Effects of auditory spatial attention in a semantic classification task

Daniel McCloy

Adrian KC Lee

Department of Speech & Hearing Sciences
Institute for Learning & Brain Sciences
University of Washington

[LABS] Laboratory for Auditory Brain Sciences & Neuroengineering

Institute for Learning & Brain Sciences

5aPP19

BACKGROUND

- Dichotic priming experiments show awareness of unattended speech streams under certain conditions. [1]
- After switching attention to a new stream, phonological information available early (~600 ms), semantic information available late (~1100 ms). [2]
- When attention is divided between spatial streams, it is unclear whether listeners use rapid switching or broadened / replicated "auditory spotlights." [3]

QUESTION

- How does spatially divided attention impact awareness of unattended streams?

HYPOTHESIS

- In a spatial divided attention task, defining targets semantically will *reduce* false alarms to unattended foils.

STIMULI

- New recordings of monosyllabic words from 7 semantic categories (6 words per category) plus several control words (see Table 1).
- Categories were similar in lexical frequency ($F(6,35)=1.188$, $p=0.33$) and neighborhood density ($F(6,35)=1.121$, $p=0.37$).
- Words monotonized to talker's mean f_0 (100 Hz) [4].
- Words RMS normalized and concatenated into 4 temporally interleaved 12-word streams per trial. Across-stream ISI: 250 ms, within-stream ISI: 750–1750 ms (see Figure 1).
- Streams were spatialized at ±15° and ±60° azimuth by convolving with anechoic HRTFs from [5].

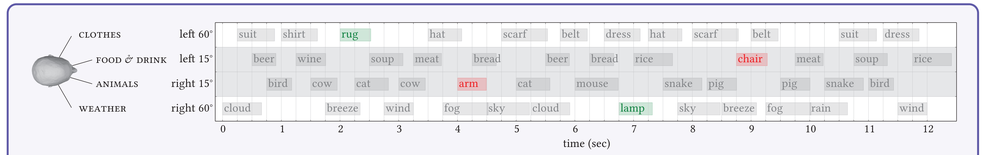
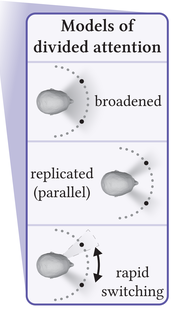


Figure 1. Sample trial structure for the semantic condition trial. To-be-attended streams have white backgrounds, ignored streams have grey backgrounds. Small rectangles are actual word durations; targets are green, and foils are red.



PROCEDURE

- Prior to experiment, all participants did a categorization task with all test words to ensure consistent categorization.
- Six participants did test (*semantic*) condition first; seven did control (*phonetic*) condition first.
- 5-step training in each phase ensured mastery of all aspects of the task (attend single streams, attend multiple streams, ignore streams, *etc*).
- On each trial, 1 stream (*selective attention*) or 2 streams (*divided attention*) marked on screen as "to be attended." Listeners responded by button press to targets in the to-be-attended stream(s).
- Four blocks per phase; within blocks, the relationship between category and spatial location was held fixed.

TABLE 1: SEMANTIC CATEGORIES & CONTROL WORDS

ANIMALS	bird	cat	cow	mouse	pig	snake
BODY	arm	chin	foot	leg	mouth	nose
CLOTHES	belt	dress	hat	scarf	shirt	suit
FOOD & DRINK	beer	bread	meat	rice	soup	wine
FURNITURE	bed	chair	couch	desk	lamp	rug
PLANTS	bark	grass	leaf	root	stick	tree
WEATHER	breeze	cloud	fog	rain	sky	wind
CONTROL WORDS	branch	cake	duck	fish	fruit	hail
	knee	moss	pants	sink	stove	wrist

PHONETIC CONDITION (CONTROL)

- Each stream comprised twelve repetitions of a different monosyllabic control word.
- Targets/foils were any word not matching the base word.

SEMANTIC CONDITION

- Each stream comprised all six words from a semantic category (see Table 1), each played twice, in random order.
- Each stream was a different semantic category.
- Targets were any word not matching the category of the stream.

RESULTS

REACTION TIME

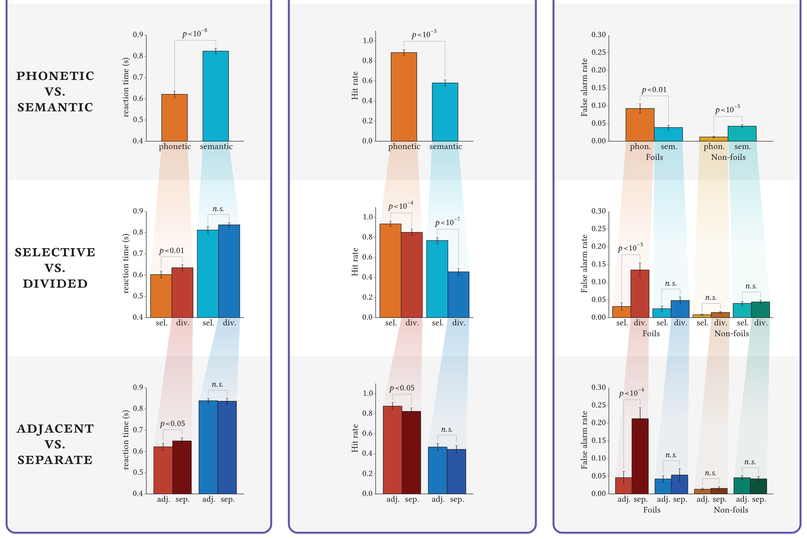
- Shorter RT in phonetic condition
- In phonetic condition, shorter RT in selective attention trials
- In phonetic divided attention cond., shorter RT in spatially adjacent trials

RESPONSE ACCURACY

- Greater accuracy in phonetic condition
- Greater accuracy in selective attention trials (esp. in semantic cond.)
- In phonetic divided attention cond., slightly greater accuracy in spatially adjacent trials

FALSE ALARM RATE

- Strong reaction to foils in phonetic trials with spatially separated attended streams
- Reaction to foils minimal in all other conditions
- Reaction to non-foils higher in semantic condition



DISCUSSION

- Reduced reaction to foils in spatially separated semantic condition supports hypothesis
- False alarms to phonetic foils are probably exogenous reorientations
- Results favor either rapid switching or multiple spotlight model of divided attention
- Listeners may use different strategies in phonetic / semantic conditions

FUTURE WORK

- Manipulate temporal ordering of streams
- Manipulate difficulty of tasks to vary cognitive load, assess with pupillometry
- Neuroimaging of semantic / phonetic task to identify level of interaction between attentional & linguistic networks

ACKNOWLEDGMENTS

This research was supported by NIH grants R01-DC013260 to Adrian KC Lee, and T32-DC000033 to the Department of Speech & Hearing Sciences, University of Washington. Special thanks to Eric Larson, Ross Maddox, Lindsey Kishline, and other members of the Laboratory for Auditory Brain Sciences & Neuroengineering for helpful discussions and feedback.

[1] Rivenez, M, Darwin, C, & Guillaume, A (2006). Processing unattended speech. *J Acoust Soc Am* 119(6), 4027–4040. DOI:10.1121/1.2190162

[2] Hafter, E, Xia, J, Kalluri, S, Poggesi, R, Hansen, C, & Whiteford, K (2013). Attentional switching when listeners respond to semantic meaning expressed by multiple talkers. *POMA* 19, 050077. DOI:10.1121/1.4801413

[3] Best, V, Gallun, F, Ihlefeld, A, & Shinn-Cunningham, B (2006). The influence of spatial separation on divided listening. *J Acoust Soc Am* 120(3), 1506–1516. DOI:10.1121/1.2234849

[4] Boersma, P, & Weenink, D (2013). Praat: Doing phonetics by computer (v5.3.41). http://www.praat.org/

[5] Shinn-Cunningham, B, Kopco, N, & Martin, T (2005). Localizing nearby sound sources in a classroom: Binaural room impulse responses. *J Acoust Soc Am* 117(5), 3100–3115. DOI:10.1121/1.1872572