



HAL
open science

ABISPA: Apprentissage Bayésien Intersensoriel de Structures Phonologiques par un Androïde bébé

Louis-Jean Boë, Jean-Luc Schwartz

► **To cite this version:**

Louis-Jean Boë, Jean-Luc Schwartz. ABISPA: Apprentissage Bayésien Intersensoriel de Structures Phonologiques par un Androïde bébé. 2005. hal-00003661

HAL Id: hal-00003661

<https://hal.science/hal-00003661>

Preprint submitted on 20 Jan 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ABISPA : Apprentissage Bayésien Intersensoriel de Structures Phonologiques par un Androïde bébé *

Responsable scientifique : Louis-Jean BOË

Louis-Jean BOË,
Université Stendhal (BP 25 38040 Grenoble Cedex)
Tel: 0476824338

Sous-thèmes dont relève ce projet :
Acquisition du langage
Phonétique, phonologie, parole, oral
Lexique

et Jean-Luc SCHWARTZ
ICP, INPG (46, Avenue Félix Viallet – 38031 Grenoble)
Tel: 0476574712

Équipes partenaires :

- Laboratoire Modyco - FRE 2329 - Université de Paris X- Nanterre
- Laboratoire TIMC - UMR 5525 - Université Joseph Fourier, Grenoble 1
- Laboratoire de Psychologie Expérimentale - Université Pierre Mendès France, Grenoble 2
- Équipe Laplace – INRIA, Grenoble
- Department of Psychology, Austin, Texas
- Speech and Hearing Center, Austin, Texas

Résumé signalétique

Le projet ABISPA a tenté d'apporter des éléments de réponse à deux questions imbriquées :

1. Comment un « androïde bébé » peut-il acquérir le contrôle des gestes pour produire des signaux associés à des représentations phonologiques ?
2. Comment cet « androïde » peut-il apprendre des cibles, des invariances et les représentations phonologiques associées à partir d'informations auditives, visuelles et orosensorielles ?

ABISPA a associé – peut-être pour la première fois en France – des chercheurs en :

- phonologie, phonétique,
- psychologie expérimentale,
- robotique,
- modélisation du vivant,

dans une problématique typiquement « sciences cognitives » : l'apprentissage auditif, visuel et orosensoriel de structures phonologiques par un androïde bébé.

La voie qui a été explorée a consisté à (re) mettre en jeu les représentations perceptives dans le cadre de la Théorie de la Perception pour le Contrôle de l'Action. Cette approche passe par la définition d'un cadre de modélisation quantitative des mécanismes de mise en relation entre perception et action. Or ce cadre existe et il a fait l'objet de nombreux développements depuis 10 ans : c'est la robotique de la parole. C'est, en effet, le lieu naturel de modélisation et d'étude des relations entre un système sensoriel et un système moteur ; le noyau central en est la mise en œuvre de processus de passage d'espaces (moteurs) à espaces (sensoriels) permettant de résoudre des problèmes complexes d'optimisation ou de satisfaction de contraintes multi-espace (cf. par exemple le projet Cog, du MIT). En postulant que la parole est construite et régulée par les interactions sensori-motrices, nous avons cherché à élaborer un « androïde bébé », agent sensori-moteur doté d'actionneurs de la parole (un conduit vocal), de capteurs auditifs, visuels et orosensoriels et de mécanismes de contrôle et d'apprentissage. Nous avons mis cet androïde en situation d'apprendre à parler, c'est-à-dire d'acquérir contrôles et représentations phonologiques, à partir de signaux externes.

Mots-clés : Acquisition de la parole • Phonétique et Phonologie des voyelles • Modélisation du conduit vocal .
Nouveau-né • Apprentissage Bayésien • Robotique • Perception des traits

Nombre de participants : enseignants-chercheurs : 9 ; chercheurs : 9 ; thésards : 5 ; stagiaires : 9.

Nombre total homme-mois : 184

* Ce projet a bénéficié d'un complément de soutien financier, de la part de Cognitique, au titre de l'aide à la valorisation.

Rappel des enjeux et objectifs fixés à l'origine

Compétence humaine majeure, la parole est un objet de recherche interdisciplinaire, au carrefour de l'étude de systèmes sensori-moteurs de production et de perception, de signaux biologiques de communication et, bien sûr, du langage. Dans le monde, se sont créés des regroupements de disciplines multiples, tant fondamentales que technologiques, qui étudient la parole avec différents angles d'attaque allant de l'acoustique, au traitement du signal, à la robotique, et à la phonétique et la phonologie, en passant par la biologie, la psychologie, mettant la parole au centre de recherches tant fondamentales qu'appliquées. Des ré-articulations se sont déjà produites et continuent à se produire entre ces domaines de recherches intéressés par l'objet parole. Se dessinent ainsi nettement de nouvelles négociations entre les sciences du langage, les sciences de la cognition et certains secteurs des

sciences physiques pour l'ingénieur pour les développements de modèles de production, de perception et d'apprentissage. On peut ainsi noter certains courants de la Phonologie qui se conjuguent avec les Sciences de parole et les Sciences cognitives. Leur approche vise à soumettre les hypothèses de l'organisation phonologique à une validation de type sciences expérimentales, dont la phonologie générative s'est affranchie par principe : « Ancrer la réalité des systèmes phonologiques dans la substance ». Il s'agit en l'occurrence de répondre aux deux questions : (i) Comment un androïde bébé peut-il acquérir le contrôle des gestes pour produire des signaux associés à des représentations phonologiques ? (ii) Comment cet androïde peut-il apprendre des cibles, des invariances et les représentations phonologiques associées à partir d'informations auditives, visuelles et orosensorielles.

Résumé des résultats effectivement atteints

- Mise en place de « l'Androïde bébé » et de ses capteurs sensoriels
Sciences de la Parole, Sciences Cognitives
- Évaluation de l'androïde
Sciences de la Parole, Phonétique
- Élaboration de prototypes vocaliques
Sciences du Langage, Phonétique
- Étude des stratégies de compensation
Sciences du Langage, Phonétique
- Implantation du paradigme Bayésien
Robotique, Sciences Cognitives
- Le parcours développemental
Psychologie, Sciences du langage
- L' limitation de cibles vocaliques isolées
Robotique, Sciences de la parole, Sciences Cognitives
- Modélisation des tissus mus et de la face
Informatique médicale
- Modélisation biomécanique de la langue
Informatique médicale
- Normalisation auditive des voyelles
Sciences Cognitives, Psychologie, Sciences de la Parole
- Des gestes de l'autre à la perception des traits
Sciences du Langage, Phonétique, Phonologie
- Apprentissage des primitives phonologiques et babillage
Sciences du Langage, Phonologie
- Implantation d'une base de données sur la croissance du conduit vocal
Informatique
- Enrichissement des données sur le babillage
Sciences du langage

Publications issues du projet

En collaboration

ICP — Laboratoire de Psychologie Expérimentale, Grenoble

Duchêne, F. (2002) *Sait-on récupérer les traits vocaliques à partir du signal de parole ?* TER de Maîtrise de Psychologie Cognitive, Université Pierre Mendès France, Grenoble.

Golanski, C., Tournafol, É. (2000) Normalisation perceptive des voyelles (de l'adulte au nouveau-né). Maîtrise MASS. UPMF, Grenoble.

Ménard, L., Schwartz, J.-L., Boë, L.J., Kandel, S., & Vallée, N. (2001) Perceptual Identification and Normalization of Synthesized French Vowels from Birth to Adulthood. *Eurospeech'2001, Aalborg Denmark*, 163-166.

Ménard, L., Schwartz, J.-L., Boë, L.J., Kandel, S., & Vallée, N. (2002) Auditory normalization of French vowels synthesized by an articulatory model simulating growth from birth to adulthood. *Journal of the Acoustical Society of America*, 111 (4), 1992-1905.

Vallée, N., Schwartz, J.-L., Uomini, T., Boë, L.J., Kandel, S. (2002) Une étude expérimentale de l'identification perceptive du trait vocalique d'aperture. *4e Journées Internationales du GDR « Phonologie » Grenoble* (résumé).

ICP — TIMC, Grenoble

Payan, Y., Chabanas, M., Pelorson X., Vilain, C., Levy, P., Luboz, V. & Perrier, P. (2002) Biomechanical models to simulate consequences of maxillofacial surgery. *C. R. Biologies*, Vol. 325, pp. 407-417.

ICP — Laplace, INRIA, Grenoble

Schwartz, J.-L., Boë, L.J., Bessière, P. (2002). Speech robotics, a computational framework for the study and modelling of speech acquisition and evolution. Submitted to *Evolution of Communication*.

ICP — Laplace, INRIA, Grenoble — Department of Communication Sciences and Disorders, Austin, Texas

Serkhane, J., Schwartz, J.-L., Boë, L.J., Davis, B., Bessière, P., Mazer, E. (2002) Étude comparative de vocalisations de bébés humains et de bébés robots. *XXIV Journées d'Étude sur la Parole, Nancy*, 149-152.

Dans chaque laboratoire

Abry C., Stefanuto M., Vilain A., Laboissière R. (2002) *What can*

- the utterance 'Tan, Tan' of Broca's patient Leborgne tell us about the hypothesis of an emergent 'babble-syllable' downloaded by SMA? In *Phonetics, Phonology and Cognition*. J. Durand and B. Laks ed., pp. 226-243. Oxford University Press, Oxford GB.
- Abry, C., Cathiard, M.-A., Vilain, A., Laboissière, R., Loevenbruck, H., Savariaux, C., Schwartz, J.-L. (2002) *Some insights in bimodal perception given for free by the natural time course of speech production*. Festschrift à Christian Benoit, eds. E. Vatikiotis-Bateson, G. Bailly, à paraître.
- Bessière, P. (2000) *Vers une théorie probabiliste des systèmes sensori-moteurs*. HDR, Université Joseph Fourier, Grenoble, France.
- Boë, L.J., Ménard, L., Maeda, S. (2000) Adaptation of control strategies during the vocal tract growth inferred from simulation studies with an articulatory model. *5th Seminar on Speech Production: Models and Data*. CREST Workshop on Models of Speech Production: Motor Planning and Articulatory Modelling. Kloster Seoon, Bavaria, 277-279.
- Lalévée, C. (2002) *Développement du contrôle des segments du corps de la parole: étude longitudinale du contrôle oral/nasal de 6 à 12 mois*. TER de Maîtrise des Sciences du Langage, Université Stendhal, Grenoble.
- Ménard, L. (2003) Acoustic variability and adaptive articulatory strategies during vocal tract growth revealed by the rounding contrast in French. *ICPhS 2003*, Barcelonne. (soumis).
- Ménard, L. (2002) *Production et perception des voyelles au cours de la croissance du conduit vocal: variabilité, invariance et normalisation*. Doctorat Sciences du Langage. Université de Grenoble III.
- Ménard, L., Boë L.J. (2000) Exploring vowel production strategies from infant to adult by means of articulatory inversion of formant data. *ICSLP*, Beijing, China, 465-468.
- Ménard, L., Boë L.J., Maeda, S. (2000) Croissance du conduit vocal et stratégies articulatoires. *XXIIIe Journées d'Étude sur la Parole*, Aussois, 421-424.
- Obolensky, D. (2001) Un modèle 3D du conduit vocal. Adaptation à la variabilité (âge, sexe, individu, espèce). Stage ENSERG, 2^e année. ICP, Grenoble.
- Rochet-Capellan, A., Boë, L.J., Schwartz, J.-L. (2002) *Le site du projet Co-Speech. Du Contrôle Orofacial des Gestes dans la communication chez les primates jusqu'à la Parole humaine*. Site <http://www.icp.inpg.fr/OHLL/>
- Schwartz, J.-L., Abry, C., Boë, L.J., Cathiard, M. (2002) Phonology in a Theory of Perception-for-Action-Control. In *Phonetics, Phonology and Cognition*. J. Durand, B. Laks ed., pp. 254-280. Oxford University Press, Oxford GB.
- Schwartz, J.-L., Boë, L.J. (2000) Predicting palatal contacts from jaw and tongue commands: A new sensory model and its potential use in speech control. *5th Seminar on Speech Production: Models and Data*. CREST Workshop on Models of Speech Production: Motor Planning and Articulatory Modelling. Kloster Seoon, Bavaria, 257-260.
- Schwartz J.-L., Boë L.J. (2001) La robotique cognitive de la parole et du langage: des structures sonores des langues du monde à l'apprentissage intersensoriel de la parole par un androïde bébé. *ASTI'2000 Premières Rencontres des Sciences et Technologies de l'Information*. Cité des Sciences, Paris. p. 60 et CR-ROM.
- Schwartz, J.-L., Boë, L.J., Paviot, Y. (2000) Modèles pour l'intégration de représentations perceptives et motrices dans l'acquisition du langage *XXIIIe Journées d'Étude sur la Parole*, Aussois, 401-404.
- Tchobanov, A. (2002) *Représentations et apprentissage des primitives phonologiques: approche neuromimétique*. Doctorat Sciences du Langage. Université de Paris X.
- Uomini, N.T. (2001) *Étude perceptive des voyelles des langues du monde, identification du trait d'aperture et influence des autres traits*. DEA Sciences du Langage, Université Stendhal, Grenoble.
- Vérin, F. (2001) *Études statistiques de données longitudinales du babillage. Cyclicités, co-occurrences, closions*. TER Sciences du Langage, Université Stendhal, Grenoble.
- Vilain A. (2000) Apports de la modélisation des degrés de liberté articulatoires à l'étude de la coarticulation et du développement de la parole. Doctorat Sciences du Langage. Université de Grenoble III.

