Primary auditory cortex's vowel representation

Andrea Fingerhut[1] & Mikael Roll[2]
[1]Medical Radiation Physics, Lund University
[2]Centre for Languages and Literature, Lund University

Abstract

Sound frequencies are represented in the primary auditory cortex (PAC) in a tonotopic structure which can be measured using functional magnetic resonance imaging (fMRI). A previous study has investigated vowels [a] and [i] and seen a correlation between the vowels' activation and the activation of simple tones corresponding to the vowels' formant frequencies. Other vowels have not yet been studied. In this study, we are investigating [a], [ɛ], [i], and [u] and compared those with activation of simple tones corresponding to their formant frequencies. This is ongoing work, and only five volunteers have participated. The preliminary results vary from high correlation between areas activated to low or no correlation. More data has to be collected to draw any further conclusions.

Introduction

The primary auditory cortex (PAC) is the first instance of sound processing in the brain's cortex. In PAC, different neurons are more or less sensitive to specific sound frequencies, and their frequency sensitivity is spatially dependent. Gradients have been found spanning from higher frequencies to lower to higher again, creating a mirror-symmetric structure along the axis of Heschl's gyrus (Formisano et al., 2009). This representation of frequencies in the brain, called 'tonotopy,' can be seen in Figure 1.

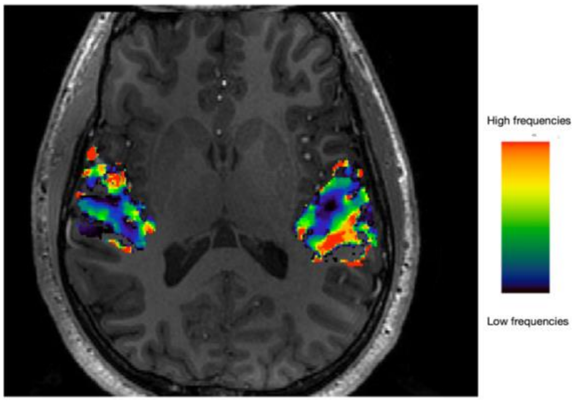

Figure 1: Figure shows a mirror-symmetric tonotopy map. Lower frequencies can be seen in the middle, surrounded by higher frequencies.

Vowels invovle resonance peaks around different frequencies on a frequency spectrum, called 'formants.' First and second formants are required to distinguish different vowels. Fisher (2019) studied vowels [i] and [ɑ], and found that they activate regions in the tonotopic structure of PAC corresponding to their formant frequencies. However, to more firmly establish a relation between vowel formants and the tonotopic structure of PAC, more than two vowels should be tested. Therefore, in this project, we are investigating the effects of the vowels [ɑ], [ɛ], [i], and [u].

Using high resolution functional magnetic resonance imaging (fMRI), a comparison between the activation pattern for simple tone and vowel activations can be made. FMRI uses MRI to study brain activation. It takes advantage of the difference in blood oxygen level to separate different states, commonly between performing a task and resting. In our case, the task is to listen to sounds, and rest is silence. By repeating sound and silence multiple times in a block paradigm, the brain areas connected to the specific task can be distinguished. This design was chosen for its robustness but is, on the other hand, very time-consuming.

Method

To map out the tonotopic structure, high resolution and fast fMRI are needed, and imaging parameters had to be optimized to reach a submillimeter resolution (0.9 mm isotropic resolution) while keeping a short imaging time (2s per volume). A 7 Tesla Philips Achieva scanner was used for this purpose. The same imaging parameters were used for both the vowel and tonotopy measurements.

So far, the auditory processing of five volunteers has been measured. They are all right-handed native speakers of Central Swedish.

Twelve simple tones of different frequencies were selected for the tonotopy measurement to match the vowel formants (300, 375, 525, 600, 675, 825, 1575, 1650, 2400, 2700, 2850, and 3300 Hz). The tones were jittered with ± 75 Hz, which has been shown to increase activation and reduce exhaustion. Each stimulus had a duration of 8s and was repeated four times. The interstimulus interval was 8s. The order of the frequencies was randomized, and the sound level was adjusted for the individual volunteers.

Three formants for every vowel were included (Table 1), as seen for [ɑ] in Figure 2. To simplify the comparison, we created synthetic vowel stimuli (complex tones, F0 = 75 Hz) in Praat that were as acoustically simple as possible but were perceived as vowels. The formant frequencies for Central Swedish in Kuronen (2000) were selected. For the complex tones to be recognizable as vowels, we had to include two harmonics surrounding the formant frequencies (formant frequency ± 75 Hz). The vowels were chosen to maximize the difference between formant 1 and 2. All the formant frequencies can be seen in table 1. We tested both a standard spectral envelope and keeping formant frequencies at the same amplitude as F0 and found little difference in acceptability of the vowels. Therefore, to increase activation for the formants, we kept their amplitude at the same level as F0.

The same block paradigm structure was used for the vowel measurement, with the same duration and number of repetitions. A task was added during this measurement where the volunteers had to answer which vowel they heard. This was both to keep their attention and to make sure that they perceived the correct vowel.

Table 1: Table shows formants 1, 2 and 3 for vowels [a], [ε], [i] and [u].

Vowel	F1 [Hz]	F2 [Hz]	F3 [Hz]
[ɑ]	525	825	2475
[ɛ]	600	1650	2700
[i]	300	2400	3300
[u]	300	675	2700

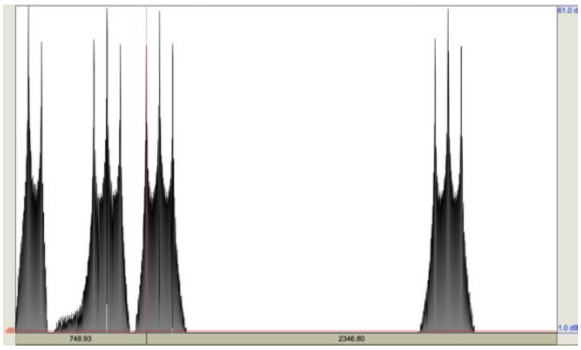

Figure 2: Figure shows frequency spectrum for vowel [a]. The fundamental frequency and the three formants can be seen with the ± 75 Hz frequency surrounding them.

For every volunteer, a tonotopy map was created by giving every voxel in the image a frequency based on which frequency had the highest activation (best frequency). Formant-specific region of interests were created for every vowel by only including voxels with a best frequency matching one of the formants. This gave a template for every vowel showing, if the hypothesis is correct, the expected vowel-activated area. The images were masked based on Harvard-Oxford cortical structural atlases, where Heschl's Gyrus was selected. The mean t-value in the overlap between the format-specific regions of interest and vowel t-maps was calculated.

Results and discussion

The results have not been consistent. We had a higher overlap between the expected formant-specific regions of interest (vowel template) and vowels for one volunteer and no or small overlap for the rest. Table 2 shows the volunteer with the highest overlap between vowel and the right template maps created from the tonotopy measurement. Two of the vowels, [ɑ], [ɛ] had the highest mean t-value in the overlap with the right vowel template. Vowel [i] had the same mean t-value in the overlap between the template for vowel [i] and [u] and vowel [u] had the highest mean t-value in the overlap with the template for vowel [ɛ]. The inconsistency in the results could of course be because the hypothesis of tonotopic vowel representation might be incorrect, but the effect is tiny, and we need to collect more data to either discard or conclude anything. Heschl's Gyrus is also small, and a little head movement or signal displacement can lead to wrong interpretations, although this can partly be solved by using robust preprocessing. The method is still under development and data will be collected during the coming period.

Table 2: Table mean t-value in overlapping area between template region of interested and vowels activation for [a], [ɛ], [i] and [u].

Template	[ɑ]	[ɛ]	[i]	[u]
[ɑ]	2.6	2.4	2.4	2.3
[ɛ]	2.7	2.9	2.5	2.6
[i]	2.2	2.0	2.4	2.4
[u]	2.6	2.7	2.3	2.4

Acknowledgements

This work was supported by Knut and Alice Wallenberg Foundation (grant number 2018.0454), Swedish Research Council (grant number 2018.00632), Crafoord Foundation (grant number 2017.0006) and Marcus and Alice Wallenberg Foundation (grant number 2018.0021).

References

Fisher J, Dick F, Levy D & Wilson S (2019). Neural representation of vowel formants in tonotopic auditory cortex. *Neuroimage*, 178: 574-582.

Formisano E, Kim D, Salle F, Moortele P, Ugurbil K & Goebel R (2003). Mirror-symmetric Tonotopic Maps In Human Auditory Cortex. *Neuron*. 40: 859-869.

Kuronen, M (2000). Vokaluttalets akustik i sverigesvenska, finlandssvenska och finska. *Studia Philologica Jyvdsfyldensia*, 49. Dissertation, University of Jyvaskyla.