# The Perception of Voice-Onset Time: An fMRI Investigation of Phonetic Category Structure

Emily B. Myers and Sheila E. Blumstein

*Brown University, Department of Cognitive and Linguistic Sciences*

## Introduction

Speech perception research has shown that although listeners identify stimuli along an acoustic-phonetic continuum in a categorical-like fashion, listeners are sensitive to distinctions within a phonetic category, perceiving some members of the category as better exemplars than others.

While a number of studies have addressed the neural basis of phonological processing (Zatorre et al., 1992, Burton et al., 2000), less is known about the involvement of neural systems underlying phonetic categorization.

This issue was explored by using event-related fMRI as subjects identified tokens along a voice onset time (VOT) continuum. Of interest were three main questions:

1. Is the left superior temporal gyrus (STG) responsive to differences between acoustically similar phonetic exemplars?

2. Do the patterns of activation differ for good members of the phonetic category compared to ambiguous ("boundary") stimuli?
(Endpoint tokens vs. Boundary tokens)

3. What neural mechanism underlies the process of mapping a phonetic token to its category or phonetic label?
(Endpoint tokens vs. Within-Category tokens)

## Methods

**TASK 1: Phonetic categorization task (/da/ or /ta/?)**
--Stimuli: Synthesized speech stimuli ranging from 0 msec (/da/) to 40 msec (/ta/) in 10 msec VOT steps (Figure 1.)

**TASK 2: Tone identification task (high or low tone?)**
--Stimuli: two pure tones at 910 Hz and 1320 Hz

**PARTICIPANTS**: 13 adult native speakers of English, one subject excluded due to image acquisition error.

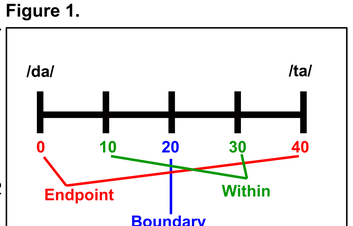
**EXPERIMENTAL DESIGN:**
Rapid event-related fMRI using a sparse-sampling design while subjects engaged in the phonetic categorization and tone identification tasks.

**Scans:**
**--EPI functional scan:** 15 3x3x5mm slices, covering perisylvian areas
**--MPRAGE** anatomical scan for coregistration

**Data Analysis:**
Used **AFNI** for all data analysis.
**Planned comparisons:**
**Endpoint vs. Boundary**
**Endpoint vs. Within**
**Endpoint vs. Tones (not shown)**

Data thresholded at a voxel level p<0.025, cluster size of 42 voxels, cluster-level threshold of p<0.01

Figure 1.



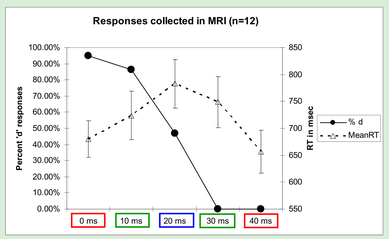

## Results

### Behavioral

**Main effect of VOT on % identification**
**--Subjects show typical 'categorical' identification function (Figure 2.)**

**Main effect of VOT on reaction time:**
**--Subjects show sensitivity to within-category VOT distinctions. (Significant difference between all VOTs except between 20 & 30 msec VOT) (Figure 2.)**

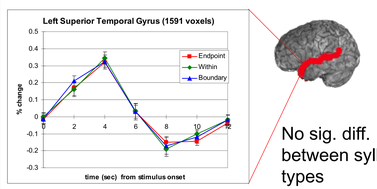
Figure 2.




### Functional

Figure 3.

Region of Interest analysis of the left STG

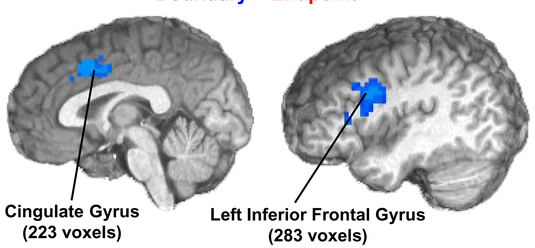



No sig. diff. between syllable types

Left STG activation is **not** modulated by VOT differences.

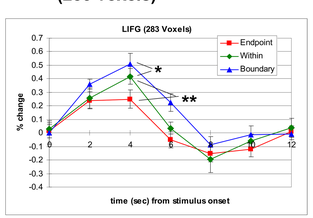
Figure 4.

**Boundary > Endpoint**




**Above:** Clusters activated for the Endpoint vs. Boundary comparison.

**Right:** Time course of activation for all syllable stimuli within the pictured left IFG cluster.




**ANOVA results based on times 2, 4 & 6**
***Endpoint vs. Within F(1,11)=8.037, p<0.016**
****Within vs. Boundary F(1,11)=8.389, p<0.015**

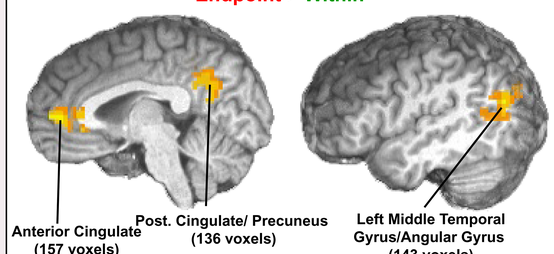
**DISCUSSION: Boundary > Endpoint**
**Activation in the LIFG and cingulate increased as stimuli approached the phonetic boundary.**

**LIFG:** This modulation suggests that the LIFG has a role in resolving phonetic category membership, particularly under conditions of phonetic competition.

**Cingulate:** This pattern is consistent with the view that the cingulate is engaged in conditions of response competition.

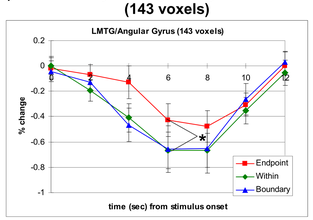
Figure 5.

**Endpoint > Within**




**Above:** Clusters activated for the Endpoint vs. Within comparison.

**Right:** Time course of activation for all syllable stimuli within the pictured left MTG cluster.




**ANOVA results based on times 4, 6 & 8 sec.**
***Endpoint vs. Within F(1,11)=18.507, p<0.001**
**n.s. Within vs. Boundary F(1,11)=0.016**

**DISCUSSION: Endpoint > Within**

**LMTG/Angular Gyrus:** An area in the posterior tip of the left MTG and extending into the angular gyrus was deactivated for all stimulus types, but less deactivated for the 'best' exemplars of the phonetic categories than for all other exemplars in the stimulus set. This may reflect the 'best match' of the acoustic input to either a phonetic category or the label for that category.

## Conclusion

**Activation in the left STG was not modulated by stimulus type**

The left STG plays a role in the initial acoustic analysis of the basic properties of speech (cf. Binder et al., 2000), but does not distinguish between good and poor category members.

**Both the left inferior frontal gyrus and cingulate showed increasing activation as stimuli approached the phonetic category boundary.**

Activation in the **LIFG** may reflect either (or both):

--**Selection** between competing phonetic categories (Thompson-Schill et al., 1997)
--**Increased resources** devoted to resolving phonetic category membership through fine-grained analysis of the stimulus.

Activation in the **cingulate** may reflect**:**

--**Competition** between responses (in this case, the 'd' button or the 't' button).

**The left middle temporal gyrus and two midline areas were responsive to the tokens that correspond to the 'best fit' to the phonetic category.**

Deactivation in the left MTG and angular gyrus may reflect:

--**Mapping** to the representation of the phonetic category, with the least deactivation for the best fit to that category.

--**Naming** of the stimulus, with the least deactivation for the stimulus that is the easiest to label.

## References

**Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000) Human temporal lobe activation by speech and nonspeech sounds.** *Cerebral Cortex*, **10, 512-528.**

**Burton, M. W., Small, S. L., & Blumstein, S. E. (2000) The role of segmentation in phonological processing: An fMRI investigation.** *Journal of Cognitive Neuroscience*, **12, 679-690.**

**Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997) Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation.** *Proceedings of the National Academy of Sciences of the United States of America*, **94, 14792-14797.**

**Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992) Lateralization of phonetic and pitch discrimination in speech processing.** *Science*, **256, 846-849.**

## Acknowledgments

The authors would like to thank Jesse Rissman and Emmette Hutchison for help with the design and execution of the study. This work was supported by NIH grant 1 F31 DC006520-01 to EBM and a Dana Foundation Grant to SEB.