Cortical Processing of Spoken Vowel-Consonant and Consonant-Vowel Syllables

Alexis G Hervais-Adelman, D Timothy Ives, Roy D Patterson

Centre for the Neural Basis of Hearing, Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, Cambridge CB2 3EG

Introduction

A recent study by Uppenkamp *et al*[1] contrasted cortical responses to natural and synthetic vowels with spectrally-matched nonspeech stimuli. They showed that the initial stage of human speech processing is located bilaterally, inferior to Heschl's Gyrus (HG), in superior temporal sulcus (STS).

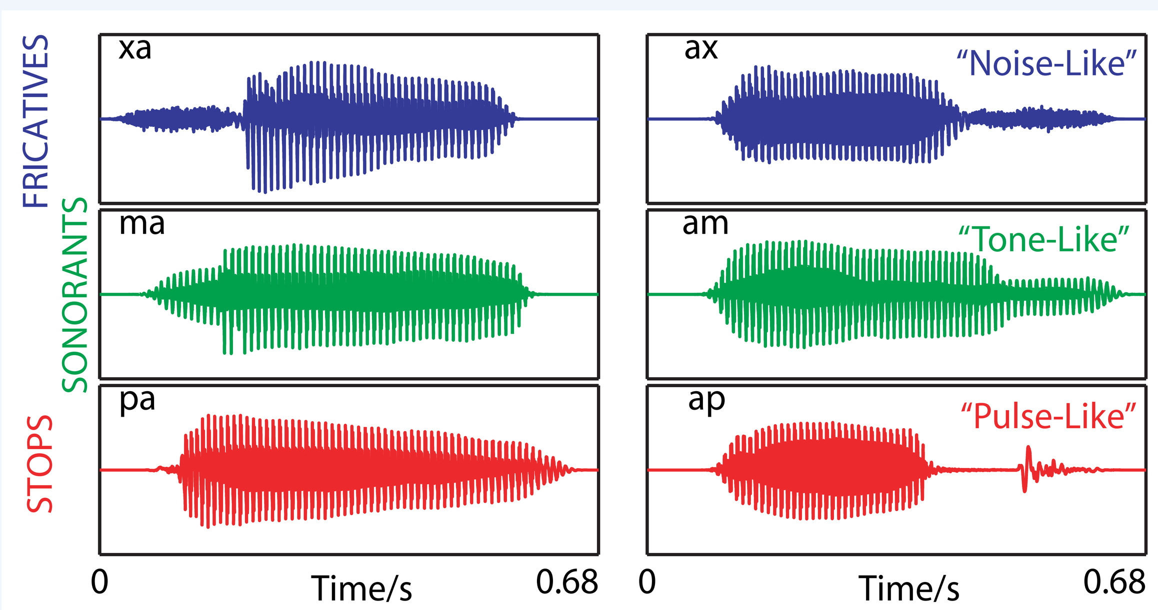
This region may be an initial stage of speech processing, but it is unclear whether it is specific to vowels or general to all speech.

In the present study, we extend this investigation by studying cortical responses to consonant-vowel (CV) and vowel-consonant (VC) pairs.

By comparing activations elicited by CV and VC stimuli with those produced by vowels alone, we hoped to reveal the cortical locus or loci of consonant processing.

Stimuli

CV and VC pairs can be considered as vowels with an additional acoustic event prepended or appended. We used three classes of consonant: Fricatives (F), Sonorants (So) and Stops (St, which can be thought of as noise-like, tone-like and pulse-like, respectively. These were combined with the 5 canonical English vowels "a", "e", "i", "o" and "u".



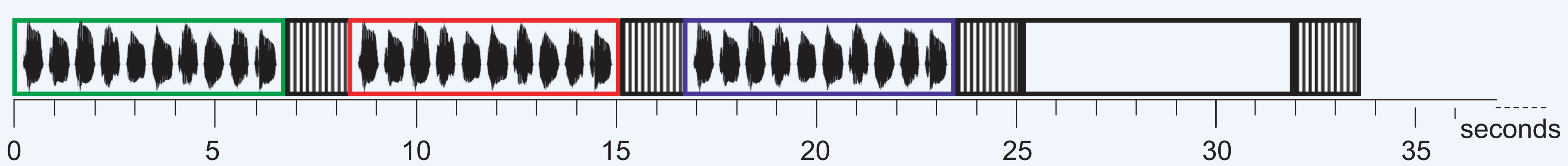
CV, VC and Vowel (V) stimuli were taken from a syllable database[2] recorded by a male speaker.

We included a nonspeech condition: musical rain (MR) and silence, yielding a total of 9 conditions: CVF, CVSo, CVSt, VCF, VCSo, VCSt, V, MR and silence.

Procedure

Sparse imaging[3], block design, TR: 8400 ms, TA: 1600 ms. 3T Bruker Medspec Scanner. Slice thickness: 4 mm, 1mm interslice gap. 21 axial slices, angled away from the eyes. 12 participants.

24 sets of 10 randomly-selected stimuli from a given condition were concatenated. These were presented in the silent intervals between scans.



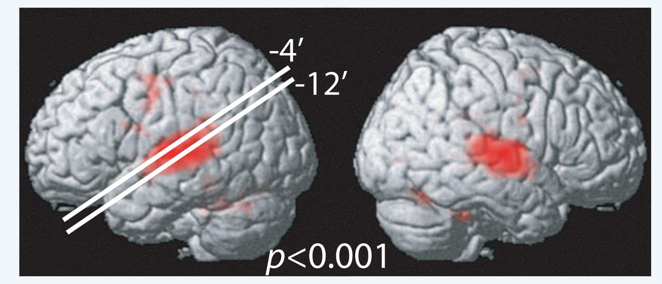
3 scanning sessions, each consisting of 8 blocks from each condition, 72 stimuli per session, presented in a randomised order. Session order was pseudo-randomly ascribed to participants.

Task: Button press in response to a quieter item in the sequence. 3 blocks of each sound condition had one of the 5th-9th item attenuated at -10dB. These trials were randomly distributed throughout the experiment.

Results

2-step analysis in SPM5. Random effects (2nd level) analyses significant at both FDR-corrected $p<0.05$ and uncorrected $p<0.001$. They are shown projected onto a canonical structural image.

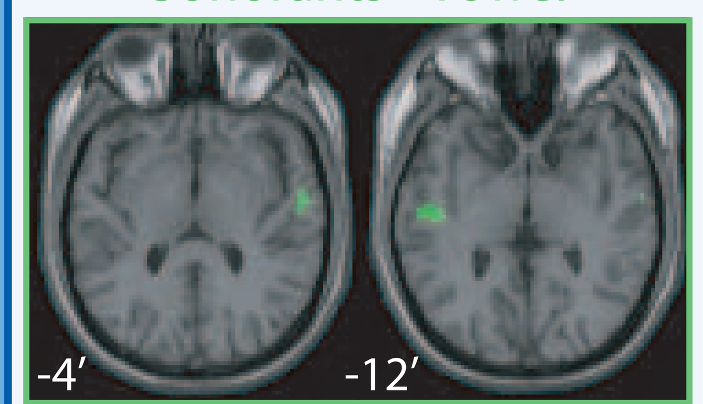
All Sound - Silence



The control contrast shows the typical pattern of bilateral superior temporal gyrus activation in response to sound.

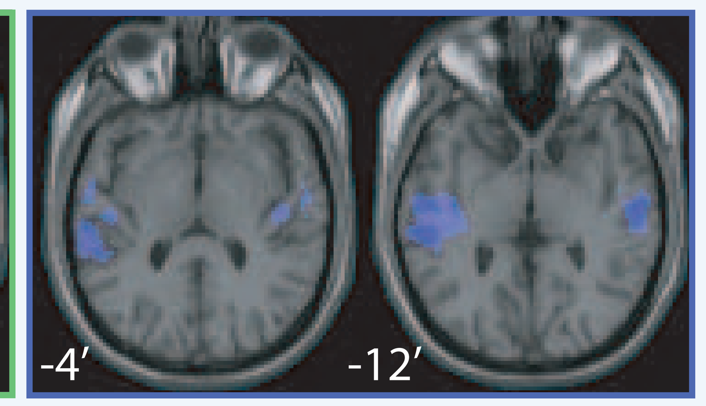
The following contrasts collapse across CV and VC. They are displayed on the two marked axial slices angled along the superior temporal lobe.

Sonorants - Vowel



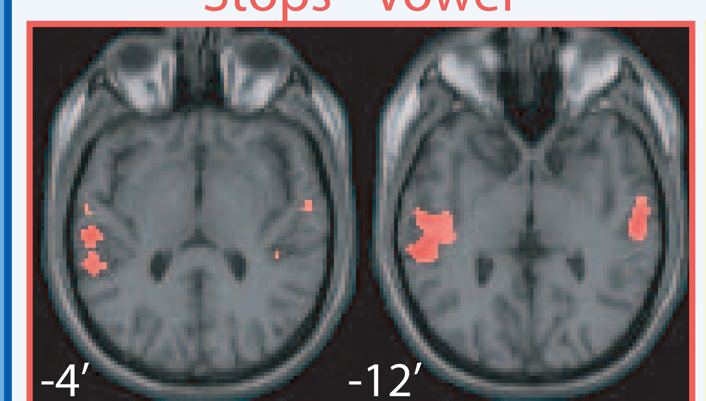
Sonorants produce activation in right anterolateral HG, and left middle STS.

Fricatives - Vowel



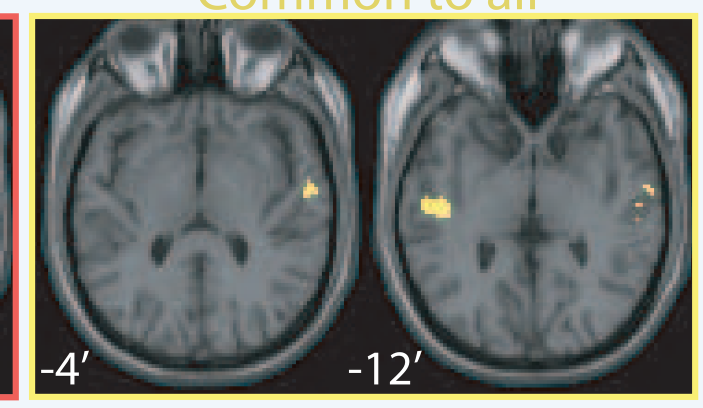
Fricatives produce activation in left Planum Temporale (PT) and bilateral activation in STS.

Stops - Vowel



Stops also produce bilateral STS activation and some left PT activation.

Common to all



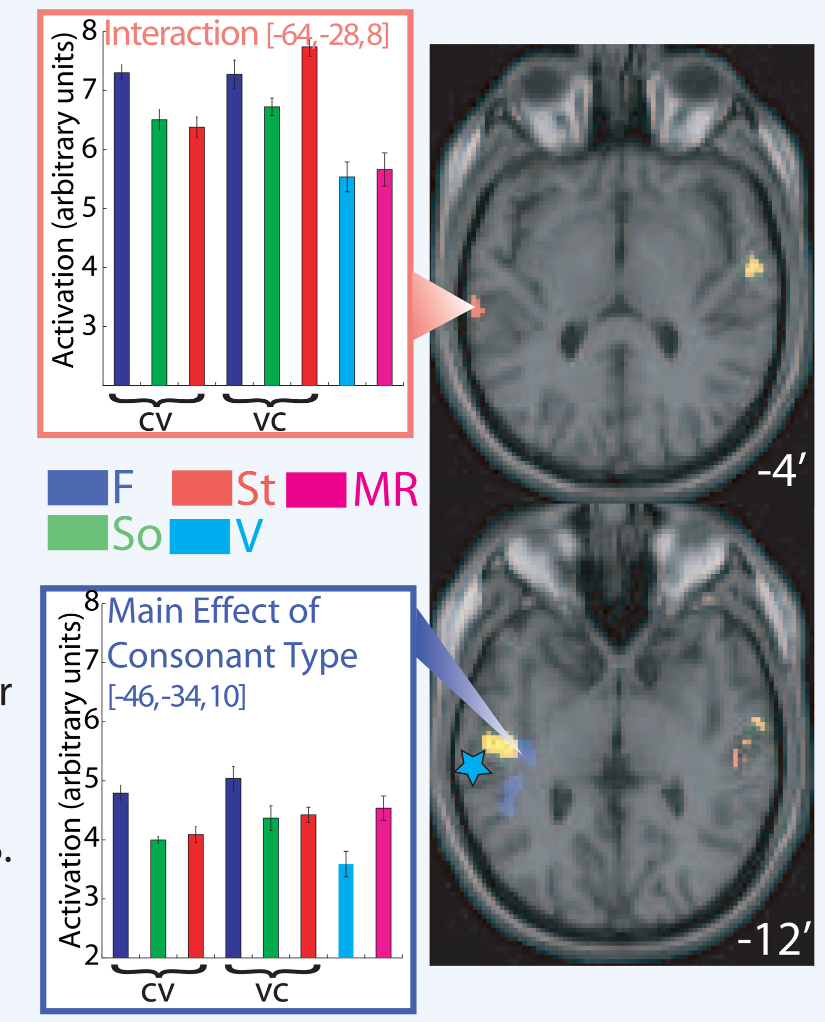
Overlap of responses to the three consonant types. This region does not show differential activation in response to different consonants

CV/VC pairs show a significant difference in response when compared with vowels. This corresponds to the additional processing related to the presence of a consonant in the CV and VC syllables.

Results (cont.)

An ANOVA was carried out to test for a main effect of consonant type, Order (CV vs VC) and the interaction between the two. There was no main effect of order. However, a significant main effect of consonant type and consonant by order interaction were found.

These are shown below alongside the overlap of the 3 consonants.



The interaction is driven by a greater response to VCSt syllables than CVSt. VCSt have a brief silent interval between V and C. This region may respond to acoustic onsets and hence be more active in this condition.

A main effect of consonant type is driven by the significantly greater activation in response to fricatives than stops or sonorants, along deep STS. This may be due to fricatives being a broadband noise-like stimulus, having more overall energy than stops or sonorants.

★ Vowels - MR [-62,-28,4] There is a significantly greater activation in response to vowels than MR in mid STS, consistent with existing results.

Consonant processing occurs outside the region found by Uppenkamp et al[1]. This suggests that the initial stages of vowel and consonant processing occur separately, consistent with the regions being found be responsible for acoustic, rather than speech-specific, processing of these stimuli.

Conclusions

These results demonstrate that initial stages of consonant processing occur separately from initial stages of vowel processing. The different activations patterns observed in response to different consonant types can be explained in terms of the acoustic differences between the stimuli. Further experiments geared towards multivariate analyses may find distinct activation patterns in response to different consonant types.

References:

[1]Uppenkamp, S., Johnsrude, I. S., Norris, D., Marslen-Wilson, W., & Patterson, R. D. (2006). Locating the initial stages of speech-sound processing in human temporal cortex. Neuroimage, 31(3), 1284-1296.

[2]Ives, D. T., Smith, D. R., & Patterson, R. D. (2005). Discrimination of speaker size from syllable phrases. J Acoust Soc Am, 118(6), 3816-3822.

[3]Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., et al. (1999). "Sparse" temporal sampling in auditory fMRI. Hum Brain Mapp, 7(3), 213-223