

SUJET PROPOSÉ A
UN CONTRAT DOCTORAL ETABLISSEMENT 2022

UNITE DE RECHERCHE

Nom de l'Unité de Recherche: LLING
N° de l'Unité de Recherche : UMR 6310
Nom du Directeur de l'Unité de Recherche : Hamida DEMIRDACHE

Ecole Doctorale : ELICC

SUJET DE THESE

Intitulé Français du sujet de thèse proposé
200 caractères maximum espaces compris

UN - Adaptation à la variation et traitement de l'incertitude en perception de la parole :
modélisation en parole naturelle et vocodée.

Intitulé Anglais du sujet de thèse proposé
200 caractères maximum espaces compris

UN - Adaptation to variation and entropy in speech perception: modelling processes for
natural and vocoded speech.

Domaine scientifique principal de la thèse	Perception de la parole
Domaine scientifique secondaire de la thèse	
Discipline	Sciences du Langage : Linguistique et Phonétique Générales

DIRECTEUR DE THESE

CIVILITE	Mme
NOM	Delais-Roussarie
PRENOM	Elisabeth
Adresse mail	elisabeth.delais-roussarie@univ-nantes.fr
Date d'obtention HDR	2005
Nombre d'encadrement en cours au 1^{er} octobre 2022	3 (1 en co-tutelle, 2 en co-encadrement)

Co-directeur le cas échéant

CIVILITE	M.
NOM	CROUZET
PRENOM	Olivier
UNIVERSITE D'AFFECTION UNITE DE RECHERCHE	Université de Nantes
Adresse mail	olivier.crouzet@univ-nantes.fr
Date d'obtention HDR	(co-encadrement)
Nombre d'encadrement en cours au 1^{er} octobre 2022	1 (celle-ci exclue)

ARGUMENTAIRE SCIENTIFIQUE

Argumentaire scientifique présentant les enjeux de la thèse :

- **problématique,**
- **contexte,**
- **méthodologie**

1 page maximum

MOTS CLES (5) : perception de la parole, variation, mécanismes adaptatifs, parole naturelle, parole vocodée

Problématique

Les travaux princeps de Ladefoged & Broadbent (1957) avaient mis en évidence l'impact de changements de fréquences de résonance dans le contexte phonétique sur l'interprétation d'une voyelle non modifiée. Par exemple, si dans la phrase « Please say this word again : bet », on élève la fréquence du premier formant (F1, la première fréquence de résonance) sur toute la phrase sauf sur le mot final « bet », un contraste perceptif se met en place qui conduit les auditeurs à percevoir le F1 de « bet » plus bas qu'il ne l'est en réalité. Ce changement perceptif se caractérise par le fait que les auditeurs identifient alors massivement « bit » (/bit/) à la place de « bet » (/bet/) alors même que le mot isolé est bien perçu comme « bet ». Ces résultats ont été répliqués dans des conditions expérimentales diversifiées (Holt, 2006 ; Sjerps, McQueen & Mitterer, 2013) et mettent en avant le caractère très général de ce phénomène. La thèse aura pour objectif de contribuer à la modélisation de ces mécanismes d'adaptation au contexte en s'intéressant notamment aux interactions entre les propriétés de la voix et la catégorisation des voyelles dans la perspective des modèles contemporains de traitement de la variation en perception de la parole.

Contexte

Les travaux princeps de Ladefoged & Broadbent ont donné lieu à des modélisations théoriques qui se sont largement développées dans les années 90 (Goldinger, 1996 ; Mullenix, 1997) et ont conduit à contraster deux grandes catégories de modèles : des modèles dits "à exemplaires" et des modèles de normalisation acoustique pour l'adaptation au locuteur. Des travaux récents remettent cependant en cause ces deux catégories de modèles. En effet, ils reposent principalement sur l'idée que la variation serait de manière générale un frein à la classification perceptive et que l'accumulation d'information devrait permettre de faire ressortir des tendances homogènes. Plusieurs propositions récentes attribuent au contraire un rôle primordial aux processus d'analyse en continu de cette variation pour expliquer la capacité des systèmes perceptifs à s'adapter au changement (Jongman & McMurray, 2011 ; Kleinschmidt & Jaeger, 2015). Ainsi, la modélisation de mécanismes d'adaptation permanente au changement serait essentielle pour comprendre la perception de la parole.

Parallèlement à ces travaux, plusieurs études sur les patients sourds portant un implant cochléaire montrent que ces personnes éprouvent des difficultés particulières à percevoir certaines informations liées à la voix du locuteur. Ainsi, les patients implantés subissent une altération de la reconnaissance des voyelles dans des contextes multi-locuteurs (Chang & Fu, 2006). De même, les patients portant un implant éprouvent des difficultés à classer le genre des locuteurs sur la base d'informations acoustiques liées à la « taille du conduit vocal » (répartition des résonance du conduit vocal sur le spectre de fréquences) alors qu'ils réussissent à utiliser des informations fondées sur la hauteur de la voix en lien avec la fréquence fondamentale de vibration des cordes vocales (f_0 , Fuller et al., 2014).

Capitalisant sur ces travaux, la thèse aura pour objet de modéliser plus précisément les mécanismes impliquant la prise en compte du contexte acoustique dans la catégorisation perceptive des voyelles en étudiant parallèlement les données observées en parole naturelle et dans des simulations d'implants

cochléaires à travers ces phénomènes d'adaptation au contexte acoustique.

Méthodologie

La première année de la thèse portera d'une part sur la finalisation des études de puissance statistique à partir des données préliminaires disponibles, d'autre part sur la mise en place de deux recueils de données en laboratoire dans une tâche de classification perceptive à choix forcé. Ces études permettront d'évaluer dans quelle mesure la classification des voyelles est altérée de manière différente en parole naturelle et dans des simulations d'implants cochléaires (parole vocodée), ce qui permettra d'évaluer le rôle de l'accès aux différents types d'informations acoustiques dans l'émergence de ces mécanismes d'adaptation (informations tonales, spectre de modulation d'amplitude, précision de la résolution spectrale). Le premier recueil de données sera réalisé en Néerlandais à l'UMCG, le second sera réalisé en français au LLING. Les méthodes seront similaires mais permettront d'évaluer la répliquabilité de nos données sur deux langues tout en fournissant des bases pour les recherches ultérieures.

Au cours de la seconde année, nous élaborerons un dispositif d'étude fondé sur le paradigme d'interférence de Garner (Garner, 1974 ; Roark & Holt, 2019). Ce paradigme consiste à faire varier deux paramètres physiques de manière soit indépendante soit reliée (par exemple hauteur de la voix et paramètres acoustiques associés à la catégorie de voyelle) sur lesquels peuvent se fonder les auditeurs pour donner une réponse perceptive. Il sera mis en œuvre afin de documenter comment les mécanismes d'analyse de la voix du locuteur et des catégories phonologiques de la langue peuvent être imbriqués dans les phénomènes d'adaptation.

Les résultats issus de ces deux paradigmes expérimentaux fourniront des éléments de comparaison de l'adaptation à la variation en parole naturelle et en parole vocodée (simulations d'implants cochléaires). En fonction des résultats obtenus, la dernière année de la thèse sera consacrée soit à des recueils de données confirmatoires auprès de patients sourds implantés, soit à des expériences de modélisation des données recueillies par simulations.

Références de l'unité de recherche sur le sujet :

programmes de recherches en cours,

bibliographie sommaire de l'unité et du porteur de projet sur le sujet proposé

1 page maximum

Le laboratoire dispose d'une expertise reconnue dans la mise en place de corpus de données orales (Delais-Roussarie & Post, 2014 ; Delais-Roussarie & Yoo, 2014), ce qui sera essentiel dans l'exploration des diverses sources impactant ces mécanismes d'adaptation perceptive (contexte prosodique, contexte articulatoire segmental, changements de locuteurs, quantité d'*input* nécessaire à la recalibration de l'espace perceptif). Une base de données de phrases destinée à évaluer les changements perceptifs en français a été développée récemment (Leprieur, Crouzet & Gaudrain, 2020). Par ailleurs, notre laboratoire a récemment acquis une cabine insonorisée avec le soutien de la région Pays de la Loire, du CNRS et de l'Université de Nantes (équipement lourd co-financé). La cabine a été installée fin 2018 et nous permet de recueillir des données acoustiques et perceptives dans des conditions expérimentales qui sont essentielles pour la mise en œuvre de ce projet.

Dans le cadre d'une collaboration avec le *University Medical Center Groningen* (UMCG, Pays-Bas ; équipe du Pr. Deniz Başkent) nous avons développé des outils informatiques qui permettent de générer des altérations des sons produits par un locuteur selon différentes dimensions comme la hauteur (f_0) de la voix, la longueur du conduit vocal, la configuration du conduit vocal. Ces méthodes permettent donc de modifier des caractéristiques sonores qui sont liées, d'une part à la voix du locuteur, d'autre part aux distinctions entre catégories phonologiques. Ce travail a fait l'objet d'études

préliminaires dont les résultats portent sur les voyelles du Néerlandais (Crouzet, Gaudrain & Başkent, 2019). Parallèlement à cette collaboration, nous travaillons également avec le *Centre de Recherches en Neurosciences de Lyon* (Dr. Étienne Gaudrain, CRNL, UMR) autour d'un projet sur les interactions entre ces mêmes mécanismes et les effets du contexte sémantique en français. Cette collaboration capitalise elle-même sur le travail effectué à l'UMCG et a donné lieu au développement d'une base de référence statistique (Gaudrain & Crouzet, 2019) et à la présentation des premiers résultats (Crouzet & Gaudrain, 2019). Ces deux collaborations ont reçu des soutiens financiers essentiellement orientés vers des missions scientifiques (co-financement « *Soutien à la Mobilité Internationale (SMI)* » - CNRS et Programme Marie Skłodowska-Curie « *PRESTIGE* », PI : Olivier Crouzet ; « *Mission pour les Initiatives Transverses et l'Interdisciplinarité - MITI* », CNRS InSHS - InSB, PI : Olivier Crouzet). Une partie de ces financements a permis l'acquisition de matériel nécessaire à la conduite des travaux associés (Sonomètre, Calibreur acoustique, Convertisseur Digital-Analogique, licences logicielles Matlab). Nous avons en outre développé une expertise dans les *simulations aléatoires* qui sera essentielle pour la planification statistique *a priori* des recueils de données (Crouzet, 2017 ; 2019). Enfin, les relations que vous avons établies avec ces deux centres de recherche mais aussi avec le service ORL du CHU de Nantes permettront, en fonction des résultats obtenus, de planifier des recueils de données ultérieurs auprès de patients sourds Néerlandophones et Francophones utilisant un implant cochléaire (UMCG et CHU de Nantes) mais aussi de patients présentant une amusie (perte de la sensibilité à la musique, CRNL).

Publications et communications associées :

- Crouzet, O. (2017). Fundamentals and applications of resampling methods for the analysis of speech production and perception data. In *Workshop on statistical methods in phonetic sciences*, a satellite workshop to the *Phonetics and Phonology in Europe (PaPE)* conference, Germany : Universität zu Köln.
- Crouzet, O. (2019, communication invitée). Small Data: Number of Features vs. Number of Observations. In *TAPAS 3rd Training Event – Data Collection, Management and Ethical Practices, Toulouse, FR (TAPAS: Training network on Automatic processing of Pathological Speech ; <https://www.tapas-etn-eu.org/>)* – Horizon 2020 Marie Skłodowska-Curie Actions Innovative Training Network N° 766287– IRIT (Institut de Recherche en Informatique de Toulouse), Université Paul Sabatier.
- Crouzet, O., Gaudrain, E., & Başkent, D. (2019). Perceptual adaptation to formant changes associated with either vowel categories or vocal tract length: Implications for speech perception in cochlear implant users. In *Conference on Implantable Auditory Prosthesis*, Lake Tahoe, CA, USA : University of Southern California.
- Crouzet, O., & Gaudrain, E. (2019). On the articulation between acoustic and semantic uncertainty in speech perception: Investigating the interaction between sources of information in perceptual classification. In *177th Meeting of the Acoustical Society of America*, Louisville, KY, USA : Acoustical Society of America.
- Crouzet, O., Borel, S., & James, C. (2019). Atelier « Implants cochléaires et variation chez l'adulte ». Atelier satellite des 8^{èmes} *Journées de Phonétique Clinique*, Mons, BE : Institut de Recherche en Sciences et Technologies du Langage, Université de Mons, <https://langage.be/JPC/indexjpc.html>.
- Crouzet, O., & Delais-Roussarie, E. (2018). Atelier « Temporalité et séquentialité dans les formes sonores. ». *Atelier international, Université de Nantes – CNRS*, <https://bit.ly/2pLy9u6>.
- Delais-Roussarie, E. & B. Post (2014) Corpus annotation and transcription systems. In Durand, J.; U. Gut & G. Kristoffersen (eds): *Handbook of Corpus Phonology*. Oxford University Press. pp. 46-88.
- Delais-Roussarie, E. & H. Yoo (2014) Corpus and research in phonetics and phonology (I): Methodological and formal considerations. In Durand, J.; U. Gut & G. Kristoffersen (eds):

Handbook of Corpus Phonology. Oxford University Press.

Gaudrain, E., & Crouzet, O. (2019). word2vec model trained on lemmatized French Wikipedia 2018. Tech. Rep., Funding : CNRS International Mobility Support program, 2017 ; European Community PRESTIGE Mobility program N o 2017–2–0044 ; NWO / ZonMW VICI : 918–17–603 ; LABEX CeLyA (ANR–10–LABX–0060) of Université de Lyon, within the program « Investissements d’Avenir » (ANR–16–IDEX–0005) operated by the French National Research Agency (ANR), <http://doi.org/10.5281/zenodo.3241447>.

Leprieur, L., Crouzet, O. & Gaudrain, E. (2020). Une base de données de phrases en français pour l’étude du rôle conjoint des incertitudes sémantique et acoustique dans la perception de la parole. Actes des 33^e Journées d’Études sur la Parole, Nancy, France. pp.353-361.

Connaissances et compétences requises pour le futur doctorant :

Le candidat ou la candidate devra avoir acquis une bonne expérience du travail expérimental en perception de la parole, ce qui inclut notamment la conception de matériel linguistique adapté aux questions scientifiques, les recueils de données orales et les activités de segmentation / annotation d’un corpus oral.

Il ou elle aura idéalement acquis des compétences dans un ou plusieurs langages de programmation adaptés au traitement de signaux audio (Matlab, Octave, Python / Numpy) et aura une expérience des méthodes d’analyse statistique descriptives et inférentielles des données avec le langage de statistique R ou S.

Il ou elle devra par ailleurs avoir des aptitudes relationnelles avérées en raison des nombreux contacts humains qui devront être mis en oeuvre pour les recueils de données auprès des participants.

Intérêt du projet quant aux perspectives d’insertion professionnelle du-de la doctorant-e :

Ce projet, de part son ouverture à des problématiques sociales impactant les situations de communication et son intégration à des collaborations impliquant des centres de recherche spécialisés dans les troubles de l’audition, ouvre des perspectives aussi bien au niveau de l’enseignement et de la recherche académique que pour le développement de solutions qui pourront concerner les aspects thérapeutiques aussi bien que technologiques des troubles de la communication orale. Il pourrait aussi conduire le ou la candidate à développer des compétences dans le domaine de l’apprentissage des langues secondes en mettant à disposition des conceptions originales de l’adaptation à la variation.

DOCTORAT / EDUCATION

BRETAGNE / LANGAGES, INTERACTIONS

LOIRE / COGNITION, CLINIQUE

