FAST PHONOTOPIC MAPPING WITH OSCILLATION-BASED FMRI
– PROOF OF CONCEPT -

Mikkel Wallentin, Torben E. Lund, Camilla M. Andersen, Roberta Rocca

Department of Linguistics, Cognitive Science and Semiotics, Aarhus University; Interacting Minds Centre, Aarhus University; Centre of Functionally Integrative Neuroscience, Aarhus University Hospital, Denmark

INTRODUCTION

The auditory cortices contain tonotopic representations of sound frequency (e.g. Saenz and Langers, 2014), but how about functional organization of speech sounds? Previous experiments have hinted at a phonotopic map (e.g. Formisano et al., 2008), but the temporal and spatial resolution of neuroimaging has made it difficult to construct protocols for mapping phonemes in the brain. Here, we describe a novel oscillation-based method, using a fast fMRI protocol, building on findings that the BOLD signal having higher temporal fidelity than hitherto imagined (Lewis et al., 2016).

MATERIALS AND METHODS

Stimuli consisted of pairs of phonemes making up syllables. Two conditions (9v5c and 9c5v) combined 9 Danish vowels/consonants with 5 consonants/vowels to create a total of 45 syllables per condition. In each condition, consonants and vowels were repeated in a fixed order, i.e. in condition 9v5c, a vowel would be repeated on every 9th trial, and the consonant would be repeated every 5th trial, making every combination new for the 45 trials, but at the same time creating two highly predictable rhythms for vowels and consonants.

Sessions consisted of 6x4 blocks: [6x9v5c, 6x9c5v, 6x9v5c, 6x9c5v], lasting 18 minutes. Three sessions were acquired in a single participant (female, 25 years) for this preliminary study, yielding a total of 72 blocks.

fMRI data was acquired at 3T using a whole-brain fast acquisition sequence (TR = 371ms, multi-band EPI acquisition) to capture signal changes at syllable resolution.
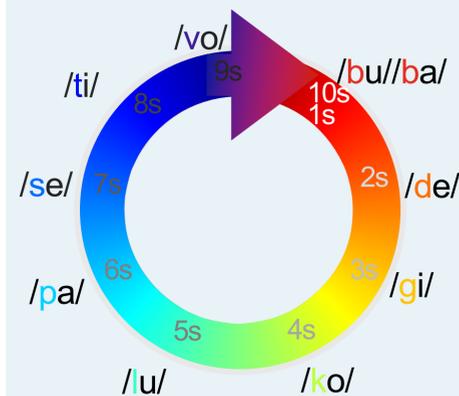
DATA ANALYSIS

Data were modelled using sine and cosine waves at the presentation rate for vowels and consonants, i.e. either 1/9 Hz or 1/5 Hz.
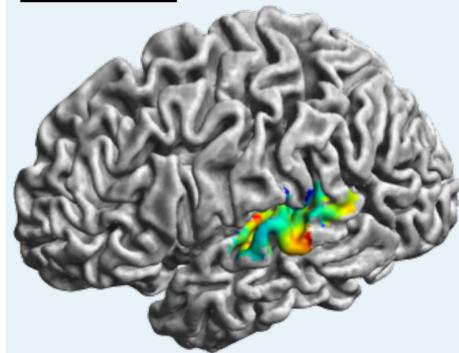
Fitted sine and cosine waves were used to generate a phase map for each 45 s block. Phase is indicative of the delay in a voxel's responsiveness to a repetitive stimulus, thus here suggesting differences in phoneme responsivity. Phase maps from each block were used to perform a multivariate classification test.

The phase maps for the 72 blocks were divided into two parts. The first half was used to conduct a search-light analysis (using the nilearn-package in Python) in order to select the 500 most predictive voxels. These voxels were used in a subsequent pattern classification test on the 2nd half of the phase maps. Both steps involved a Gaussian naïve-bayes classifier. Cross-validation and permutation tests were used to determine significance.

fMRI BOLD signals that oscillates with these frequencies are assumed to be related to processing of phonemes

Vowels and consonants are combined and repeated at two different frequencies (1/9 Hz and 1/5 Hz)

Oscillation Condition 9c5v

/vo/ /bu/ /ba/ /ti/ /de/ /se/ /gi/ /pa/ /ko/ /u/

9 consonants: f=1/9 Hz
5 vowels: f=1/5 Hz

Phase map: 9 consonants at f=1/9 Hz

Oscillation Condition 9v5c

/kɔ/ /la/ /ba/ /gø/ /de/ /dɛ/ /gi/ /by/ /ki/ /u/
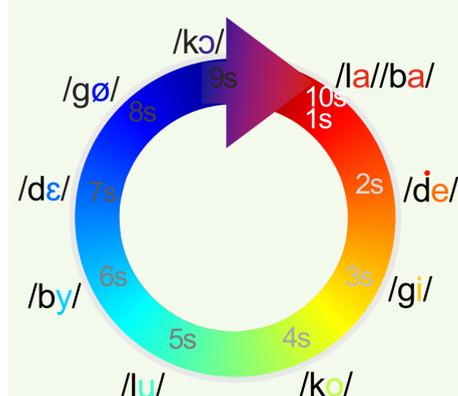
9 vowels: f=1/9 Hz
5 consonants: f=1/5 Hz
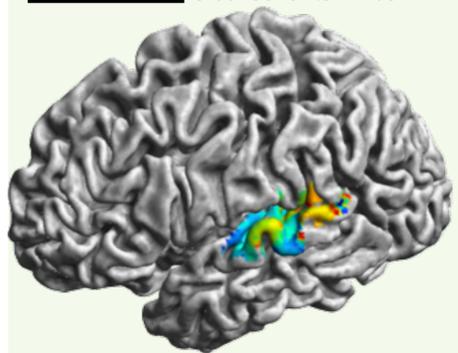
Phase map: 9 vowels at f=1/9 Hz

Time (t in seconds)
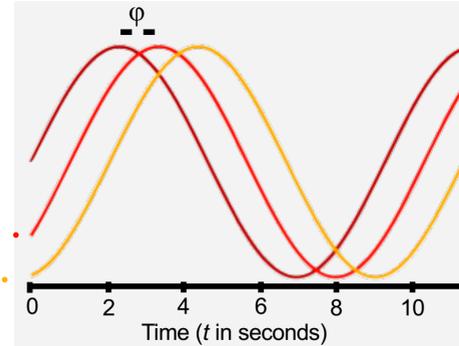
$$Y(t)=\alpha\sin(2\pi ft)+\beta\cos(2\pi ft)$$
$$\varphi=\arctan2(\alpha, \beta)$$

Brain signal delay: Phase (φ)
(index of differential phoneme sensitivity)
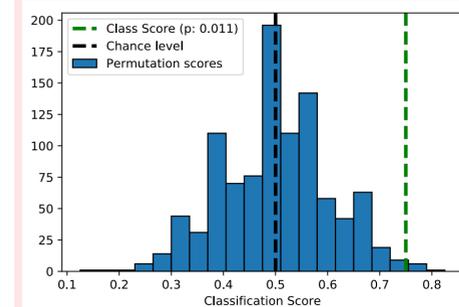
Machine learning:
9 consonants or 9 vowels pattern?

Delay in the oscillation (known as "phase" - φ) is indicative of brain areas sensitive to particular phonemes.
Fast data acquisition allows for the estimation of phase.

Phase maps can be used to classify vowel or consonant components of stimuli

RESULTS

Univariate SPM-analysis across all data showed a significant difference (P<0.05, FWE-corrected) between consonants and vowels in the left auditory cortex. The same areas also differentiated between phonemes oscillating at 1/5 Hz and 1/9 Hz, regardless of phoneme type.

Classification tests were able to classify 45 second phase maps from the 9s condition into consonants and vowels with 77% accuracy (p<0.03) and 5s condition with 75% accuracy (p<0.02).

CONCLUSION

This protocol provides the first step towards mapping a phonotopic "fingerprint" across multiple phonemes simultaneously at the individual participant level. This map may be used to predict native language, foreign language exposure as well as literacy. It is also the first step towards making use of fMRI signal for decoding at near speech rate temporal resolution.

REFERENCES

Formisano, E., De Martino, F., Bonte, M., Goebel, R., 2008. "Who" is saying "what"? Brain-based decoding of human voice and speech. Science (New York, NY) 322, 970-973.
Lewis, L.D., Setsompop, K., Rosen, B.R., Polimeni, J.R., 2016. Fast fMRI can detect oscillatory neural activity in humans. Proc Natl Acad Sci U S A 113, E6679-E6685.
Saenz, M., Langers, D.R., 2014. Tonotopic mapping of human auditory cortex. Hear Res 307, 42-52.

POSTER COPY: mikkel@cc.au.dk

SCHOOL OF COMMUNICATION AND CULTURE
AARHUS UNIVERSITY

DComm
Deictic Communication