



**HAL**  
open science

## Interaction entre voix et parole dans la perception de voyelles vocodées

Olivier Crouzet, Etienne Gaudrain, Sem Foreman, Deniz Baskent

► **To cite this version:**

Olivier Crouzet, Etienne Gaudrain, Sem Foreman, Deniz Baskent. Interaction entre voix et parole dans la perception de voyelles vocodées. 3èmes Journées Perception Sonore, Jun 2017, Brest, France. , 2017. hal-01542966

**HAL Id: hal-01542966**

**<https://hal.univ-brest.fr/hal-01542966>**

Submitted on 20 Jun 2017

**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.

# Interaction entre voix et parole dans la perception de voyelles vocodées.

Crouzet, Olivier<sup>1,3</sup> Gaudrain, Étienne<sup>2,3</sup> Foreman, Sem<sup>3</sup> Başkent, Deniz<sup>3</sup>

<sup>1</sup> Laboratoire de Linguistique de Nantes – LLING / UMR6310, Université de Nantes / CNRS, France  
<sup>2</sup> Centre de Recherche en Neurosciences de Lyon – CRNL / UMR 5292, CNRS, Inserm U1028, Université Lyon 1, France  
<sup>3</sup> University Medical Center Groningen – University of Groningen, Groningen, The Netherlands.

## Introduction

- Analyse perceptive de la parole  $\Rightarrow$  2 dimensions distinctes :
  1. Propriétés de la voix...
    - Caractéristiques vocales,
    - Identité du locuteur,
    - genre,
    - état émotionnel...
  2. Classification en catégories linguistiques...
    - Classes phonémiques,
    - Traits distinctifs...
- Impacts du traitement de la « voix » sur la perception de la parole :
  - Difficultés à séparer des locuteurs dans des situations de « cocktail-party »  $\Rightarrow$  Interaction entre voix et séparation de sources [5];
  - Performances de reconnaissance des mots réduites quand diversité de locuteurs [11]  $\Rightarrow$  Interaction entre voix et classification phonologique;
  - Modèles de normalisation [13], modèles à exemplaires [7];
- Perception de parole vocodée :
  - Baisses de performance importantes dans des contextes multi-locuteurs [3]...
  - Impact sur les difficultés des patients portant un implant cochléaire ?

## Objectifs

- Comprendre les mécanismes d'interaction entre perception de la voix (caractéristiques du locuteur) et de la parole (catégories linguistiques);
- Ne pas raisonner en termes de *performance* (« y-a-t-il de moins bonnes performances de catégorisation avec de la parole vocodée –resp. un implant cochléaire– ? »);
- Mais en termes de *modélisation* des mécanismes perceptifs :
  - Les deux dimensions interagissent-elles ? (intégralité des dimensions vs. indépendance)
  - Si oui, est-ce de manière plus ou moins marquée qu'avec de la parole naturelle ?
  - L'une d'elles est-elle « dominante » ?
  - Les différences entre parole naturelle et parole vocodée peuvent-elles rendre compte des difficultés observées en reconnaissance de parole ?

## Travaux en cours

- Préparation en cours d'une série d'expériences...
- Étudier les mécanismes d'adaptation au locuteur (Longueur du Tractus Vocal, VTL) avec de la parole vocodée;
- Réplication du travail princeps de [10] sur l'adaptation en reconnaissance de parole : Impact de modifications contextuelles de VTL sur l'identification des voyelles [10, 15];
- Mise en place d'une expérience de type « paradigme d'interférence de Garner » [6, 8] permettant d'analyser les modalités d'interaction entre deux dimensions liées;

VTL : influences homogènes sur les résonances (formants);

Catégories vocaliques : changements de forme du conduit vocal pouvant donner lieu à des changements homogènes ou non-homogènes des résonances [4, 9];

## Contenu de la présentation

1. Analyses acoustiques des voyelles du Néerlandais;
2. Variations inter-locuteurs;
3. Organisation de l'espace formantique en lien avec les canaux de vocodage;
4. Identification des voyelles adéquates pour l'élaboration du matériel;

## Analyses acoustiques des voyelles du Néerlandais

- Données issues de [14] et [12];
- Fréquences des 3 premiers formants de voyelles du Néerlandais produites en contexte /hVt/ isolé;
- 75 locuteurs (25 femmes vs. 50 hommes);
- Les données sont extraites de Praat [2];
- Trois types de classification pour les voyelles du Néerlandais : courtes / longues (durées physiques?), tendues / relâchées (définition phonétique?), *checked* / *free* (définition distributionnelle);
- Données acoustiques de [1] : relation non-parallèle entre tendu / relâché et courte / longue (= *checked* / *free*);

## Espaces vocaliques et Variation

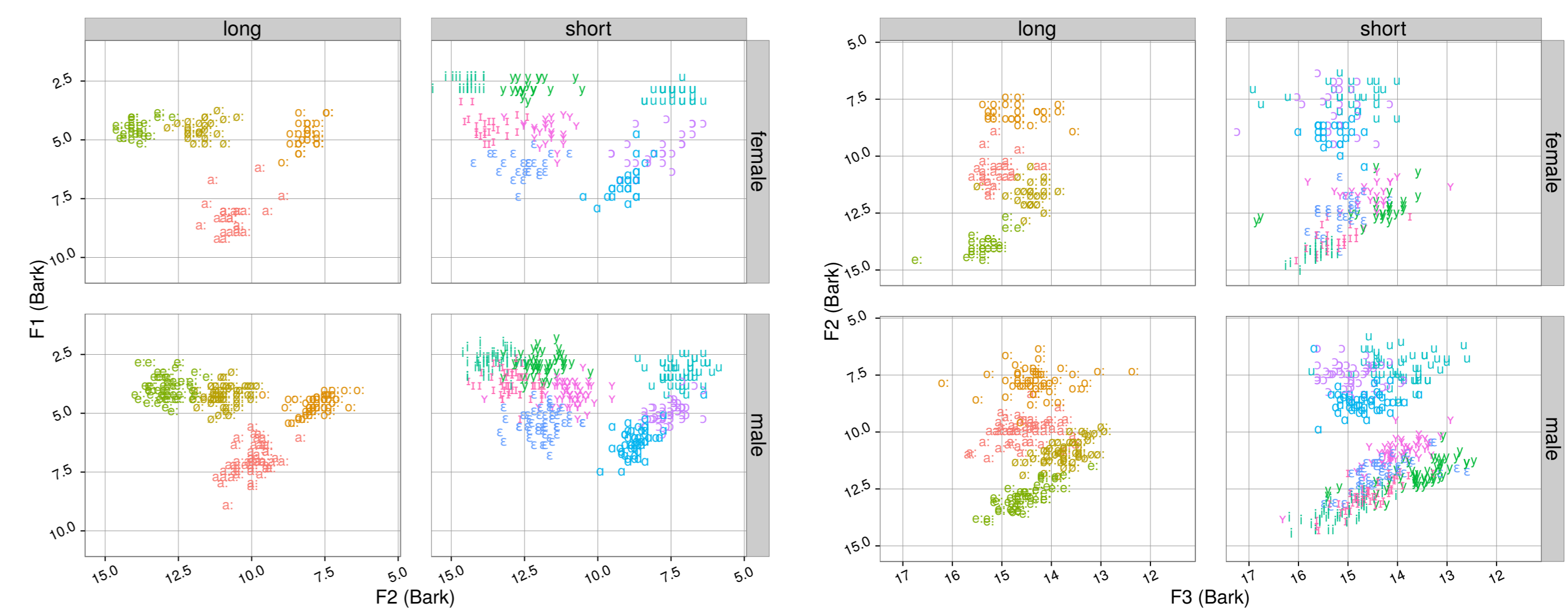


FIGURE 1: Mesures des formants ( $F_1$  et  $F_2$ , en Bark) des voyelles du Néerlandais.

FIGURE 2: Mesures des formants ( $F_2$  et  $F_3$ , en Bark) des voyelles du Néerlandais.

## Espaces vocaliques et canaux de vocodage (12 canaux)

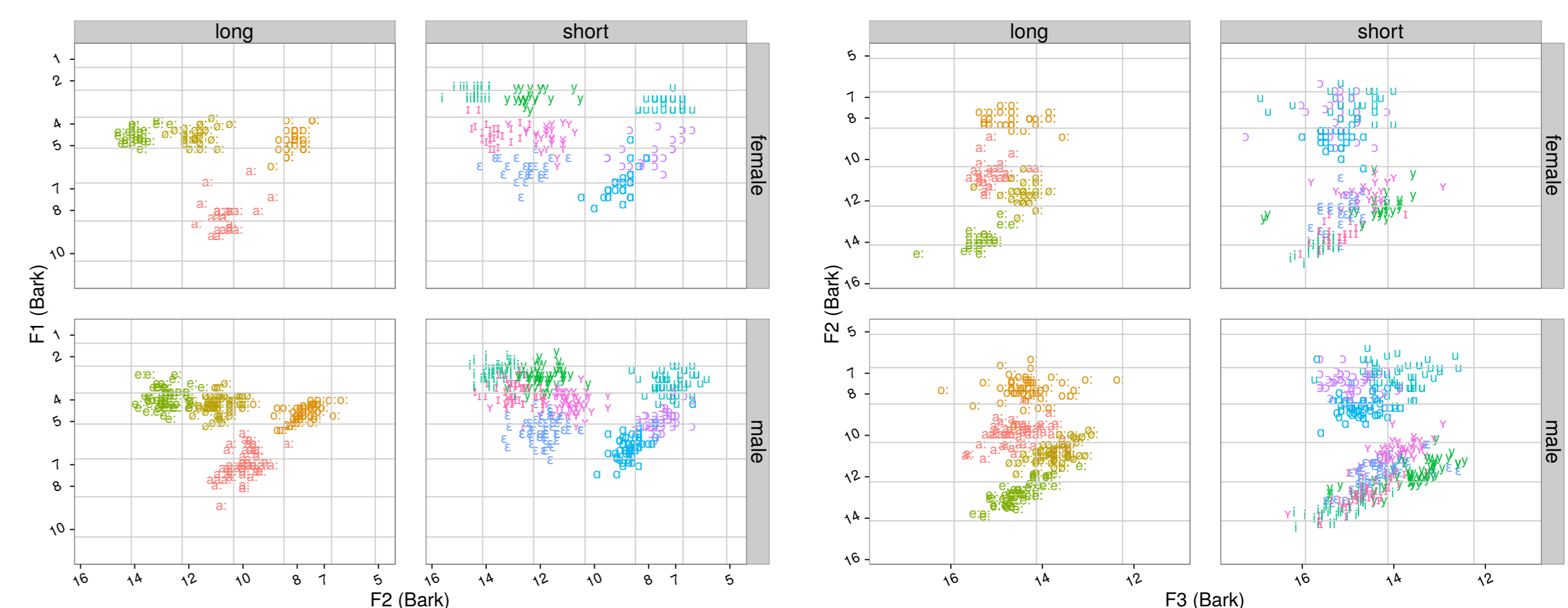


FIGURE 3: Relations de fréquences entre un vocoder à 12 canaux et les formants ( $F_1 \sim F_2$ ) des voyelles du Néerlandais.

FIGURE 4: Relations de fréquences entre un vocoder à 12 canaux et les formants ( $F_2 \sim F_3$ ) des voyelles du Néerlandais.

## Sélection des voyelles

- Utilisation de ces données pour la sélection des voyelles étudiées;
- Focalisation sur les voyelles courtes afin d'éviter aux auditeurs de recourir à la durée pour l'identification;
- Autre argument : plus de chevauchements entre les espaces (plus intéressant pour étudier les changements de catégorie perceptive);
- Permettent également d'étudier deux types de relations entre les formants :
  1. Pour certaines paires de voyelles,  $F_1$  et  $F_2$  évoluent de manière cohérente dans la même direction (p. ex. /o/  $\rightarrow$  /a/, /u/  $\rightarrow$  /e/). Similitude avec les effets du variation de VTL sur les résonances;
  2. Pour d'autres,  $F_1$  et  $F_2$  évoluent de manière cohérente dans des directions opposées (p. ex. /i/  $\rightarrow$  /e/ or /u/  $\rightarrow$  /a/. Contrast avec les effet de changement de VTL.

## Travail en cours

- Manipulation des paires de voyelles sélectionnées : génération d'un continuum acoustique entre deux extrêmes afin de faciliter les changements de catégories;
- Évaluation des courbes d'identification psychométriques pour la sélection des versions adéquates. Critères :
  1. Taux d'identification correcte satisfaisants en isolation;
  2. Taux de changement de catégorie en parole naturelle suffisamment élevés en contexte induisant une adaptation;

## Références

- [1] P. ADANK, R. VAN HOUT et R. SMITS. "An acoustic description of the vowels of Northern and Southern Standard Dutch". Anglais. In : *The Journal of the Acoustical Society of America* 116.3 (2004), p. 1729. [2] P. BOERSMA et D. WEENING. *Praat : Doing phonetics by computer*. Computer program. Version 5.4. 2014. [3] Y.-p. CHANG et Q.-J. FU. "Effects of Talker Variability on Vowel Recognition in Cochlear Implants". Anglais. In : *Journal of Speech Language and Hearing Research* 49.6 (déc. 2006), p. 1331-1341. [4] G. FANT. *Acoustic theory of speech production : With calculations based on X-Ray studies of Russian articulations*. Anglais. The Hague : Mouton, 1960. [5] C. D. FULLER et al. "Gender Categorization Is Abnormal in Cochlear Implant Users". Anglais. In : *Journal of the Association for Research in Otolaryngology* 15.6 (déc. 2014), p. 1037-1048. [6] W. R. GARNER. *The Processing of Information and Structure*. 1<sup>re</sup> éd. Psychology Press, août 1974. [7] S. D. GOLDINGER. "Echoes of echoes? An episodic theory of lexical access". In : *Psychological Review* 105 (1998), p. 251-279. [8] N. KAGANOVICH, A. L. FRANCIS et R. D. MELARA. "Electrophysiological evidence for early interaction between talker and linguistic information during speech perception". In : *Brain Research* 1114 (2006), p. 161-172. [9] R. D. KENT. "Vocal tract acoustics". In : *Journal of Voice* 7.2 (1993), p. 97-117. [10] P. LADEFOGED et D. E. BROADBENT. "Information conveyed by vowels". In : *The Journal of the acoustical society of America* 29.1 (1957), p. 98-104. [11] J. W. MULLENIX, D. B. PISONI et C. S. MARTIN. "Some effects of talker variability on spoken word recognition". In : *The Journal of the Acoustical Society of America* 85.1 (1989), p. 365-378. [12] D. J. P. J. van NIEROP, L. C. W. POLS et R. PLOMP. "Frequency Analysis of Dutch Vowels from 25 Female Speakers". Anglais. In : *Acta Acustica united with Acustica* 29.2 (août 1973), p. 110-118. [13] D. B. PISONI. "Some Thoughts on « Normalisation » in Speech Perception". In : *Talker Variability in Speech Processing*. Sous la dir. de K. JOHNSON et J. W. MULLENIX. San Diego : USA : Academic Press, 1997. [14] L. C. W. POLS, H. R. C. TROMP et R. PLOMP. "Frequency analysis of Dutch vowels from 50 male speakers". Anglais. In : *The Journal of the Acoustical Society of America* 53.4 (1973), p. 1093. [15] M. J. SJERPS, J. M. MCQUEEN et H. MITTERER. "Evidence for precategorical extrinsic vowel normalization". Anglais. In : *Attention, Perception, & Psychophysics* 75.3 (avr. 2013), p. 576-587.