ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Speech Communication
Session 5aSCb: Production and Perception II: The Speech Segment (Poster Session)

5aSCb10. Multidimensional scaling of English fricatives using the acoustic change complex of Electroencephalogram (EEG) recordings
Paul Iverson*, Marta Mulyak and Anita Wagner

*Corresponding author’s address: Speech Hearing and Phonetic Sciences, University College London, 2 Wakefield St, London, WC1N 1PF, London, United Kingdom, p.iverson@ucl.ac.uk

In EEG recordings, there is a characteristic P1-N1-P2 complex after the onset of a sound, and a related complex, called the Acoustic Change Complex (ACC), when there is a change within a sound (e.g., a formant transition between two vowels). In the present study, the ACC was measured for all possible pairs of eight sustained voiced and voiceless English fricatives, in EEG recordings from native speakers of British English. The magnitude of the ACC was used as a similarity measure for multidimensional scaling (MDS), producing a two-dimensional perceptual space that related to both voicing and place of articulation. The results thus demonstrate that this combination of ACC and MDS can be effective for mapping multidimensional phonetic spaces at relatively early levels of auditory processing, which may be useful for evaluating the effects of language experience in adults and infants.

Published by the Acoustical Society of America through the American Institute of Physics
INTRODUCTION

Research on speech perception has long sought to separately measure auditory sensitivity and phonemic categorization (e.g., categorical perception experiments), but in practice all levels of processing are active when a listener gives a post-perceptual behavioral response. That is, one can modulate the role of phonological categorization by manipulating tasks and stimuli (e.g., reducing discrimination sensitivity peaks for vowels when the tasks have fewer demands on memory and attention; Fry et al., 1962), but it is implausible that one can eliminate the role of higher-level linguistic processing given that speech recognition is so automatic. Early research on the mismatch negativity (MMN) of EEG recordings was exciting because it appeared to offer a direct measure of auditory processing that occurred too early in time for later linguistic processes to play a role, but now it appears that the MMN can be affected by higher-level linguistic processes in addition to auditory processing (e.g., Näätänen et al., 2007). The present study examined a lower-level EEG response instead, the Acoustic Change Complex (ACC; Martin and Boothroyd, 1999). In EEG recordings, there is a characteristic P1-N1-P2 complex after the onset of a sound, and the ACC is a related complex that occurs when there is a change within a sound (e.g., a vowel formant change). The ACC has not been examined extensively, but it offers the possibility of measuring the sensitivity of the auditory system to acoustic change in speech.

One advantage of the ACC is that it is very time efficient compared to the MMN; one can measure the ACC for concatenated sequences of stimuli that change about twice a second, allowing for the measurement of 120 stimulus changes in only one minute of recording (about the number of changes that are needed to be averaged in order to measure the ACC reliably; Martin, 2010). Most research using the ACC has tried to exploit this efficiency in clinical assessments of hearing (e.g., Martin et al., 2008). The present study instead used the efficiency of this measure in order to measure perceptual sensitivity for large numbers of stimulus pairs, generating similarity matrices that can be analyzed via multidimensional scaling (MDS; Kruskal, 1964). Our aim is to use the ACC to generate MDS spaces for speech that are analogous to those from our previous behavioral work (Iverson et al., 2003) but that more directly measures early auditory sensitivity. The present study played a set of eight voiced and voiceless English fricatives in concatenated sequences to native English-speaking listeners. Listeners heard sequences based on all possible pairs during EEG recording, and the magnitude of the ACC was used as an index of similarity. The resulting similarity matrix was then analyzed by non-metric MDS.

METHOD

Subjects

The listeners were 12 native speakers of southern British English who had no self-reported hearing or neurological disorders.

Stimuli and presentation

Four voiceless (/f/, /θ/, /s/, /ʃ/), and four voiced (/v/, /ð/, /z/, /ʓ/) fricatives were recorded from an adult female native English speaker. They were recorded in a sustained fashion (i.e., holding the frication for several seconds) such that a relatively static section of 433 ms in length could be found by spectrogram inspection. This static section was further processed to flatten the amplitude envelope. These sections were concatenated with overlapping raised-cosine ramps of 5 ms in duration. A set of 56 sequences was created (all possible pairs of the eight stimuli, omitting the same stimulus paired with itself).

EEG recording and analysis

A 64-channel Biosemi EEG system was used for the recordings, with the electrode location Cz being used to assess the ACC. The recordings were analyzed using a combination of SPM and EEGLab. The recordings were high-pass filtered at 0.15 Hz, low-pass filtered at 30 Hz, artifact-corrected using Independent Components Analysis, epoched, threshold artifact-rejected, and averaged. The ACC response was almost entirely restricted to the P2 component (about 200 ms after the stimulus change) with little P1 or N1, so the magnitude of the ACC for each pair was calculated by taking the RMS amplitude in the region of P2.
RESULTS AND DISCUSSION

The matrix of ACC magnitudes for the stimulus pair was mapped into two dimensions using non-metric MDS, placing stimulus pairs with low ACC magnitudes close together in the MDS solution and those with large ACC magnitudes far apart. The solution is displayed in Figure 1. The stimuli are separated vertically on a voiced-voiceless dimension (e.g., /v/ vs. /f/) and horizontally on a dental-sibilant dimension (e.g., /f/ vs. /s/). This thus demonstrates that the ACC is sensitive to the same kinds of phonetic dimensions that are thought to be important in speech perception, and that there is a reliably monotonic relationship between ACC magnitude and acoustic/perceptual similarity that allows for MDS analysis.

FIGURE 1. Two-dimensional MDS solution for voiced and voiceless English fricatives.

REFERENCES