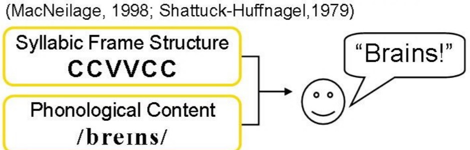
Dissociated neural representations of phonological content and syllabic frame structure

Jennifer A. Segawa[1*], Jason A. Tourville[2], Deryk S. Beal[3], & Frank H. Guenther[1,2,4]

Background

- The **frame/content** theory and the **slot/filler** theory both posit that:
  - Syllabic frame structure and phonological content of a speech utterance are represented **in parallel** (MacNeilage, 1998; Shattuck-Huffnagel,1979)



- These theories have been implemented in several influential models of speech (GODIVA, WEAVER++, Coupled Oscillator)
- **However, a neural basis of the syllabic frame is largely unexplored.**
- Previous work (Peeva et al., 2010) showed a syllabic representation in the ventral premotor cortex, but did not differentiate between representations of a syllabic frame and a full syllable (with phonological content)

[1] Graduate Program for Neuroscience, Boston University
[2] Dept. of Speech, Language, & Hearing Sciences, Boston University
[3] Institute for Stuttering Treatment & Research and Dept. of Speech Pathology and Audiology, University of Alberta
[4] Dept. of Biomedical Engineering, Boston University
* Corresponding author email: jsegawa@bu.edu

Buckner, R.L. et al., (2011) *J. Neurophysiol.* 105, 5, 2322-45.
Diedrichsen, et al. (2009). *NeuroImage.* 46, 1, 39-46.
Fischl, B., et al. (2002). *Cereb. Cortex.* 14, 1, 11-22.
Giraud, A.-L. et al. (2007). *Neuron.* 56, 1127-34.
Guenther, F.H., et al. (2006). *Brain and Language.* 96, 280-301.
Levelt, W.J. & Wheeldon, L. (1994). *Cognition.* 50, 239-269.
MacNeilage, P.F. (1998). *Beh. Brain Sci.* 21, 499-511.
O'Reilly, J.X., et al. (2010) *Cereb. Cortex.* 20 (4): 953-965.
Peeva, M., et al. (2010). *NeuroImage.* 50, 628-638.
Poeppel, D. (2003). *Speech Comm.* 41, 245-255.
Shattuck-Huffnagel, S. (1979). In *Sentence Processing: Psycholinguistic studies presented to Merrill Garrett.* Hillsdale, NJ: Erlbaum.
Tourville, J.A. & Guenther, F.H. (2003). *Boston University Technical Report CAS/CNS-03-022.* Boston, MA: Boston University

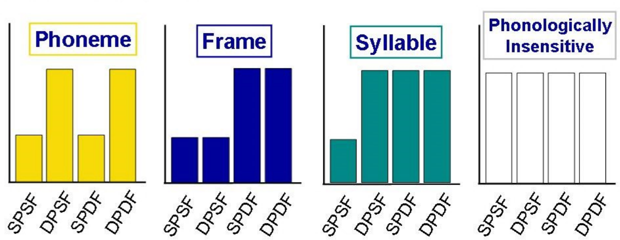
Methods

- *Subjects*: 17 American English speaking subjects
- *Stimuli & Paradigm*
  - **FMRI repetition suppression paradigm**
    - BOLD response decreases across repeated presentations of stimulus in a region that processes that stimulus
  - **4 conditions** (fMRI block: 3 repetitions of a pair of utterances)

	SPSF	DPSF	SPDF	DFDF
	Same phonemes **Same** frames	**Different** phonemes **Same** frames	**Same** phonemes **Different** frames	**Different** phonemes **Different** frames
time ↓	TWAI	FAS	RAUD	DEEF
	TWAI	REEN	DRAU	GLAI
	TWAI	FAS	RAUD	DEEF
	TWAI	REEN	DRAU	GLAI
	TWAI	FAS	RAUD	DEEF
	TWAI	REEN	DRAU	GLAI

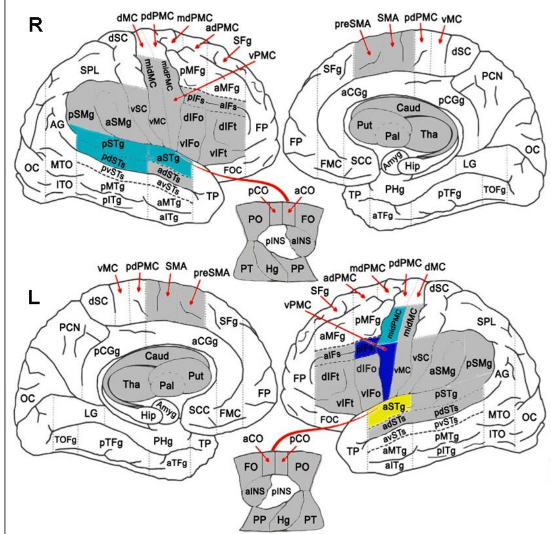
  - Timecourse of a block
    1. Auditory & orthographic targets (0.7 & 1.5 s) } 6 times
    2. Subject produces sequence (0.5 s) } 6 times
    3. Pause (3 s)
- *fMRI acquisition*
  - Siemens Trio Tim 3T, 32 channel head coil
  - TR: 2.5 s, 41 slices, Skip: 25%, 200mm coverage
  - Voxel size: 3.1 x 3.1 x 3.0 mm
- *fMRI analysis*
  - **Nipype** (http://nipy.sourceforge.net/nipype) provides an interface between neuroimaging software tools
  - Functional volumes realigned to subject's anatomical volume, corrected for slice acquisition timing, and first level model estimated with **SPM8**.
  - ROI parcellation: cortical (surface-based speech-focused, Tourville & Guenther, 2003), subcortical (Fischl et al., 2002), and cerebellar (Diedrichsen, 2001)
  - Contrast values – 2nd *and* 3rd *pair vs. baseline (silent fixation)* for each speaking condition – extracted and averaged across each ROI using **REX** (http://web.mit.edu/swg/rex)
  - Each contrast value normalized by average activity across ROIs
- *ROI pattern matching* (Peeva et al., 2010)
  - *A priori* models defined by BOLD activity comparisons across conditions
  - Model fit quantified by conjunction test comparing the 4 speaking conditions based on 4 predicted patterns of repetition suppression at each speech production ROI
  - Significance threshold of $P_{FDR} < 0.05$
    - Where P is derived from distribution of values from Monte Carlo simulation with 10,000 trials and using the noise means and variance of the data

4 predicted patterns of repetition suppression across conditions



This work is supported by NIH grant R01 DC007683.

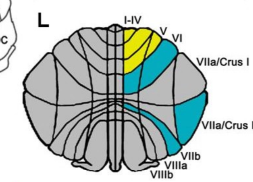
Repetition Suppression Pattern Matches

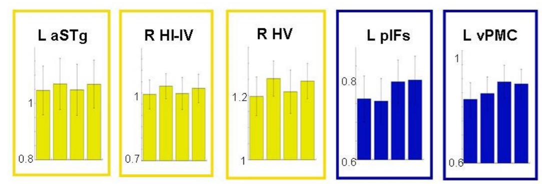


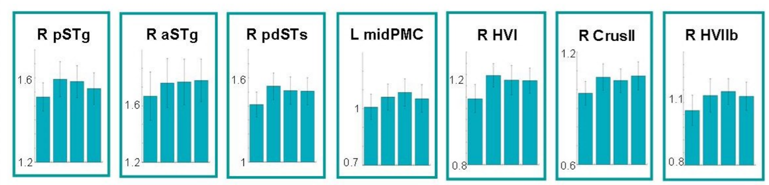
Phoneme:
- Left anterior superior temporal gyrus (**aSTg**)
- Right cerebellum, Lobules I-IV
- Right cerebellum, Lobule V

Frame:
- Left posterior inferior frontal gyrus (**L pIFg**)
- Left ventral premotor cortex (**vPMC**)

Syllable:
- Right posterior superior temporal gyrus (**pSTg**)
- Right anterior superior temporal gyrus (**aSTg**)
- Right posterior dorsal superior temporal sulcus (**pdSTs**)
- Left mid-premotor cortex (**midPMC**)
- Right cerebellum, Lobule VI
- Right cerebellum, Crus II
- Right cerebellum, Lobule VIIb







Discussion

- **Superior temporal cortex**
  - Phoneme representation in left hemisphere
  - Syllable representation in right hemisphere
  - Agrees with studies of endogenous oscillatory activity in auditory cortex (Giraud et al., 2007; Poeppel, 2003)
    - Left hemisphere: low-gamma (25-45Hz) → **Phonemic signal sampling**
    - Right hemisphere: theta (4-8Hz) → **Syllabic signal sampling**
- **Left premotor cortex**
  - Syllable representation in mid-premotor cortex
    - Stores **feedforward motor commands for speech**
    - E.g. Speech sound map (Guenther, 2006), mental syllabary (Levelt & Wheeldon, 1994)
  - Frame representation in ventral premotor/posterior inferior frontal sulcus
    - Reflects **syllabic frame representation** without phonological content
- **Right cerebellar hemisphere**
  - **Fine-tuning of cerebral sensori-motor processing**
  - Functional/anatomical connections (Buckner, et al., 2011; O'Reilly, et al., 2010)
  - Anterior: phoneme representation (L superior temporal cortex ↔ R anterior cerebellum)
  - Posterior: syllable representation (L prefrontal/premotor ↔ R posterior cerebellum)