

# La discrimination des langues par la prosodie : Modélisation linguistique et études comportementales

Franck Ramus<sup>1</sup>

Laboratoire de Sciences Cognitives et Psycholinguistique (EHESS/CNRS)  
ramus@lscp.ehess.fr

## Abstract

Spoken languages have been classified by linguists according to their rhythmic properties. Although researchers have measured many speech signal properties, they have failed to identify reliable acoustic characteristics for language classes. This paper presents instrumental measurements based on a consonant/vowel segmentation for eight languages. The measurements suggest that intuitive rhythm types reflect specific phonological properties, which in turn are signaled by the acoustic/phonetic properties of speech. The data support the notion of rhythm classes and also allow the simulation of language discrimination experiments with human subjects. Four such experiments are reported and establish the overall consistency of the model. Consequences for automatic language identification are considered.

## Résumé

Les langues du monde ont été classées par les linguistes selon leurs propriétés rythmiques. Bien que de nombreuses mesures aient été effectuées sur le signal de parole, aucune n'a permis de rendre compte correctement des classes rythmiques de langues. Dans cet article nous proposons des mesures basées sur une segmentation de la parole en consonnes/voyelles effectuée en huit langues. Ces mesures suggèrent que les types de rythme reflètent des propriétés phonologiques précises, qui sont elles-mêmes détectables au niveau acoustique/phonétique. Nos données sont compatibles avec la notion de classes de rythme, et permettent la simulation d'expériences de discrimination de langues chez des sujets humains. Quatre expériences sont présentées qui renforcent la cohérence globale du modèle. Des applications à l'identification automatique des langues sont envisagées.

## 1 Introduction

Le rythme de la parole semble être un bon moyen de caractériser les langues du monde et de les classer, au moins en un petit nombre de groupes. En effet, les linguistes ont traditionnellement distingué les langues accentuelles (*stress-timed*), englobant notamment les langues germaniques, slaves, ainsi que l'arabe, et les langues syllabiques (*syllable-timed*), comprenant les langues latines, ou encore le yoruba et le telegu [Abercrombie 1967, Pike 1945]. Un troisième groupe, les langues moraiques (*mora-timed*), comprenant le japonais ou le tamoul, a également été proposé [Ladefoged 1975]. Il était supposé que toutes les langues du monde avaient une organisation rythmique bien déterminée, appartenant à l'une de ces trois classes.

L'intuition derrière cette classification était que la production de la parole repose sur la répétition d'unités semblables, comme le pied, la syllabe ou la more, chaque langue utilisant un seul type d'unité, d'où l'existence de trois classes distinctes. Il était par ailleurs supposé que ces unités se répétaient à intervalles réguliers, les accents toniques étant régulièrement espacés

---

<sup>1</sup> Je remercie Marina Nespor et Jacques Mehler pour leur collaboration, la Délégation Générale pour l'Armement pour son soutien financier, et les participants de la 1<sup>ère</sup> journée sur l'identification automatique des langues pour une discussion intéressante de mes résultats.

dans les langues accentuelles, et de même pour les syllabes dans les langues syllabiques et les mores dans les langues moraiques: c'est l'hypothèse d'*isochronie*. Suivant cette hypothèse, il serait possible, en mesurant les durées séparant les accents, les syllabes ou les mores dans un échantillon d'une langue, de déterminer la classe rythmique de celle-ci. Malgré de nombreuses recherches, cette hypothèse n'a pas été validée empiriquement, les accents n'étant pas plus régulièrement espacés dans les langues accentuelles que dans les langues syllabiques, ni vice versa pour les syllabes [Bolinger 1965, Dauer 1983, Roach 1982]. A ce stade, la caractérisation rythmique des langues est donc assez incertaine. Dans ce qui suit nous proposons une nouvelle approche de cette caractérisation.

## 2 Corrélat du rythme dans le signal de parole<sup>2</sup>

### 2.1 Bases phonologiques du rythme

Notre approche repose sur une conception du rythme de parole, non plus comme primitive de l'organisation temporelle des langues, mais comme conséquence de certaines de leurs propriétés phonologiques [Bertinetto 1981, Dasher 1982, Dauer 1983], notamment: la complexité des syllabes, la corrélation entre poids syllabique et accent, la présence ou non de réduction vocalique... Selon cette conception, les langues dites syllabiques sont des langues n'autorisant que des syllabes simples et n'admettant pas de réduction vocalique. Les syllabes sont donc de taille relativement stable, donnant ainsi l'impression d'un rythme syllabique régulier. Dans les langues accentuelles, au contraire, des syllabes complexes sont autorisées, et celles-ci portent en général l'accent tonique. Les syllabes plus simples, en revanche, ne sont pas accentuées, et font au contraire l'objet d'une réduction vocalique, accentuant le contraste entre les syllabes fortes et les syllabes faibles, ce qui induit un rythme syllabique moins régulier, porté par les seules syllabes accentuées.

### 2.2 Etude empirique du rythme en 8 langues

L'approche phonologique du rythme de parole prédit qu'une analyse de la complexité syllabique d'une langue devrait permettre de déterminer sa classe rythmique. C'est ce que nous nous proposons de tester, en mesurant la complexité syllabique par le biais d'une segmentation de la parole en consonnes/voyelles.

Le matériel utilisé a été extrait du corpus multilingue du LSCP [Nazzi 1997]. Huit langues (Anglais, Néerlandais, Polonais, Français, Espagnol, Italien, Catalan, Japonais) et quatre locutrices natives de chaque langue ont été choisies. Un corpus de 160 phrases a été constitué, 20 par langue, 5 par locutrice. Les phrases ont été sélectionnées pour avoir un nombre de syllabes et des durées comparables à travers les langues. Les phonèmes de chaque phrase ont été marqués manuellement et alignés avec le signal de parole, puis classifiés en consonne ou voyelle<sup>3</sup>. Afin de mesurer plus directement la complexité syllabique, nous ne nous sommes pas intéressés aux durées des phonèmes individuels, mais aux durées des intervalles vocaliques (du début à la fin d'une séquence de voyelles) et consonantiques (du début à la fin d'une séquence de consonnes). La complexité syllabique est donc capturée notamment par la durée des groupes consonantiques.

De ces durées nous avons dérivé 3 variables prenant une valeur par phrase:

- %V la proportion (en durée) d'intervalles vocaliques dans la phrase<sup>4</sup>;

---

<sup>2</sup> . Une partie des données présentées dans cette section sont tirées de F. Ramus, M. Nespors, & J. Mehler, "Correlates of linguistic rhythm in the speech signal," *Cognition*, (sous presse)

<sup>3</sup> Les semi-voyelles pre- ou inter-vocaliques ont été considérées comme des consonnes, les semi-voyelles post-vocaliques comme des voyelles.

<sup>4</sup> La proportion d'intervalles consonantiques est égale à 1-%V et n'apporte donc pas d'information supplémentaire.

- $\Delta V$  l'écart type des durées d'intervalles vocaliques par phrase;
- $\Delta C$  l'écart type des durées d'intervalles consonantiques.

La figure 1 donne les valeurs moyennes de %V,  $\Delta C$  et  $\Delta V$  par langue, les barres d'erreur représentant l'erreur standard de la moyenne.

**Figure 1**

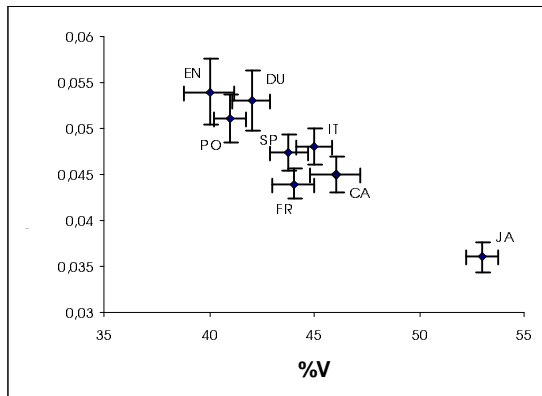


Figure 1a. Plan (%V,  $\Delta C$ )

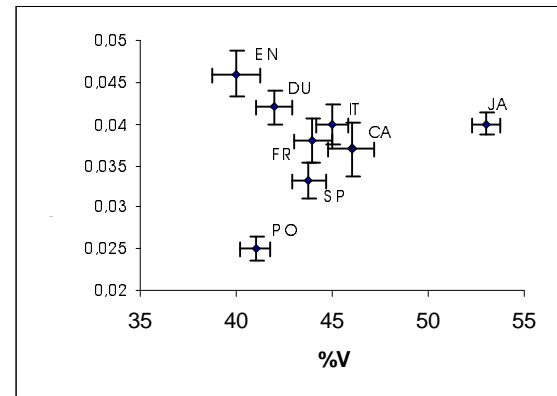


Figure 1b. Plan (%V,  $\Delta V$ )

Le plan (%V,  $\Delta C$ ) fait ressortir clairement 3 groupes, et il s'avère que ces groupes correspondent aux classes rythmiques décrites dans la littérature : Anglais, Néerlandais et Polonais pour les langues accentuelles, Espagnol, Italien, Français et Catalan pour les langues syllabiques, et Japonais pour les langues moraiques. Une ANOVA introduisant un facteur « classe de rythme » montre un effet de classe significatif à la fois pour %V et  $\Delta C$  ( $p < 0.001$ ). Des comparaisons multiples (Tukey test) montrent par ailleurs que chaque classe est significativement différente des deux autres, pour %V ( $p < 0.001$ ) et  $\Delta C$  ( $p < 0.005$ ). Il semble donc que les variables %V et  $\Delta C$  capturent l'essentiel des propriétés rythmiques de ces langues, et cela peut s'interpréter par les propriétés phonologiques mentionnées ci-dessus. En effet, plus une langue admet des syllabes complexes, plus elle admet des groupes de consonnes de taille importante, donc plus le rapport consonne/voyelle augmente. En conséquence les langues accentuelles ont un faible %V. De plus, les langues admettant des syllabes complexes admettent aussi les syllabes plus simples, ce qui implique que les groupes consonantiques peuvent être de taille plus variable que dans les langues n'admettant que des syllabes simples. Les langues accentuelles ont donc un  $\Delta C$  plus élevé que les langues syllabiques.

Ainsi que le montre le plan (%V,  $\Delta V$ ), la variable  $\Delta V$  semble moins directement liée aux classes de rythme (pas d'effet de classe significatif). Par rapport à  $\Delta C$ ,  $\Delta V$  apporte un élément supplémentaire d'information, suggérant que le Polonais a des différences importantes avec les autres langues accentuelles. De fait, le statut rythmique du Polonais ne fait pas l'unanimité, certains auteurs le considérant comme une langue syllabique ou encore intermédiaire [Nespor 1990, Rubach 1985]. Ce désaccord est dû à l'absence de réduction vocalique en Polonais, alors que cette propriété est généralement associée aux langues accentuelles. Cela se traduit dans nos données par un plus faible  $\Delta V$  pour le Polonais (moins de variations des durées de voyelles).

### 2.3 Analyse discriminante

Jusqu'ici nous avons montré que, malgré l'échec des recherches sur l'isochronie, il est possible de trouver dans le signal de parole des corrélats phonétiques des classes rythmiques. On peut maintenant se demander si, étant donné une phrase, il est possible de déterminer

fiablement la classe rythmique de la langue en question. Dans ce but nous avons effectué une analyse discriminante avec cross-validation. Dans cette procédure, chaque phrase est classifiée à l'aide de fonctions discriminantes calculées sur l'ensemble des données sauf la phrase en question. Ici nous utilisons exclusivement les variables %V et  $\Delta C$ , qui sont les mieux reliées aux classes de rythme. Les pourcentages de classifications sont présentés dans le tableau 1, le niveau du hasard étant 33%.

**Tableau 1**

classe réelle	classe prédite		
	moraique	syllabique	accentuelle
moraique	95%	5%	0%
syllabique	13.8%	53.8%	32.5%
accentuelle	5%	31.7%	63.3%

Ainsi, il est possible, étant donnée une seule phrase (environ 3 s.) d'une langue indéterminée, de prédire avec une bonne fiabilité la classe rythmique à laquelle elle appartient. Afin de mieux apprécier la valeur absolue des scores rapportés ici, notons que :

- Les modèles de classes de langue sont calculés à partir de 20 phrases (environ 60 secondes) par langue, ce qui est peu. Un corpus plus important permettrait l'élaboration de meilleurs modèles et donc une meilleure classification.
- Les échantillons de parole testés sont relativement courts (une phrase, en moyenne 3 secondes). L'utilisation d'échantillons plus longs permettrait une évaluation plus précise de %V et  $\Delta C$  pour chaque échantillon et améliorerait donc également la classification.

Le même type d'analyse peut être restreint à des paires de langues et ainsi prédire si la discrimination est possible pour une paire donnée. Le tableau 2 donne les résultats de classification correcte prédits par l'analyse discriminante pour les paires de langues testées dans la section suivante (hasard = 50%).

**Tableau 2**

Paire de langues	Variables	Pourcentage de phrases correctement classifiées
Anglais/Japonais	% V, $\Delta C$	95%
Anglais/Espagnol	% V, $\Delta C$	62.5%
Polonais/Espagnol	% V, $\Delta C$	62.5%
Anglais/Polonais	% V, $\Delta C$	40%
Anglais/Polonais	% V, $\Delta V$	90%

Comme la figure 1a le suggère, les variables %V et  $\Delta C$  prédisent une discrimination lorsque les langues appartiennent à 2 classes différentes, mais pas pour 2 langues de la même classe (Anglais/Polonais). Néanmoins, si l'on incorpore  $\Delta V$  dans le modèle, l'Anglais et le Polonais deviennent alors discriminables. Dans un but de discrimination automatique,  $\Delta V$  apporte donc une information supplémentaire. Dans un but de modélisation de résultats comportementaux, l'intérêt de  $\Delta V$  reste une question empirique, que l'on peut aborder en testant la discrimination Anglais/Polonais.

### 3 Résultats comportementaux de discrimination de langues

Les données décrites ci-dessus prédisent qu'il devrait être facile de distinguer des langues appartenant à des classes rythmiques différentes, et difficile de distinguer des langues appartenant à la même classe, si toutefois l'on ne dispose que du rythme. C'est également l'hypothèse faite par [Mehler 1996] pour le nouveau-né, étant supposé que ce dernier est

particulièrement sensible à la prosodie. Cette hypothèse a été étayée par de nombreuses expériences chez le nouveau-né, montrant que celui-ci pouvait distinguer le Français du Russe, l'Anglais de l'Italien, de l'Espagnol ou du Japonais, mais pas l'Anglais du Néerlandais [Mehler 1988, Moon 1993, Nazzi 1998] Il a même été montré que la discrimination était possible au niveau de la classe, des nouveau-nés discriminant un mélange d'Anglais et Néerlandais contre un mélange d'Espagnol et d'Italien, mais pas Anglais + Italien contre Néerlandais + Espagnol [Nazzi 1998]. La plupart des expériences sus-citées ayant été effectuées ou répliquées avec de la parole filtrée, il semble que le nouveau-né discrimine bien à l'aide des propriétés suprasegmentales de la parole, et donc probablement du rythme. Cependant, le filtrage de parole n'est pas idéal pour étudier le rythme dans la mesure où il préserve également certaines informations phonétiques, de même que la fréquence fondamentale, donc l'essentiel de l'intonation.

### 3.1 Stimuli

La resynthèse est une technique plus souple pour délexicaliser la parole, car elle permet d'en manipuler les propriétés phonétiques, phonotactiques ou prosodiques [voir Ramus 1999 pour de plus amples détails]. Ici les stimuli ont été resynthétisés de manière à n'en préserver que le rythme. Les phrases d'origine sont celles décrites au 2.2. Leur segmentation en consonne/voyelle a été utilisée pour les resynthétiser en remplaçant toutes les consonnes par /s/ et toutes les voyelles par /a/ [cf. Dutoit 1996 pour des informations sur le logiciel de synthèse MBROLA]. Les stimuli étaient donc des phrases de type /sasasasas.../ ayant précisément les mêmes intervalles vocaliques et consonantiques que les phrases d'origine, et une fréquence fondamentale constante à 230 Hz. Ainsi, seul le rythme des phrases d'origine était préservé<sup>5</sup>.

### 3.2 Procédure

La discrimination des paires de langues suivantes a été testée : Anglais/Japonais, Anglais/Espagnol, Polonais/Espagnol, Anglais/Polonais. Chaque paire de langue a fait l'objet d'une expérience sur 16 sujets pour la plupart de langue française. Pour la paire Anglais/Japonais, il s'agissait d'une tâche de catégorisation de langue avec entraînement et feedback [voir Ramus 1999]. Pour les autres paires, il s'agissait d'une tâche de détection d'intrus AAX: un essai consistait en 2 phrases de familiarisation dans la même langue, puis une phrase dans une langue indéterminée, dont le sujet devait dire s'il elle était ou non de la même langue que les 2 premières. Du feedback était fourni après chaque réponse. Les essais étaient regroupés en 2 blocs pendant lesquels la langue de familiarisation était maintenue constante. Les expériences ont été programmées à l'aide du langage EXPE [Pallier 1997].

L'analyse du paragraphe 2.3 prédit que les trois premières paires peuvent être distinguées sur la base du rythme. Le résultat de la quatrième dépend des variables que l'on incorpore dans le modèle : si l'on suppose que seules %V et  $\Delta C$  modélisent la perception du rythme, alors l'Anglais et le Polonais devraient être confondus ; si l'on suppose que  $\Delta V$  joue également un rôle, alors les 2 langues devraient être discriminées.

### 3.3 Résultats

Les indices de discrimination A' [Snodgrass 1988] pour les phases de test ainsi que la significativité des comparaisons (test T) par rapport au niveau de chance (50%) sont présentés dans le tableau 3.

---

<sup>5</sup> Des échantillons des types de stimuli utilisés peuvent être écoutés à l'adresse Internet suivante : <http://www.ehess.fr/centres/lscp/persons/ramus/resynth/ecoute.htm>.

**Tableau 3**

Paire de langues	A'	Significativité
Anglais/Japonais	0.72	p<0.001
Anglais/Espagnol	0.65	p<0.001
Polonais/Espagnol	0.74	p<0.001
Anglais/Polonais	0.61	p=0.006

Conformément aux prédictions, le rythme de l'Anglais est discriminé de celui de l'Espagnol et du Japonais, de même que pour l'Espagnol et le Polonais. De plus l'Anglais et le Polonais sont également discriminés, ce qui suggère que  $\Delta V$  aussi reflète des propriétés pertinentes pour les sujets.

#### 4 Discussion et application à l'identification automatique

Partant de l'hypothèse que le rythme de parole a pour source des propriétés phonologiques, nous avons montré que ces propriétés ont des corrélats phonétiques qui peuvent être mesurés directement sur le signal de parole. Ainsi, partant du signal, il est possible de retrouver la typologie rythmique des langues et de prédire quelles paires peuvent être distinguées sur la base du rythme.

On peut maintenant envisager un certain nombre d'améliorations et d'extensions au modèle actuel :

- Augmenter la taille du corpus de base, et ce dans plusieurs directions :
  - ⇒ Augmenter la taille du corpus pour chaque langue ;
  - ⇒ Augmenter le nombre de langues ;
  - ⇒ Etendre le corpus à de la parole spontanée ;Ces trois extensions, qui permettraient d'asseoir la validité empirique et l'intérêt théorique du modèle, ne peuvent s'envisager raisonnablement dans le cadre d'une segmentation manuelle du signal, ce qui nous amène au point suivant.
- Etendre le corpus à de la parole segmentée automatiquement. En principe il devrait s'agir d'une segmentation indépendante de la langue et donnant des frontières de phonèmes robustes, ce qui peut être problématique pour une segmentation phonétique détaillée. Néanmoins le présent modèle ne requiert qu'une segmentation en consonnes/voyelles, ce qui rend le problème sensiblement plus simple [voir Corredor-Ardoy 1997 et Adda-Decker dans ce volume pour une approche de la segmentation automatique multilingue].

Ainsi, l'extension du corpus par segmentation automatique permettrait 1) de tester à plus grande échelle la validité empirique du présent modèle, 2) d'augmenter sa plausibilité psychologique, étant entendu qu'il ne peut contribuer à modéliser la perception du rythme par l'humain que si ce dernier peut disposer effectivement d'un algorithme effectuant la segmentation en consonnes/voyelles, 3) d'améliorer les modèles de langues déterminés ici par analyse discriminante, 4) d'explorer les propriétés rythmiques de langues jusqu'à présent peu étudiées.

On peut à partir de là envisager l'élaboration d'un système automatique pour reconnaître le rythme des langues. Un tel système devrait comporter plusieurs niveaux :

- Extraction des silences de la parole spontanée et normalisation pour le débit du locuteur (important pour  $\Delta V$  et  $\Delta C$ ) ;
- Segmentation automatique de la parole en consonnes/voyelles.
- Exploitation des durées des intervalles vocaliques et consonantiques mesurés pour déterminer la classe rythmique de l'échantillon par rapport à des modèles de classes de

langues prédéterminés [voir Dominey 1998 et Dominey dans ce volume pour une implémentation de ce niveau par des réseaux de neurones].

Il va sans dire que cette reconnaissance du rythme ne permettra pas d'identifier des langues au sein des classes de rythme, mais pourrait néanmoins réduire l'espace de recherche pour des systèmes d'identification des langues se basant sur d'autres approches (acoustiques, phonétiques ou phonotactiques).

Il reste à déterminer si la caractérisation rythmique d'un échantillon de parole peut réellement apporter des informations qui ne sont pas déjà obtenues par les autres approches. Par exemple, on peut penser que les informations de type phonotactique, notamment les probabilités de transition entre consonnes et voyelles, incluent de fait les informations rythmiques liées à la complexité des syllabes et aux groupes consonantiques. Néanmoins, le rythme de parole n'est peut-être pas réductible aux probabilités de transition entre consonnes et voyelles, notamment parce que les informations de durée ne sont pas prises en compte actuellement dans les modèles phonotactiques. Les données ci-dessus en donnent un exemple: l'Anglais et le Polonais ont la même complexité syllabique, mais diffèrent par la présence (en anglais) ou non (en polonais) de réduction vocalique, ce qui se traduit dans nos données par un  $\Delta V$  plus élevé pour l'Anglais. Il s'agit donc là d'une différence rythmique reposant strictement sur la durée des voyelles et donc non accessible aux modèles phonotactiques.

Dans le cas particulier de l'Anglais et du Polonais, on pourrait bien sûr argumenter que la discrimination est néanmoins possible sur la base de différences purement phonétiques, mais cela n'est pas nécessairement vrai de toutes les paires comportant des différences rythmiques. De plus, les modèles phonétiques peuvent manquer de robustesse sur de courts échantillons de parole, car un échantillon court ne comporte qu'un ensemble réduit de phonèmes et donc sous-détermine grandement le répertoire phonétique de la langue. La robustesse des caractéristiques rythmiques que nous avons montrée sur de courts échantillons nous laisse donc penser qu'elles peuvent apporter de l'information significative au moins dans ce cas précis.

## 5 Bibliographie

- [Abercrombie 1967] D. Abercrombie, *Elements of general phonetics*. Chicago: Aldine, (1967).
- [Bertinetto 1981] P. M. Bertinetto, *Strutture prosodiche dell' italiano. Accento, quantità, sillaba, giuntura, fondamenti metrici*. Firenze: Accademia della Crusca, (1981).
- [Bolinger 1965] D. Bolinger, "Pitch accent and sentence rhythm," in *Forms of English: Accent, morpheme, order*. Cambridge, MA: Harvard University Press, (1965).
- [Corredor-Ardoy 1997] C. Corredor-Ardoy, J. L. Gauvain, M. Adda-Decker, & L. Lamel, "Language identification with language-independent acoustic models," presented at Eurospeech, Rhodes, (1997).
- [Dasher 1982] R. Dasher & D. Bolinger, "On pre-accentual lengthening," *Journal of the International Phonetic Association*, vol. 12, pp. 58-69, (1982).
- [Dauer 1983] R. M. Dauer, "Stress-timing and syllable-timing reanalyzed," *Journal of Phonetics*, vol. 11, pp. 51-62, (1983).
- [Dominey 1998] P. Dominey & F. Ramus, "Neural network processing of natural language: I. Sensitivity to prosodic structure in the new-born," (soumis).
- [Dutoit 1996] T. Dutoit, V. Pagel, N. Pierret, F. Bataille, & O. van der Vrecken, "The MBROLA Project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes," presented at ICSLP'96, Philadelphia, (1996).
- [Ladefoged 1975] P. Ladefoged, *A course in phonetics*. New York: Harcourt Brace Jovanovich, (1975).

- [Mehler 1996] J. Mehler, E. Dupoux, T. Nazzi, & G. Dehaene-Lambertz, "Coping with linguistic diversity: The infant's viewpoint," in *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, J. L. Morgan & K. Demuth, Eds. Mahwah, NJ: Lawrence Erlbaum Associates, (1996), pp. 101-116.
- [Mehler 1988] J. Mehler, P. Jusczyk, G. Lambertz, N. Halsted, J. Bertoncini, & C. Amiel-Tison, "A precursor of language acquisition in young infants," *Cognition*, vol. 29, pp. 143-178, (1988).
- [Moon 1993] C. Moon, R. P. Cooper, & W. P. Fifer, "Two-day-olds prefer their native language," *Infant Behavior and Development*, vol. 16, pp. 495-500, (1993).
- [Nazzi 1997] T. Nazzi, "Du rythme dans l'acquisition et le traitement de la parole.," Paris: Ecole des Hautes Etudes en Sciences Sociales, (1997).
- [Nazzi 1998] T. Nazzi, J. Bertoncini, & J. Mehler, "Language discrimination by newborns: towards an understanding of the role of rhythm," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 24, pp. 756-766, (1998).
- [Nespor 1990] M. Nespor, "On the rhythm parameter in phonology," in *Logical issues in language acquisition*, I. M. Roca, Ed. Dordrecht: Foris, (1990), pp. 157-175.
- [Pallier 1997] C. Pallier, E. Dupoux, & X. Jeannin, "EXPE: An expandable programming language for on-line psychological experiments," *Behavior Research Methods, Instruments, & Computers*, vol. 29, pp. 322-327, (1997).
- [Pike 1945] K. L. Pike, *The intonation of American English*. Ann Arbor, Michigan: University of Michigan Press, (1945).
- [Ramus 1999] F. Ramus & J. Mehler, "Language identification with suprasegmental cues: A study based on speech resynthesis," *Journal of the Acoustical Society of America*, vol. 105, pp. 512-521, (1999).
- [Roach 1982] P. Roach, "On the distinction between "stress-timed" and "syllable-timed" languages," in *Linguistic controversies*, D. Crystal, Ed. London: Edward Arnold, (1982).
- [Rubach 1985] J. Rubach & G. E. Booij, "A grid theory of stress in Polish," *Lingua*, vol. 66, pp. 281-319, (1985).
- [Snodgrass 1988] J. G. Snodgrass & J. Corwin, "Pragmatics of measuring recognition memory: Applications to dementia and amnesia," *Journal of Experimental Psychology: General*, vol. 117, pp. 34-50, (1988).