b). The beginning of word segmentation in English-learning infants. *Cognitive Psychology*, 39, 159-207.
- Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D. G., Kennedy, L., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, 24, 252-293.
- Jusczyk, P.W., Luce, P.A. & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, 33, 630-645.
- Karzon, R.G., & Nicholas, J.G. (1989). Syllabic pitch perception in 2- to 3-month-old infants. *Perception and Psychophysics*, 45, 10-14.
- Kelly, M.H., & Bock, J.K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 389-403.
- Klatt, D.H. (1979). Speech perception : A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, 7, 279-312.
- Klatt, D.H. (1989). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical Representation and Process* (pp. 169-226). Cambridge, MA : MIT Press.

- Kooijman, V., Hagoort, P., & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research*, 24, 109-116.
- 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.
- Kuijpers, C., Coolen, R., Houston, D., & Cutler, A. (1998). Using the head-turning technique to explore cross-linguistic performance differences. *Advances in Infancy Research (Vol. 12)*, (pp. 205-220). Stamford, CT : Ablex.
- Mattys, S., Jusczyk, P.W., Luce, P.A., & Morgan, J.L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, 38, 465-494.
- Mattys, S., & Jusczyk, P.W. (2001a). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, 78, 91-121.
- Mattys, S., & Jusczyk, P.W. (2001b). Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology : Human Perception and Performance*, 27, 644-655.
- Mattys, S.L., White, L., Melhorn, J.F. (2005). Integration of Multiple Speech Segmentation Cues: A Hierarchical Framework. *Journal of Experimental Psychology : General*, 134, 477-500.
- McLennan, S. (2005). A dynamic adaptive model of linguistic rhythm. In *Proceedings of the I.S.C.A. Workshop on Plasticity in Speech Perception* (p. 130), London, U.K., June 15-17 2005.
- McQueen, J.M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition : Spotting words in other words. *Journal of Experimental Psychology : Learning, Memory and Cognition*, 20, 621-638.
- Mehler, J., Dommergues, J. Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, 20, 298-305.
- Mehler, J., Dupoux, E., Nazzi, T. & Dehaene-Lambertz, G. (1996). Coping with linguistic diversity : The infant's viewpoint. Dans J.L. Morgan & K. Demuth (Eds.), *Signal to syntax* Mahwah, N.J. : Lawrence Erlbaum Associates. 101-116.
- Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertocini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 144-178.
- Ménard, L., Ouelon, C., & Dolbec, J. (1999). Prosodic markers of regional group membership : The case of French of Quebec versus France. In *Proceedings of the XIVth International Congress of Phonetic Sciences (ICPhS99)*, San Francisco (pp. 1601-1604).
- Moon, C., Panneton-Cooper, R. & Fifer, W.P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, 16, 495-500.
- Morgan, J.L. & Saffran, J.R. (1995). Emerging integration of sequential and supra-segmental information in preverbal speech segmentation. *Child Development*, 66, 911-936.
- Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words : differences between consonants and vowels. *Cognition*, 98, 13-30.
- Nazzi, T., Bertocini, J., & Mehler, J. (1998). Language discrimination by newborns : Towards an understanding of the role of rhythm. *Journal of Experimental Psychology : Human Perception and Performance*, 24, 756-766.
- Nazzi, T., Dille, L.C., Jusczyk, A.M., Shattuck-Hunagel, S., & Jusczyk, P.W. (2005). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech*, 48, 279-298.
- Nazzi, T., Floccia, C., & Bertocini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior and Development*, 21, 779-784.
- Nazzi, T., Iakimova, I., Bertocini, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French : emerging evidence for

- crosslinguistic differences. *Journal of Memory and Language*, 54, 283-299.
- Nazzi, T., Iakimova, G., Bertoni, J., Polka, L., Sundara, M., & Girard, S. (2006). The development of word segmentation in infants acquiring different dialects of French. Poster présenté à la 31st Annual Boston University Conference on Language Development, Boston MA, USA, 3-5 novembre 2006.
- Nazzi, T., Jusczyk, P.W., & Johnson, E.K. (2000). Language discrimination by English learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, 43, 1-19.
- Nazzi, T., Kemler Nelson, D.G., Jusczyk, P.W., & Jusczyk, A.M. (2000). Six-month-olds' detection of clauses in continuous speech: Effects of prosodic well-formedness. *Infancy*, 1, 123-147.
- Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, 41, 233-243.
- Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, 32, 258-278.
- Peretz, I., Lussier, I., & B el and, R. (1998). The differential role of syllabic structure in stem completion for French and English. *European Journal of Cognitive Psychology*, 10, 75-112.
- Perruchet, P., & Vinter, A. (1998). PARSER: a model for word segmentation. *Journal of Memory and Language*, 39, 246-263.
- Pike, K. (1945). *The intonation of American English*. Ann Arbor, MI: University of Michigan Press.
- Polka, L. & Sundara, M. (2003). Word segmentation in monolingual and bilingual infant learners of English and French. *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, 1021-1024.
- Ramus, F. (2002). Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition*, 2, 85-115.
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, 288, 349-351.
- Ramus, F., Nesp or, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73, 265-292.
- Saffran, J.R., Aslin, R.N. & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.
- Sebasti an-Gall es, N. & Bosch, L. (2002). Building phonotactic knowledge in bilinguals: role of early exposure. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 974-989.
- Sebasti an-Gall es, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language*, 31, 18-32.
- Segui, J. & Ferrand, L. (2000). *Le ons de parole*. Paris: Odile Jacob.
- Shafer, V. L., Shucard, D. W., & Jaeger, J. J. (1999). Electrophysiological indices of cerebral specialization and the role of prosody in language acquisition in three-month-old infants. *Developmental Neuropsychology*, 15, 73-109.
- Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P.W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, 49, 249-267.
- Spring, D. R. & Dale, P. S. (1977). Discrimination of linguistic stress in early infancy. *Journal of Speech and Hearing Research*, 20, 224-232.
- Thiessen, E.D., Hill, E.A., & Saffran, J.R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, 7, 53-71.
- Thiessen, E.D. & Saffran, J.R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, 39, 706-716.

- Tincoff, R., & Jusczyk, P. W. (1996). Are word-final sounds perceptually salient for infants? Paper presented at the Fifth Conference on Laboratory Phonology, Evanston, IL, July, 1996
- Turk, A.E., Jusczyk, P.W., Gerken, L. (1995). Do English-learning infants use syllable weight to determine stress? *Language and Speech*, 38, 143-158.
- van de Weijer, J. (1998). *Language input for word discovery*. Doctoral dissertation, University of Nijmegen (MPI Series in Psycholinguistics, 9).
- Vroomen, J., van Zon, M., & de Gelder, B. (1996). Metrical segmentation and inhibition in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 98-108.
- Wauquier-Gravelines, S. (2002). Segmentation de la chaîne parlée et détermination des frontières gauches dans l'acquisition du lexique précoce. *Rencontres Internationales du GDR Phonologies, Grenoble, 6-8 juin 2002*.

