

1048

C.T. Best

---

Perception de Parole de Différentes Langues par le Bébé

**DÉVELOPPEMENT DE LA  
PERCEPTION DES SONS  
DE PAROLE DE DIFFÉRENTES  
LANGUES AU STADE PRÉVERBAL**

**Catherine T. BEST**  
Wesleyan University  
et  
Haskins Laboratories  
U.S.A.

CATHERINE T. BEST, Ph.D.  
DEPARTMENT OF PSYCHOLOGY  
WESLEYAN UNIVERSITY  
MIDDLETOWN, CT 06457

*On s'attend d'une très grande punition.*

*Gal*

# LE DÉVELOPPEMENT LANGAGIER: UNE PRÉDICTION PRÉCOCE EST-ELLE POSSIBLE ?

ACTES DU COLLOQUE  
« Le premier langage peut-il être prédictif  
du développement langagier ultérieur ? »

organisé par le **GRI DESYCOLE** \*  
Université de Franche-Comté  
Besançon 6 juin 1994

**G. KONOPCZYSKI**  
**S. VINTER**

*publischer : l'Ortho-Editions ; Besançon, France*

\*GRI DESYCOLE = Groupe de Recherche Interdisciplinaire sur le  
Développement des Systèmes de Communication chez l'Enfant.

Pour répondre à ces questions, j'ai proposé un modèle de perception inter-langues de la parole, le « modèle d'assimilation perceptive » (ou PAM pour Perceptual Assimilation Model) (BEST, 1993, 1994, sous presse). Il suppose que le type d'information que les enfants et les adultes détectent dans la parole est une information sur les gestes distals du conduit vocal qui produit le signal de parole c'est-à-dire, sur des événements articulatoires. Ce point de vue est fondé sur la théorie écologique de la perception de James GIBSON (1966, 1979), c'est à dire sur la théorie de la perception directe du monde réel. Enfants et adultes diffèrent dans leur capacité à reconnaître des coordinations entre gestes articulatoires spécifiques à une langue, qui relèvent d'un niveau d'abstraction supérieur. Par exemple, la coordination temporelle entre le geste laryngal dévoisant une consonne et le relâchement de la fermeture des lèvres ; ou bien la séquence des gestes dans les patterns syllabiques qui sont phonotactiquement admis dans une langue donnée. Le système phonologique comporte des patterns complexes de ce genre, que les adultes ont appris à détecter dans leur langue maternelle. Les jeunes enfants, en revanche, ne détectent que des patterns simples de gestes articulatoires et des différences qui sont potentiellement utilisées dans n'importe quelle langue, avant de pouvoir commencer à reconnaître des patterns de coordination de gestes spécifiques à une langue. L'information sur les gestes simples que les jeunes enfants peuvent appréhender est représentée comme un espace phonétique universel des gestes articulatoires possibles. Pour les auditeurs adultes, la perception est contrainte par une focalisation perceptive sur les catégories divisant l'espace phonologique natif. Ces catégories sont spécifiques à la langue native, et relèvent d'un niveau de complexité supérieur.

D'autre part, le modèle PAM suppose que les auditeurs adultes perçoivent non seulement leur langue selon leurs catégories phonologiques natives, mais tendent aussi fortement à percevoir des productions de parole dans une autre langue selon les mêmes catégories natives. En d'autres termes, ils assimilent perceptivement les catégories non natives aux catégories phonologiques natives les plus proches, selon les similarités et dissimilarités perçues entre les catégories natives et non natives. Un phonème non natif peut être assimilé à une catégorie phonologique native, soit comme un bon exemplaire de la catégorie native, soit comme un exemplaire déviant. Il est aussi possible qu'un tel phonème ne puisse être assimilé de façon univoque à aucune catégorie native, et qu'il tombe donc dans un

« espace phonétique inexploité » entre catégories, ou encore, qu'il soit perçu comme un son de parole inclassable. Enfin, il peut tomber à l'extérieur de l'espace phonologique natif et ne pas être perçu comme un son de parole.

Les implications de ces possibilités d'assimilation pour les divers types de contrastes phonologiques non natifs, et les prédictions associées concernant les performances de discrimination, sont les suivantes : 1) *TC* (**T**wo-**C**ategory, ou Deux Catégories), où la discrimination est excellente. 2) *CG* (**C**ategory **G**oodness difference, ou différence de qualité d'une Catégorie), où la discrimination est bonne à très bonne. 3) *SC* (**S**ingle **C**ategory, ou Catégorie unique), où la discrimination est mauvaise et 4) *NA* (**N**on-**A**ssimilable), où la discrimination est bonne à très bonne, comme CG.

Les résultats de mes études perceptives inter-langues avec des adultes supportent fortement les prédictions du modèle PAM. La première étude, avec McROBERTS et SITTHOLE (BEST, McROBERTS & SITTHOLE 1988), était centrée sur le type *NA* (non assimilable) de contrastes non natifs pour des sujets anglophones américains ; les 18 contrastes utilisés portaient sur le voisement et la place d'articulation de clicks consonantiques non nasalisés du Zulu. Les sujets dans leur grande majorité ont décrit ces clicks comme n'étant pas des sons de parole. Autrement dit, les clicks étaient *non assimilables* à des catégories de l'espace phonologique de l'anglais américain. De plus, les sujets ont discriminé tous les contrastes de clicks beaucoup mieux que ne l'auraient laissé prévoir des études inter-langues plus classiques (cf. aussi BEST, McROBERTS, LAFLEUR & SILVER-ISENSTADT sous presse). Nous avons ensuite examiné trois autres contrastes du Zulu qui devaient correspondre aux trois autres cas de figure de notre modèle : *TC* (un contraste de fricatives latérales) ; *CG* (un contraste d'occlusives vélaires éjective *versus* non voisée) ; et *SC* (un contraste de « b » occlusif *versus* implosif). Comme prédit par le modèle, les adultes américains ont assimilé les fricatives latérales à deux catégories de l'anglais, bien que nombre de sujets aient décrit ces sons comme des combinaisons phonotactiquement illégales telles que « shl : zhl » ou « thl : thl » (/θ/~~/ð/). La plupart des sujets ont décrit les contrastes de vélaires comme relevant d'une différence de bon exemplaire (donc *CG*) entre un « k » anglais et un « k » déviant. Ils décrivaient généralement les occlusives bilabiales « b » comme quasiment identiques au « b » anglais. Les performances de discrimination confirment les prédictions du modèle : discrimination à

un niveau proche du plafond pour les fricatives latérales ; très bonne discrimination pour les occlusives vélares, mais significativement moins bonne que pour les fricatives latérales ; enfin, discrimination beaucoup plus pauvre pour les occlusives bilabiales. Nous avons ensuite confirmé le pattern de performances du cas de *TC* pour un second contraste non natif : [p'e] : [t'e] (consonnes initiales éjectives) en Tigrinya (Éthiopie).

Les premières analyses d'une étude récente sur des sujets adultes français (avec A. LEVITT, P. HALLÉ, et B. DE BOYSSON-BARDIES) sont aussi en accord avec le modèle PAM, montrant en particulier les différences inter-langues entre groupes de sujets prédites par le modèle : la plupart des sujets français assimilent le contraste des occlusives bilabiales du Zulu — qui était difficile à percevoir par les américains — à une différence entre les consonnes *b* et *p* du français, donc le perçoivent comme un contraste *TC*, comme on pouvait s'y attendre étant donné les caractéristiques phonétiques de ces consonnes [du français et celles du contraste Zulu].

Plus récemment, les prédictions du modèle PAM ont été étendues à un ensemble de contrastes vocaliques non natifs (avec A. LEVITT et A. FAHNER) : contrastes /ə/ : /œ/, /œ/ : /y/, et /o/ : /ò nasalisé/ du français ; /i/ : /y/ et /i/ : /ʉ/ du norvégien ; et /ə/ : /u/ du Thai. Les patterns d'assimilation des sujets américains différaient selon ces contrastes. Le contraste norvégien /i/ : /y/ était la plupart du temps assimilé à une seule catégorie /i/ (*SC*), tandis que le contraste norvégien /i/ : /ʉ/ était régulièrement assimilé à un contraste natif /i/ : /U/. Comme prédit, la discrimination était la plus médiocre pour le contraste norvégien /i/ : /y/, mais était excellente pour /i/ : /ʉ/. A nouveau, les premiers résultats obtenus avec des sujets français montrent des différences de perception que prédisent bien les différences entre les répertoires de voyelles français et anglais. En particulier, la plupart des français assimilaient le contraste norvégien /i/ : /y/ (très déroutant pour les américains) au contraste français correspondant /i/ : /y/ ; ils discriminaient assez bien les voyelles de ce contraste. Il faut rappeler que le /y/ norvégien diffère phonétiquement du /y/ français à la fois pour l'arrondissement des lèvres et pour la hauteur de la langue. Quelques sujets français furent déroutés par cette différence phonétique et ont assimilé les deux voyelles du contraste norvégien au /i/ français ; ces sujets étaient aussi ceux qui discriminaient le moins bien les deux voyelles norvégiennes.

Bien entendu, nous avons aussi étudié les performances de dis-

crimination d'enfants apprenant l'anglais sur les contrastes consonantiques décrits précédemment. Les résultats montrent, au cours de la première année, des changements spécifiques à la langue apprise en partie compatibles avec les performances des adultes américains, et en partie différents pour certains aspects. Dans une première étude, nous avons trouvé que des enfants apprenant l'anglais continuent à bien discriminer différents clics Zulu jusqu'à l'âge d'au moins 14 mois, ce qui est compatible avec la bonne perception par les adultes des contrastes *non assimilables* correspondants. Il faut souligner que ce résultat est *incompatible* avec le déclin de la perception des contrastes non natifs trouvé par WERKER pour d'autres contrastes consonantiques vers 10-12 mois. Nous avons fait l'hypothèse que les contrastes étudiés par WERKER correspondaient au cas de figure *SC*. Effectivement, dans une seconde étude avec de nouveaux groupes d'enfants de 6-8 mois et 10-12 mois, nous avons pu répliquer d'une part le déclin en discrimination pour l'un des contrastes étudiés par WERKER (un contraste Nthlakampx) et d'autre part l'*absence* de déclin que nous avons trouvé pour les clics.

Dans l'étude suivante, nous avons trouvé non seulement un changement développemental entre 6-8 mois et 10-12 mois pour la perception des trois contrastes Zulu portant sur d'autres consonnes que des clics, mais aussi certaines différences entre les performances des adultes anglophones et des enfants de 10-12 mois apprenant l'anglais. Les enfants de 6-8 mois percevaient comme prévu les trois contrastes. Mais, alors que les adultes discriminaient les fricatives latérales (contraste *TC*) et les occlusives vélares (contraste *CC*), les 10-12 mois en étaient incapables.

Pourquoi cette difficulté des enfants les plus âgés à percevoir des contrastes relativement faciles pour les adultes ? Des résultats récents obtenus par JUSZYK suggèrent une réponse possible. On se rappelle que la plupart des sujets adultes assimilaient les fricatives latérales Zulu à des combinaisons de consonnes phonotactiquement illégales en anglais. Le changement développemental trouvé par JUSZYK dans la perception des patterns phonotactiques natifs *versus* non natifs, vers l'âge de 9 mois, suggère que nos sujets de 10-12 mois ne parvenaient pas à discriminer les fricatives latérales parce qu'ils ne pouvaient détecter aucun pattern phonotactique familier ou natif dans ces sons. Reste à vérifier directement cette interprétation. En tout cas, elle est en accord avec les résultats d'une étude complémentaire avec les éjectives de l'éthiopien. Les adultes percevaient

ces consonnes comme phonotactiquement admises en anglais, assimilant le contraste à *pay* : *tay* (contraste *TC*). Un nouveau groupe d'enfants de 10-12 mois discriminaient bien ce contraste, bien que ne parvenant pas, eux non plus, à discriminer les fricatives latérales Zulu.

Je propose l'hypothèse de travail suivante, fondée sur les patterns développementaux qui viennent d'être décrits : les très jeunes enfants discriminent les contrastes de sons de parole qu'ils soient natifs ou non, parce qu'ils détectent des différences entre gestes simples qui ne reflètent pas encore la connaissance de patterns de coordination gestuelle plus abstraits, spécifiques à une langue. Vers la seconde moitié de la première année, les enfants commencent à détecter ces patterns abstraits, peut-être plus tôt pour les contrastes vocaux que pour les contrastes consonantiques, comme le suggèrent les récents travaux de Kuhl (GRIESER & Kuhl 1989, Kuhl et al 1992) et de Polka & Werker (1994). Mais cette étape précoce ne reflète pas encore la connaissance de la valeur fonctionnelle de ces contrastes dans le système phonologique adulte, ni non plus le niveau de détail phonétique que les adultes reconnaissent comme caractérisant leurs catégories phonologiques natives. Le développement d'une connaissance plus complète du système phonologique natif requiert probablement plusieurs années entre la fin de la première enfance et le stade préscolaire. Parmi les « étapes » de ce développement, l'émergence d'une conscience phonologique vers l'âge de 5 ans apparaît comme un indice fiable de l'éveil d'une bonne aptitude pour la lecture.

Une philosophie de la perception directe du monde réel conduit naturellement à s'intéresser de plus près aux relations entre la phonétique (les propriétés physiques « de surface » de la parole) et la phonologie (les catégories abstraites, linguistiques, qui sous-tendent les détails phonétiques et sur lesquelles reposent les oppositions de sens dans chaque langue particulière). La phonologie d'une langue est considérée comme faisant partie de droit de la grammaire qui lui est spécifique, et donc comme étant linguistique par nature. Par opposition, la phonétique a traditionnellement été cantonnée à la mise en correspondance automatique du code phonologique abstrait avec les mouvements de l'appareil vocal, comme le proposent Chomsky et Halle dans « *Sound Pattern of English* ». Les faits phonétiques ont ainsi été considérés comme un aspect superficiel et **non linguis-**

**tique** du langage. Cependant, l'intérêt théorique d'une « interface » phonétique-phonologie a récemment resurgi avec l'observation que les langues, voire les dialectes, diffèrent dans les spécifications phonétiques de catégories phonologiques identiques (définies par les mêmes traits). Par exemple, Keating (1990) a montré que les diverses langues n'utilisent pas les mêmes indices phonétiques pour réaliser les distinctions de voisement des consonnes post vocaliques, comme dans *bat* : *bad* ; Lindau (1982) rapporte que les occlusives éjectives diffèrent par l'intensité et le délai du relâchement glottal entre le [t'a] Navajo et le [t'a] Quechua. Qui plus est, les réalisations phonétiques du même phonème peuvent varier selon les variétés dialectales d'une même langue, comme l'a montré van Riemsdijk (1982) pour le timing de l'abaissement du vélum dans la production des voyelles nasales en français parisien par opposition au français du Québec. Ces observations remettent en question l'idée selon laquelle la conversion du code phonologique en ses détails phonétiques est « automatique » et universelle. Par conséquent, la structuration phonétique d'une langue, tout comme son système phonologique, doit être spécifique et, en toute logique, devrait être considérée comme une composante véritablement linguistique de la grammaire de cette langue.

Étant donné que l'enfant ne parvient à posséder le système phonologique de sa langue que par le biais de l'input phonétique qui est spécifique de sa langue et même de son dialecte, il est inévitable que les deux domaines — le physique et l'abstrait — soient inextricablement liés, du moins dans les étapes précoces de développement du langage. Autrement dit, les caractéristiques phonétiques physiques doivent servir de base au développement du système phonologique spécifique à une langue. Nos hypothèses sur la perception des informations gestuelles sont compatibles avec ces notions d'interface phonétique-phonologie et de développement de la compétence phonologique.<sup>1</sup>

\* Nos remerciements à Pierre Hallé pour la traduction de cet article en français.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- BEST C.T. (1993) : Emergence of language-specific constraints in perception of non-native speech : A window on early phonological development. In B. DE BOYSSON-BARDIES, S. DE SCHONEN, P. JUSCZYK, P. MACNEILAGE, AND J. MORTON (eds.) *Developmental Neurocognition : Speech and Face Processing in the First Year of Life* (289-304). Dordrecht, the Netherlands : Kluwer Academic Publishers.
- BEST C.T. (1994) : The emergence of native-language phonological influences in infants : A perceptual assimilation model. In J. GOODMAN & H. C. NUSBAUM (eds.) *The Development of Speech Perception : The Transition from Speech Sounds to Spoken Words*. (167-224). Cambridge MA : MIT Press.
- BEST C.T. (in press) : Learning to perceive the Sound Pattern of English. To appear in C. ROVEE-COLLIER & L. LIPSITT (eds.) *Advances in Infancy Research*. Vol. 8 (1994). Norwood, NJ : Ablex Publishing Corporation.
- BEST C.T., McROBERTS, G.W., LAFLEUR R., AND SILVER-ISENSTADT J. (in press) : Divergent developmental patterns for infants' perception of two non-native consonant contrasts. *Infant Behavior and Development*.
- BEST C.T., McROBERTS G.W., & SITHOLE N.N. (1988) : The phonological basis of perceptual loss for non-native contrasts : Maintenance of discrimination among Zulu clicks by English-speaking adults and infants. *Journal of Experimental Psychology : Human Perception and Performance*, 14, 345-360.
- GIbson J.J. (1966) : *The Senses Considered as Perceptual Systems*. Boston, MA : Houghton-Mifflin.
- GIbson J.J. (1979) : *The Ecological Approach to Visual Perception*. Boston, MA : Houghton-Mifflin.
- GRIEser D.L., AND Kuhl P.K. (1989) : Categorization of speech by infants : Support for speech-sound prototypes. *Developmental Psychology*, 25, 577-588.
- JUSCZYK P.W., KEMLER-NELSON D.G., HIRSH-PASEK K., KENNEDY L., WOOWWARD A., PIWOZ J. (1992) : Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, 24, 252-293.
- KEATING P.A. (1990) : Phonetic representations in a generative grammar. *Journal of Phonetics*, 18, 321-334.
- Kuhl P.K., WILLIAMS K.A., LACERDA F., STEVENS K.N., & LINDHOLM B. (1992) : Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255, 606-608.
- LINDAUM. (1982) : Phonetic differences in glottalic consonants. *UCLA Working Papers in Phonetics*, 54, 66-77.
- LISKER L., & ABRAMSON A.S. (1970) : The voicing dimension : Some experiments on comparative phonetics. *Proceedings of the 6th International Congress of Phonetic Sciences*. Prague : Academia.
- MACKAIN K.S., BEST C.T., & STRANGE W. (1981) : Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, 2, 369-390.
- MEHLER J., JUSCZYK P., LAMBERTZ G., HALSTED N. BERTONCINI J., & AMIEL-TISON C.A. (1988) : A precursor of language acquisition in young infants. *Cognition*, 29, 143-178.
- MIYAWAKI K., STRANGE W., VERBRUGGE R., LIBERMAN A.M., JENKINS J.J., & FUJIMURA O. (1975) : An effect of linguistic experience : The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics*, 18, 331-340.
- POLKA L. (1991) : Cross-language speech perception in adults : Phonemic, phonetic, and acoustic contributions. *Journal of the Acoustical Society of America*, 89, 2961-2977.
- POLKA L. (1992) : Characterizing the influence of native experience on adult speech perception. *Perception & Psychophysics*, 52, 37-52.
- POLKA L. and Werker, J. F. (1994) : Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology : Human Perception & Performance*, 20, 421-435.
- TEES R.C., & WERKER J.F. (1984) : Perceptual flexibility : Maintenance or recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology*, 38, 579-590.
- TREHUB S.E. (1976) : The discrimination of foreign speech contrasts by adults and infants. *Child Development*, 47, 466-472.
- VAN REENEN P. (1982) : *Phonetic Feature Definitions : Their Integration into Phonology and their Relation to Speech, a Case Study of the Feature Nasal*. Dordrecht : Foris Publications.
- WERKER J.F., GILBERT J.H.V., HUMPHREY K., & TEES R.C. (1981) : Developmental aspects of cross-language speech perception. *Child Development*, 52, 349-355.
- WERKER J.F., & LALONDE C. E. (1988) : Cross-language speech perception : Initial capabilities and developmental change. *Developmental Psychology*, 24, 672-683.
- Werker J., & Logan J. (1985) : Cross-language evidence for three factors in speech perception. *Perception and Psychophysics* 37, 35-44.
- Werker J.F., & Tees R.C. (1984a) : Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75, 1866-1878.
- Werker J.F., & Tees R.C. (1984b) : Cross-language speech perception : Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7, 49-63.