

Psychophysical and Physiological Studies of Synthetic Vowel Harmonic Structure



Julian Jenkins III^{1,2}, Jonathan Z. Simon^{1,3}, David Poeppel⁴, William J. Idsardi^{2,5}

¹Department of Biology, University of Maryland, College Park; ²Cognitive Neuroscience of Language Laboratory, University of Maryland, College Park;

³Department of Electrical and Computer Engineering, University of Maryland, College Park; ⁴Department of Psychology, NYU; ⁵Department of Linguistics, University of Maryland, College Park

MOTIVATIONS

• **Timbre** of vowel tokens as the means to probe category formation [1,2].

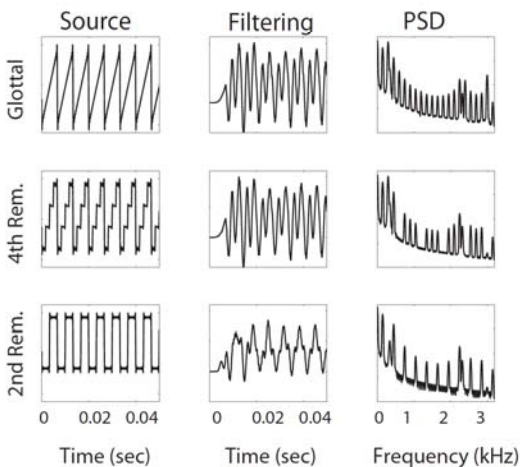
• We can control the physical structure of synthesized vowels and begin to test which aspects contribute to perceptual classification [3,4,5].

• Psychophysics and MEG used to probe response characteristics [6,7,8,9,10,11].

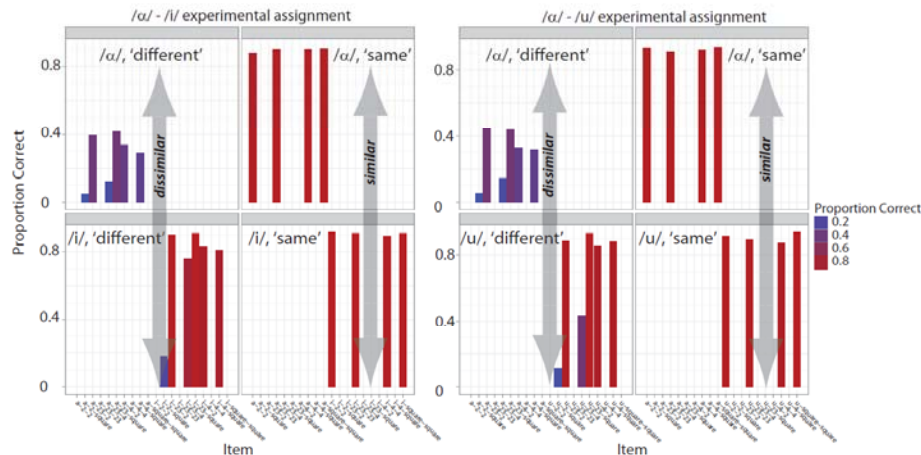
• Null hypothesis: vowels should be, on balance, equally discriminable.

SYNTHETIC VOWELS

1. Ecological validity: source waveforms based on glottal excitation pattern (sawtooth wave comprised of 23 harmonics; F0 = 150 Hz) via Fourier Synthesis
2. Filters (transfer functions) taken from formant values (F1,F2,F3) of male vocal tract measurements - /a/, /i/, /u/
3. Create a timbral continuum by removing harmonics (every 6th, 5th, 4th, 3rd, 2nd)



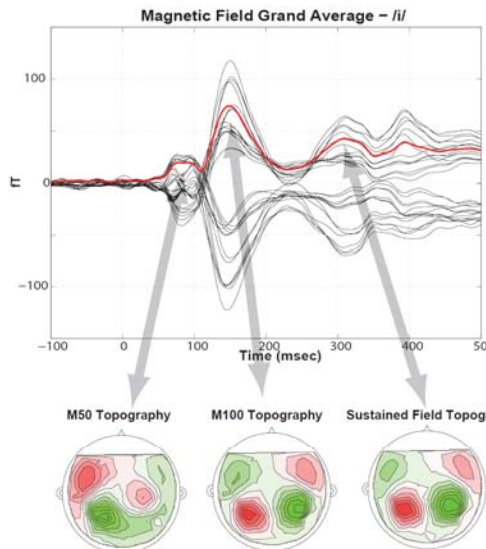
PSYCHOPHYSICAL RESULTS



1. **Dramatic asymmetry in vowel proportion correct - /a/ signals correctly discriminated much less than either /i/ or /u/**

2. **Asymmetry arises from signals in middle of the continuum**

MEG RESULTS



1. Three major components in the evoked response (M50, M100, sustained field)
2. **M50 very left-lateralized**
3. **M100 latency and peak RMS depend on F0 and transfer function structure**
4. **At least in preattentive state, evoked response not adequate to examine within-category timbral differences**
5. **/a/ - /i/:** no statistically different peak RMS and latency (M100) **/a/ - /u/:** /a/ signals exhibited faster latency, larger amplitudes (M100)

PSYCHOPHYSICS

1. AX discrimination task; participants randomly placed in one of two experimental pools
2. Subset of continuum signals – glottal excitation pattern, 4th removed, 2nd removed, square sources
3. Analyses: Proportion Correct (Vowel, Vowel Height, Vowel Position, Trial); GLMs: proportion correct (factors: Vowel, Vowel Height, Vowel Position, Trial)

PHYSIOLOGY: MEG DESIGN

1. Passive presentation; participants randomly placed in one of two experimental pools
2. Calculated Peak RMS and latency (M50, M100, P2m)
3. Analyses: Peak RMS and latency (M50 – vowel bandwidth; M100 – F0, transfer function structure; P2m – harmonics contained); Bootstrapped RMS: amplitude differences between condition; grand averages: response consistency

DISCUSSION

- M100 – the major auditory evoked component – latency and peak RMS determined by F0 and vowel structure [9,12]
- Evoked responses within vowel categories consistent

• Why is /a/ treated so differently?

Hypothesis 1: 'Averaging' of the first two formants – smears the source spectrum leading to discrimination impairment

Hypothesis 2: signal dynamics (temporal evolution) differs based on transfer function; some filters produce more self-similar waveforms [13]

Timbral discrimination of vowels is based on how the filter (transfer function structure) affects overall signal dynamics; salient (?) dimensions reflected in evoked response preattentively

REFERENCES

1. Binder et al. 2004. Nat Neurosci. 7(3): 295-301
2. Zatorre et al. 2004. J Neurosci. 24(14): 3637-42
3. Klein et al. 1970. J Acoust Soc Am. 48(4): 999-1009
4. Pols et al. 1969. J Acoust Soc Am. 46(2): 458-67
5. Caclin et al. 2005. J Acoust Soc Am. 118(1): 471-82
6. Slawson et al. 1968. J Acoust Soc Am. 43(1): 87-101
7. Vihla et al. 2000. PNAS. 97(19): 10590-4
8. Vihla et al. 2003. Brain Res Cogn Brain Res. 16(2): 167-73
9. Poeppel et al. 1997. Neurosci Lett. 221(2-3): 145-8
10. Eulitz et al. 1995. J Neurosci. 15(4): 2748-55
11. Diesch and Luce 1997. Psychophysiology. 34(5): 501-10
12. Roberts et al. 2000. J Clin Neurophysiol. 17(2): 114-29
13. Baizano 1986. Music Perception. 3(3): 297-314

