Project

Reconnaissance des mots parlés et impact de la variabilité acoustique

Sophie Dufour[1], Jonathan Grainger[2]

(1) LPL, (2) LPC

Abstract

Un problème majeur auquel est confronté notre système de perception est de reconnaître un même mot sous différents modes de réalisation. Les modèles abstractionnistes supposent que les mots dans le lexique mental sont stockés sous la forme de séquences linéaires consistant en des traits (Marslen-Wilson & Warren, 1994), des phonèmes (McClelland et Elman, 1986) ou des syllabes (Mehler, Dommergues, Frauenfelder & Segui, 1981). Le signal de parole serait dans un premier temps converti en une séquence de segments discrets écartant ainsi tous les détails acoustiques fins non pertinents pour l'identification, et le résultat de cette analyse serait ensuite projeté sur les représentations symboliques abstraites stockées en mémoire. Au contraire, selon les modèles exemplaristes (Goldinger, 1998), les mots seraient stockés sous la forme de traces acoustiques détaillées encodant des informations fines liées par exemple à la voix du locuteur. Chaque mot serait alors associé à de multiples « tokens » et reconnaître un mot consisterait à trouver l'appariement le plus proche dans une vaste collection d'exemplaires. A ce jour, nous disposons dans la littérature de preuves expérimentales en faveur de l'un ou de l'autre type de représentations (Goldinger, 1998; McQueen, Norris & Cutler, 2006) et des modèles hybrides de la reconnaissance des mots parlés combinant représentations abstraites et détaillées ont été proposées (McClennan & Luce, 2005; Pierrehumbert, 2001). L'enjeu maintenant n'est donc plus de déterminer quel type de représentations rend compte au mieux de la reconnaissance des mots parlés, mais plutôt de déterminer sous quelles conditions et à quel moment l'un ou l'autre type de représentations a le plus de chance de dominer et d'influencer le traitement. C'est précisément l'objectif de cette recherche. A l'aide d'une étude en EEG, nous examinerons plus précisément a) le décours temporels de l'utilisation d'indices acoustiques fins liés à un changement de voix, b) et comment les effets liés à un changement de voix interagissent avec des facteurs lexicaux comme la fréquence des mots.

Publications

-

Fiche-résumé contribution CREx

Variation Acoustique

Reconnaissance de mots parlés et impact de la variabilité acoustique.

Investigateurs : S. Dufour (LPL), J. Grainger (LPC)

Durée : 9 mois (Février 2016 – Octobre 2016)

Contribution :

- Passation des expériences EEG,
- Prétraitement des données EEG acquises
- Analyses statistiques des résultants,
- Préparation des figures pour publication.

Objectif : *Examiner le décours temporels de l'utilisation d'indices acoustiques fins en effectuant une expérience éléctrophysiologique (EEG) pour analyser des potentiels évoqués (PE).*

Passation – La totalité des passations a été effectuée par l'assistante ingénieur et l'ingénieur CREx. Au total, 40 participants ont effectué l'expérience ; chaque séance expérimentale a duré environ 1h30.

Prétraitement – Les données électrophysiologiques ont été traitées avec la boîte à outil Matlab, EEGLAB. L'ingénieure CREx a effectué le traitement au début avec la chercheuse pour lui introduire l'utilisation de ce logiciel pour le traitement des données EEG. Le traitement a été effectué de manière semi-automatique en utilisant un script de prétraitement programmé en Matlab par l'ingénieure CREx. Les étapes de traitement ont compris le ré-échantillonnage, le filtrage (filtre FIR), la segmentation des données continues selon les conditions, le repérage et le suppression des électrodes défectueuses, l'interpolation des électrodes supprimées ainsi que la correction des artéfacts en utilisant PCA-ICA. Le repérage des électrodes et des segments bruités a été effectué à l'aide d'un outil programmé en Matlab par l'ingénieure CREx (une description se trouve via ce lien : https://blricrex.hypotheses.org/ressources/eeg/pre-processing-for-erps#badchannels).

Analyse – Les fenêtres temporelles d'intérêts ont été définies en appliquant un analyse statistique non-paramétriques (test de permutation avec correction fdr) sur la grande moyenne par condition expérimentale. Cette analyse a été effectuée dans Matlab par un script programmé par l'ingénieure CREx (https://github.com/blri/CREx_EEG/tree/master/EEG_GroupAnalysis). Une fois les fenêtres d'intérêts repérées, la chercheuse a effectué des analyses statistiques paramétriques (ANOVA mesures répétées) sur les données PE avec le logiciel *Statistica* (www.statsoft.com) pour étudier les interactions. Les moyennes par fenêtre d'intérêt et régions topographiques ont été calculées et exportées dans un fichier en format Excel à l'aide d'un outil programmé par l'ingénieure CREx (`CREx_WritetoExcel.m`). Ces fichiers ont pu être importés dans Statistica pour effectuer les analyses statistiques.

Les analyses des données électrophysiologiques ont donné des résultats cohérents avec l'expérience comportementale initiale, répliquant ces résultats tout en précisant la dynamique temporelle cérébrale à l'œuvre. Ces résultats ont permis à la chercheuse de préparer une publication.

Diffusion – L'ingénieur CREx a aidé la chercheuse pour la rédaction de la section « Méthodologie » et la préparation des figures (figure 1) de l'article titré, « On the locus of talker-specificity effects in spoken word recognition : An ERP study » . L'article a été soumis à Language, Cognition and Neuroscience et est actuellement en révision.

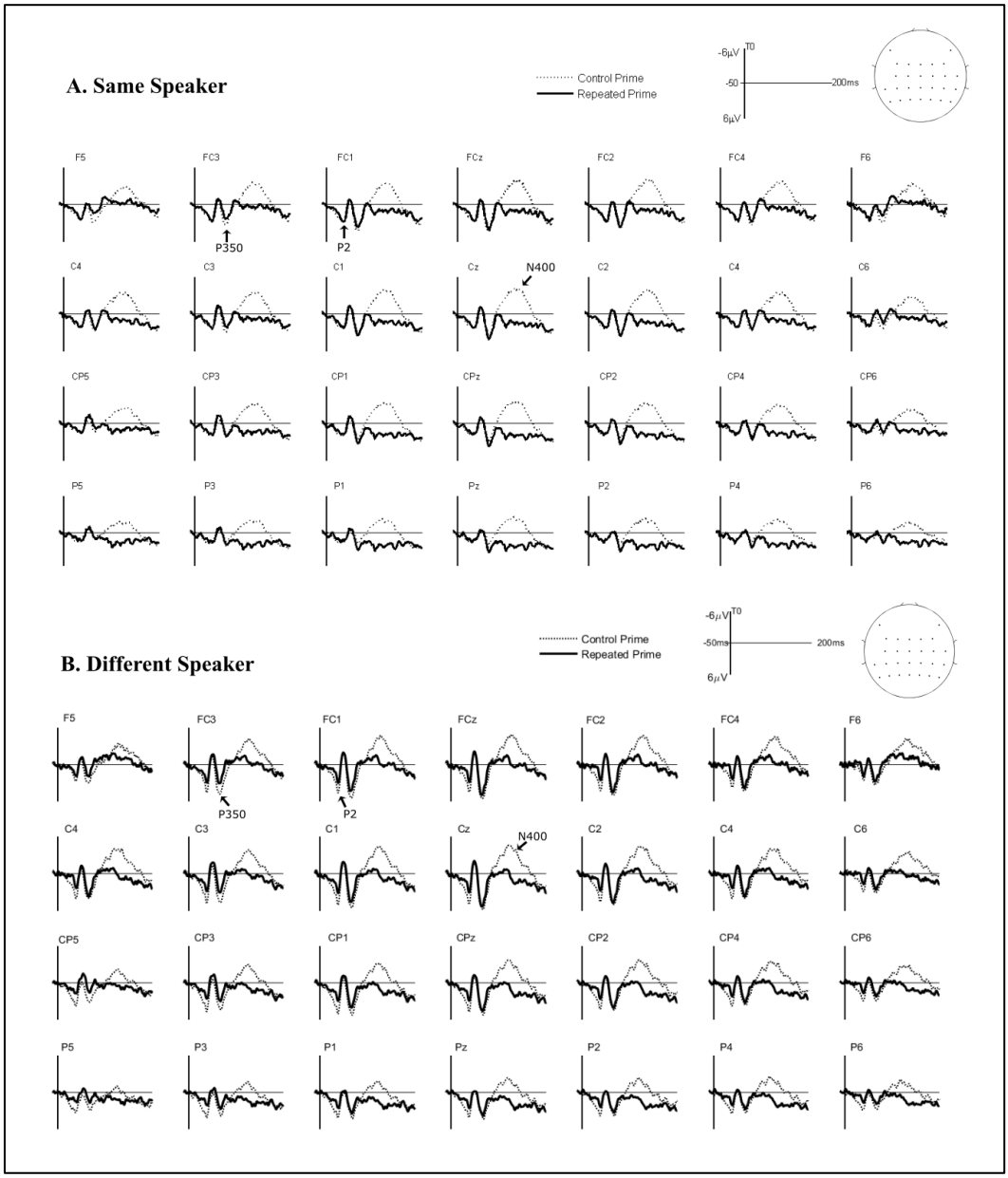

Figure 1 : Résumé des résultats des analyses PE correspondant à la condition « Mot de Basse Fréquence : sans changement de voix » montrant un effet **P2-like**, **P350** et **N400** pour la condition « *same speaker* » et « *different speaker* » provoqué par un changement d'amorce (*Control Prime* vs. *Repeated Prime*).