Acoustical cues used to segment phonemically identical speech sequences:
An EEG study

Stéphane Pota[1], Véronique Boulenger[1], Emmanuel Ferragne[1,2], Elsa Spinelli[3] & Fanny Meunier[1]

1) Laboratoire Dynamique du Langage, CNRS, Université Lumière, Lyon, France 2) Centre de Linguistique Inter-langues, de Lexicologie, de Linguistique Anglaise et de Corpus, EA 3967, Paris, France
3) Laboratoire de Psychologie et NeuroCognition, CNRS, Université Pierre Mendès France, Grenoble, France Institut Universitaire de France.   **Contact: stephane.pota@gmail.com**

INTRODUCTION

- There are no clear word boundaries in spoken language.
- Due to elision, some spoken utterances in French are phonemically identical (e.g., "l'allocution" vs. "la locution", both [lalokysjɔ̃]).
- Correct segmentation into discrete word units is necessary for comprehension.
- Some acoustic differences between members of ambiguous sequences may be exploited by listeners.
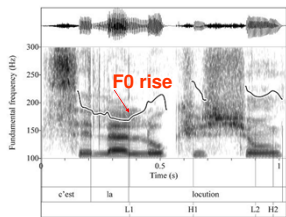
C'est la locution   **C'est l'allocution**

[selalokysjɔ̃]
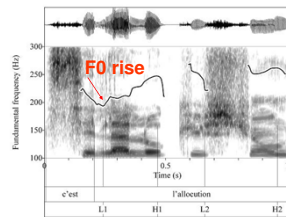
For example, among those, there are clear intonational differences between the two sequences. There is often a rise in fundamental frequency beginning at the left edge of the first content word syllable.

- However psycholinguistic models of speech recognition such as TRACE [1] and SHORTLIST [2] postulate that the acoustical stream is converted into phonemes before lexical access and therefore leave the system deaf to fine acoustic cues.

=> In this study, we examined the electrophysiological correlates of fine acoustical details that could cue segmentation.

METHODS EXPERIMENTS 1 & 2

Participants

32 healthy right-handed French native speakers, aged 18-24, with no hearing or language impairment.

Paradigm

A modified version of the Oddball paradigm [3].

Stimuli

Experiment 1: with CV syllable

Standard: "l'a", *Deviant 1*: "la" and *Deviant 2*: "li"

Experiment 2: with words

Standard: "l'allocution", *Deviant 1*: "la locution" and *Deviant 2*: "l'illocution"

EEG

32 electrodes – referenced to linked mastoids – [1-30 Hz] – segmentation [-100; 900ms]

ERPs analysis

- ERPs time-locked to target CV/word onset
- 40-ms-window centered at peak latency [200-300 ms]
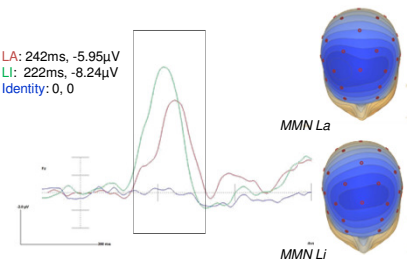
Statistical analysis

Repeated-measures ANOVAs (Condition x Spatial Domain x Lateralization)

- Conditions: Deviant Identity, LA, and LI
- Spatial Domain: Frontocentral and Centroparietal
- Lateralization: Left, Right

RESULTS

CV Experiment

MMN LA: 242ms, -5.95µV
MMN LI: 222ms, -8.24µV
MMN Identity: 0, 0

MMN La

MMN Li

- Mean ERP amplitudes significantly different from zero for LA and LI (p<.001).

- MMN LA longer latency (+20ms) than MMN LI (p<.001).

- MMN LA smaller peak (2.3µV) and mean (1.24µV) amplitudes than MMN LI (respectively p<.001 and p=.07).

- The MMN was fronto-central, whatever the condition (p<.001).

- Lateralization in the right fronto-central region for the conditions LA and LI (p=.003).

Word Experiment

MMN LA: 251ms, -3.37µV
MMN LI: 257ms, -5.87µV
MMN Identity: 0, 0
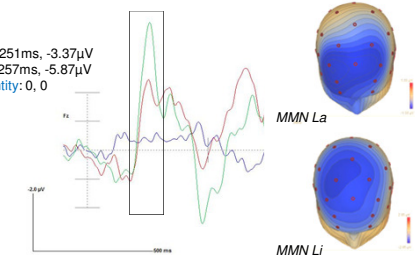
MMN La

MMN Li

- Mean ERP amplitudes significantly different from zero in all conditions (p<.001).

- No difference in peak latencies between conditions.

- MMN LA smaller peak (2.5µV) and mean (1.85µV) amplitudes than MMN LI (respectively p<.001 and p<.001).

- The MMN was fronto-central, whatever the condition (p<.001).

- Comparable lateralization in the right fronto-central region for both conditions (p=.06).

CONCLUSIONS

➤ Even in a passive listening situation (no task), the acoustic cues that differentiate "l'a" from "la" are relevant to the system recognition: **importance of fine acoustic cues (e.g. F0) that could be used for correct word segmentation.**

➤ Our results complement previous behavioral findings [4] by showing that these **fine acoustic details are available as early as phonemic information**, and therefore could be used **automatically** to disambiguate between phonemically identical spoken words and thus to bias lexical access.

➤ Acoustic cues may play a robust role in French word segmentation and, quite probably, in other languages. Role in semantic integration?

REFERENCES

[1] McClelland, J.L. & Elman, J.L. (1986). The TRACE model for speech perception. *Cognitive Psychology, 18*, 1-86.
[2] Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*, 189-234.
[3] Brunellière, A., Dufour, S., Nguyen, N., & Frauenfelder, U.H. (2009). Behavioral and electrophysiological evidence for the impact of regional variation on phoneme perception. *Cognition, 111(3)*, 390-396.
[4] Spinelli, E., Grimault, N., Meunier, F., & Welby, P. (2010). An intonational cues to word segmentation in phonemically identical sequences. *Attention, Perception, & Psychophysics, 72(3)*, 775-787.