Attention and salience in lexically-guided perceptual learning

Michael McAuliffe

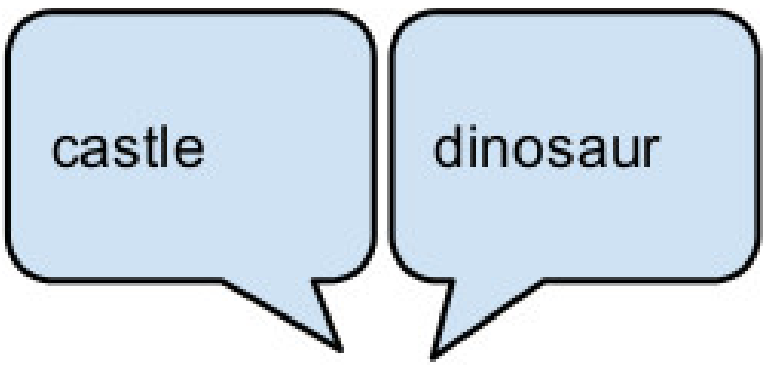
PhD Defense

Perceptual constancy

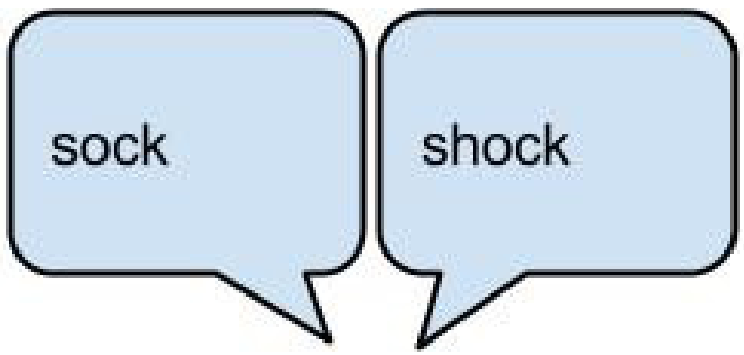
Despite variation, listeners can interpret variable productions as a single word type

Shankweiler et al. (1977); Kuhl (1979); Sumner and Kataoka (2013)

Perceptual learning

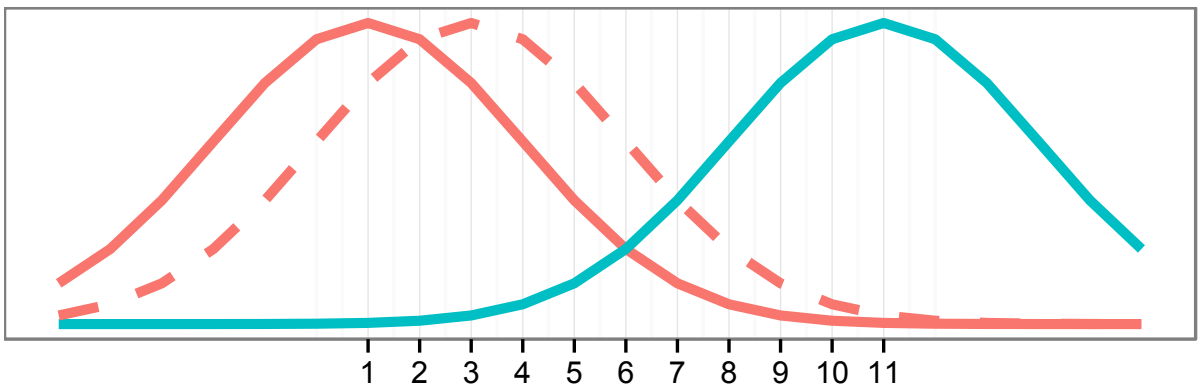


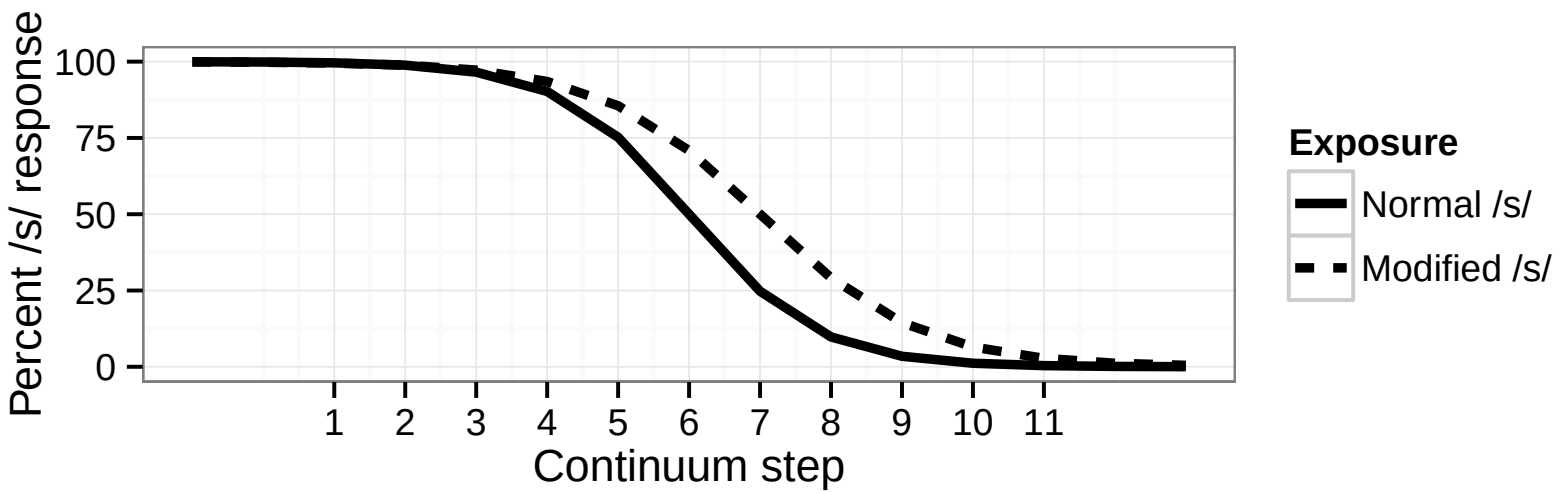
Exposure



Categorization

Categorization





Sources of variation

Example: /s/

SPEAKER

- Indexical
  - Accent
  - Gender

LISTENER

- Indexical
  - Accent
  - Perceived accent
  - Perceived gender

Strand and Johnson (1996); Li et al. (2011); Kraljic and Samuel (2005)

Sources of variation

Example: /s/

SPEAKER

- Contextual
  - Speaking rate
  - Coarticulation (/stɹ/)
  - Word position
  - Predictability

LISTENER

- Contextual
  - Speaking rate
  - Coarticulation (/stɹ/)
  - Word position
  - Predictability

Lieberman (1963); Kraljic et al. (2008); Clopper and Pierrehumbert (2008); Pitt and Szostak (2012)

Sources of variation

Example: /s/

SPEAKER

- Attention

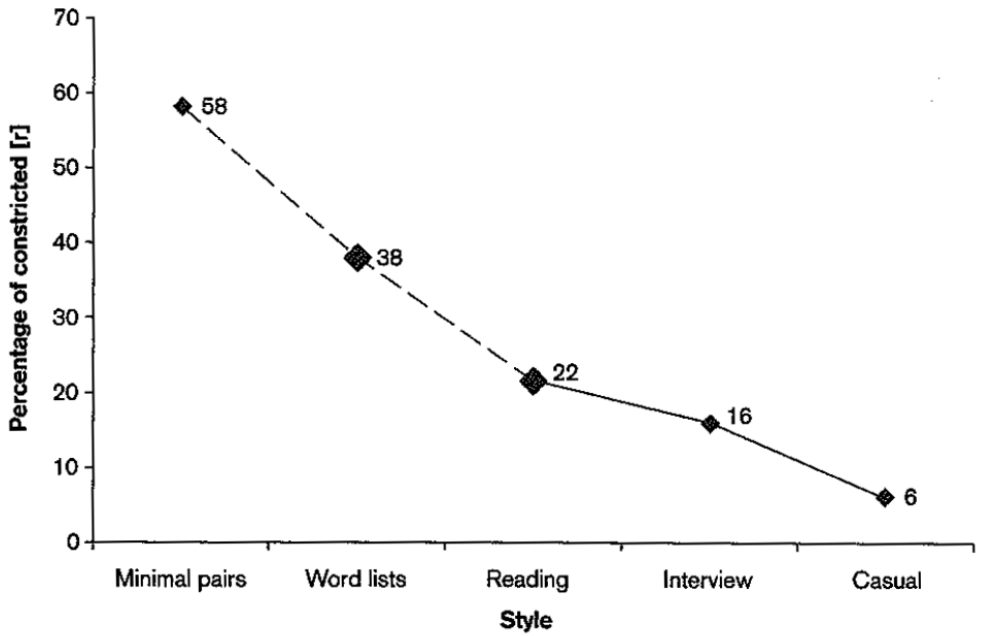


Figure 3.2 Occurrence of constricted [r] in New York City English in five speech styles. (Dashed line indicates the qualitatively different activities involving the use of unconnected speech.) (Source, Labov 1966: 221.)

LISTENER

- Attention
  - Comprehension
  - Perception

Labov (1997); Pitt and Szostak (2012)

Real world example

- Students taught by a professor with non-native accent
  - Some students will try to comprehend the lecture
  - Some students will get distracted by pronunciation
  - Some students will daydream
  - Do they all perceptually learn?

Research question

How do changes to a listener's attention in exposure affect perceptual learning in future categorization?

Interaction of variation

Example: /s/

Speaker

- Contextual
  - Speaking rate
  - Coarticulation (/stɹ/)
  - **Word position**
  - **Predictability**
- Attention

Salience

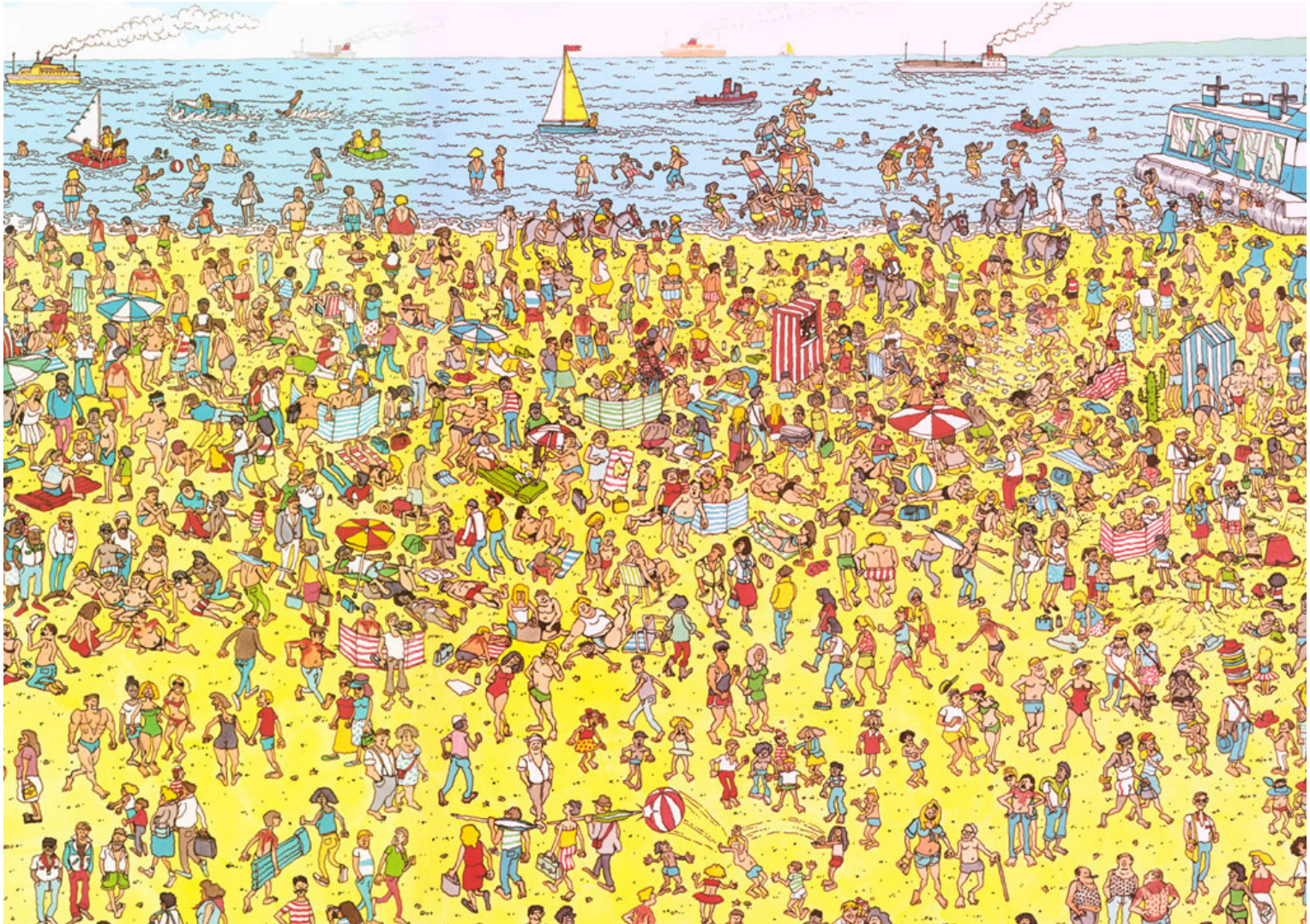
Listener

- Contextual
  - Speaking rate
  - Coarticulation (/stɹ/)
  - Word position
  - Predictability
- **Attention**
  - Comprehension
  - Perception

Outline

Attentional sets

Strategies to parse our perceptual experience



Attentional sets

Comprehension-oriented

- Focus on comprehending meaning
- Real world example:
  - Students in lecture
  - Primary focus is comprehending the professor (we hope)

Pitt and Szostak (2012)

Attentional sets

Perception-oriented

- Focus on perceiving a specific pronunciation
- Real world example:
  - Students in lecture
  - Professor with a non-native accent
  - Primary focus might shift from comprehension

Pitt and Szostak (2012)

Attentional sets in perceptual learning

- Comprehension-oriented tasks
  - Lexical decision
  - Sentence transcription
- Perception-oriented tasks
  - Psychophysical perceptual learning
  - Audio-visual lipreading (nonwords)

Ahissar and Hochstein (1993); Norris et al. (2003); Vroomen et al. (2007); Bradlow and Bent (2008); Reinisch et al. (2014)

Generalization in perceptual learning

Comprehension-oriented tasks generalize

- New words or nonwords
- (Sometimes) new voices

Perception-oriented tasks do not generalize as readily

- Exposure specificity

Ahissar and Hochstein (1993); Norris et al. (2003); Kraljic and Samuel (2005); Bradlow and Bent (2008); Pitt and Szostak (2012); Reinisch et al. (2013)

Hypothesis

Comprehension-oriented attentional sets allow for greater generalization than perception-oriented attentional sets.

Experimental paradigm

Comprehension-oriented tasks

- Lexical decision
- Word identification in sentences

Manipulations to promote perception-oriented attentional sets

- Instructions
- Salience
  - Unpredictability or low expectations
  - Increase the likelihood of listeners noticing modification
  - Assumption: similar to increasing the number of /s/ trials relative to filler trials

Explicit instructions

- “This speaker’s ‘s’ sounds are ambiguous”
- “Please listen carefully to ensure you make the correct choice”
- Promote perception-oriented attentional set

Pitt and Szostak (2012)

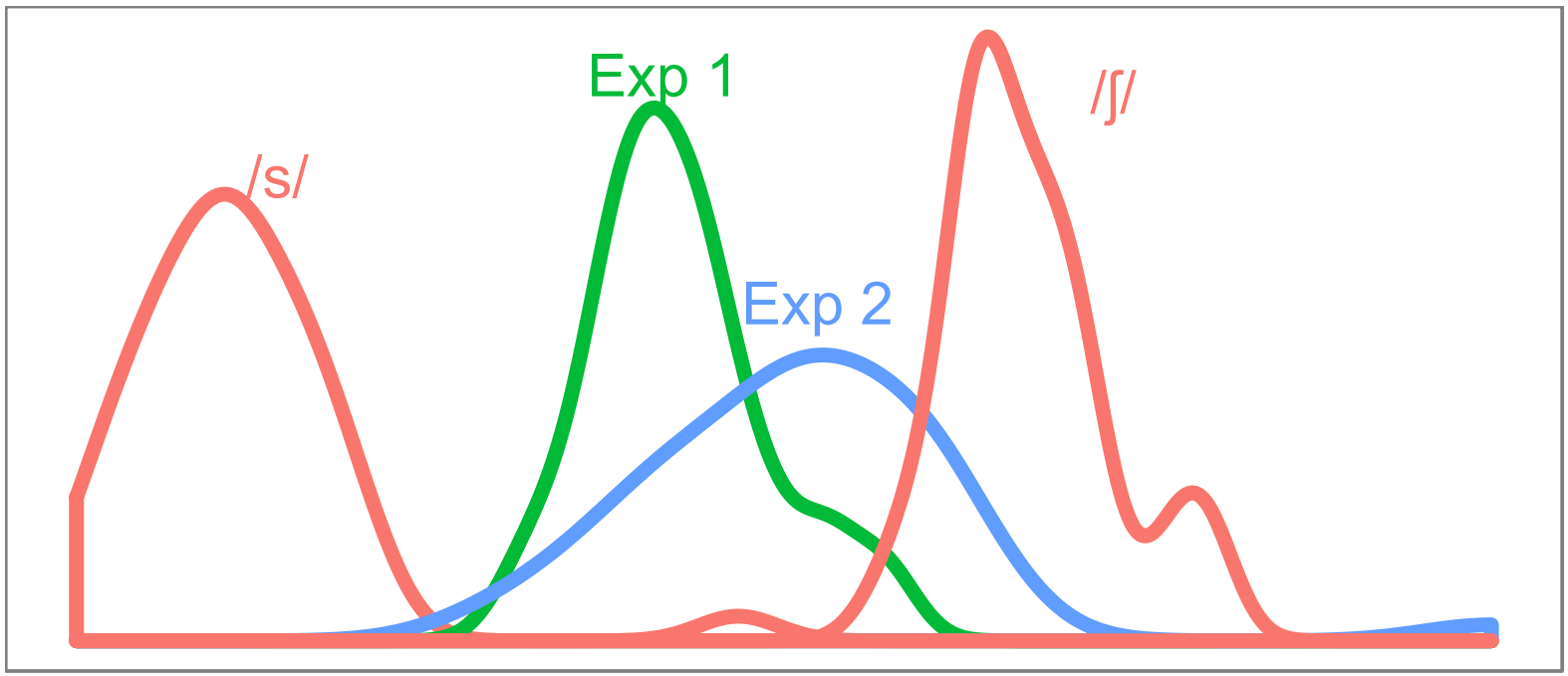
Salience - Word position

- Listeners are more tolerant of variation later in the word
- Word-initial modified /s/ should be more salient
- Examples
  - Word-initial: *submarine*
  - Word-medial: *whistle*

Pitt and Szostak (2012)

Salience - Category typicality

- Productions farther from the mean of a category are more salient



Experiments 1 and 2

Experiment 1

- Lexical decision exposure task
- 94 native English participants
- Across subject factors
  - Instructions
  - Position of modified /s/ in words (Word-initial vs word-medial)
- 50% word response rate in a pre-test (n = 20)

Experiment 2

- Same design and materials as Experiment 1
- 96 native English participants
- 30% word reponse rate in the pre-test (more atypical /s/)

Sample trials

Exposure

- Hear: whistle (Experiment 1) (Experiment 2)
- Hear: submarine (Experiment 1) (Experiment 2)
- Word or nonword?

Categorization

- Hear: sock-shock (continuum), sin-shin, sack-shack, sigh-shy
- Sock or shock? Sin or shin? etc.

Experiment 1 and 2 predictions

Possible categorization outcomes:

- Equal perceptual learning effects across all conditions
- Less perceptual learning when perception-oriented attentional sets are promoted
  - Primary hypothesis
- Perceptual learning effects stronger in Word-initial exposure condition
  - More similar to categorization items

Experiment 1 - Word-initial exposure

Ambiguous /s/

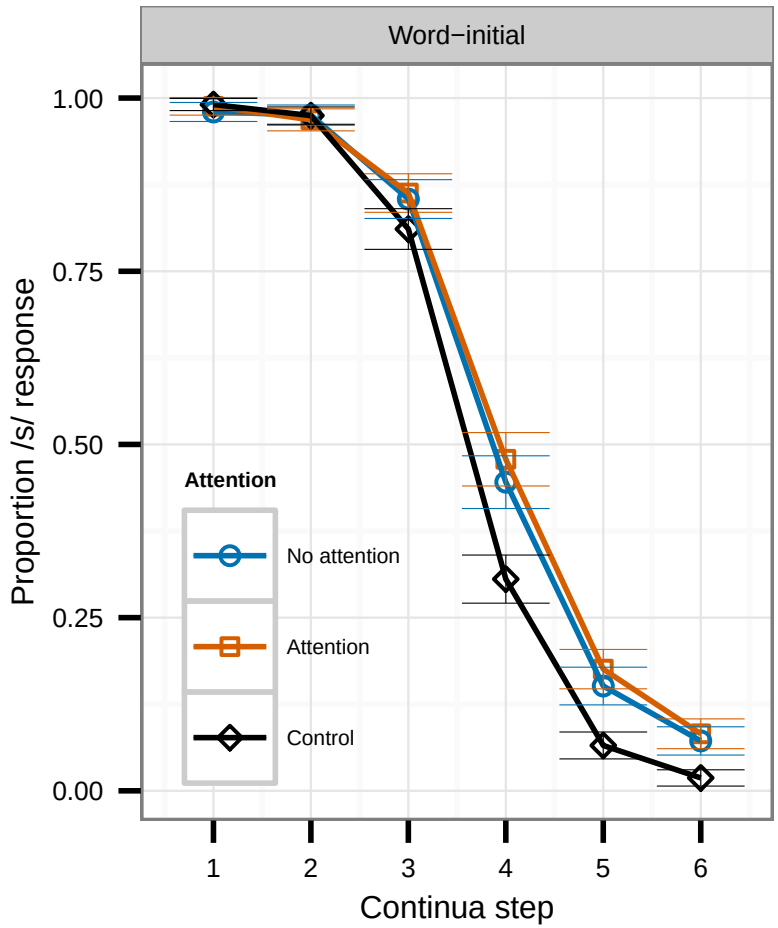
- 50% between /s/ and /ʃ/

Attention

- No /s/-oriented instructions
- Told /s/ would be ambiguous

Position of /s/

- *Word initial*
- Word medial



Experiment 1 - Word-medial exposure

Ambiguous /s/

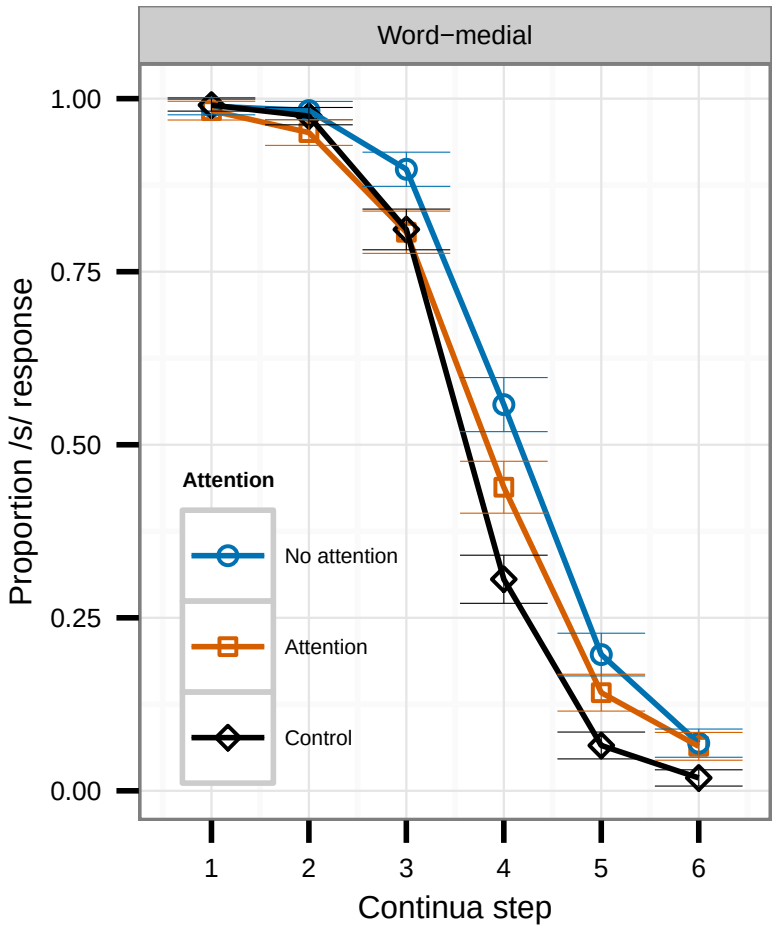
- 50% between /s/ and /ʃ/

Attention

- No /s/-oriented instructions
- Told /s/ would be ambiguous

Position of /s/

- Word initial
- *Word medial*



Experiment 2 - Word-initial exposure

Ambiguous /s/

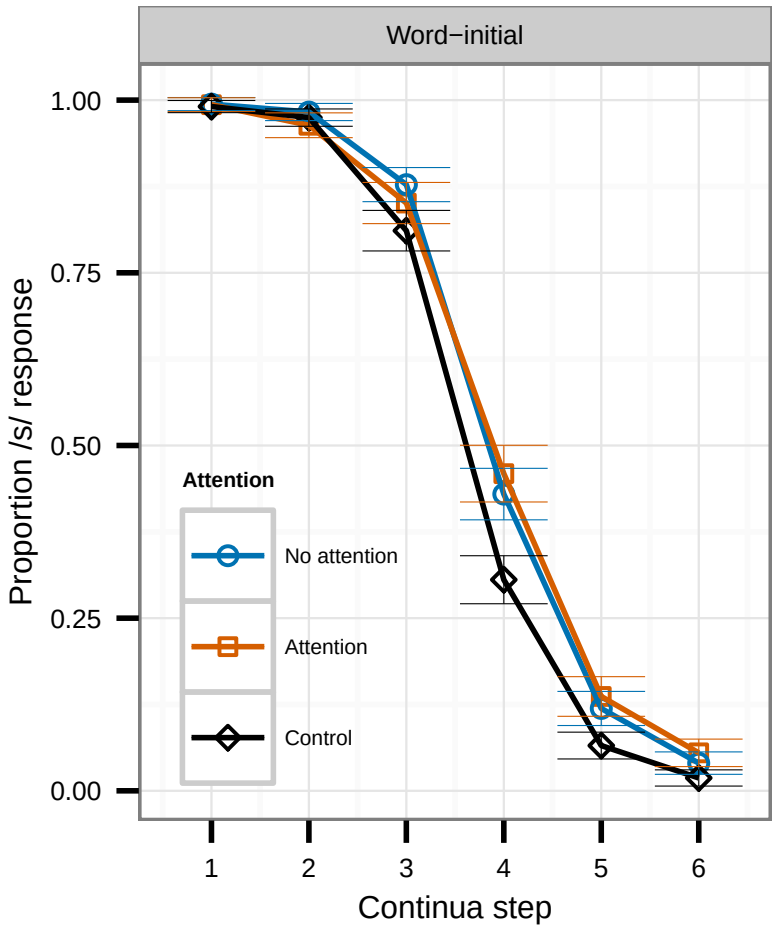
- More like /ʃ/ than /s/

Attention

- No /s/-oriented instructions
- Told /s/ would be ambiguous

Position of /s/

- *Word initial*
- Word medial



Experiment 2 - Word-medial exposure

Ambiguous /s/

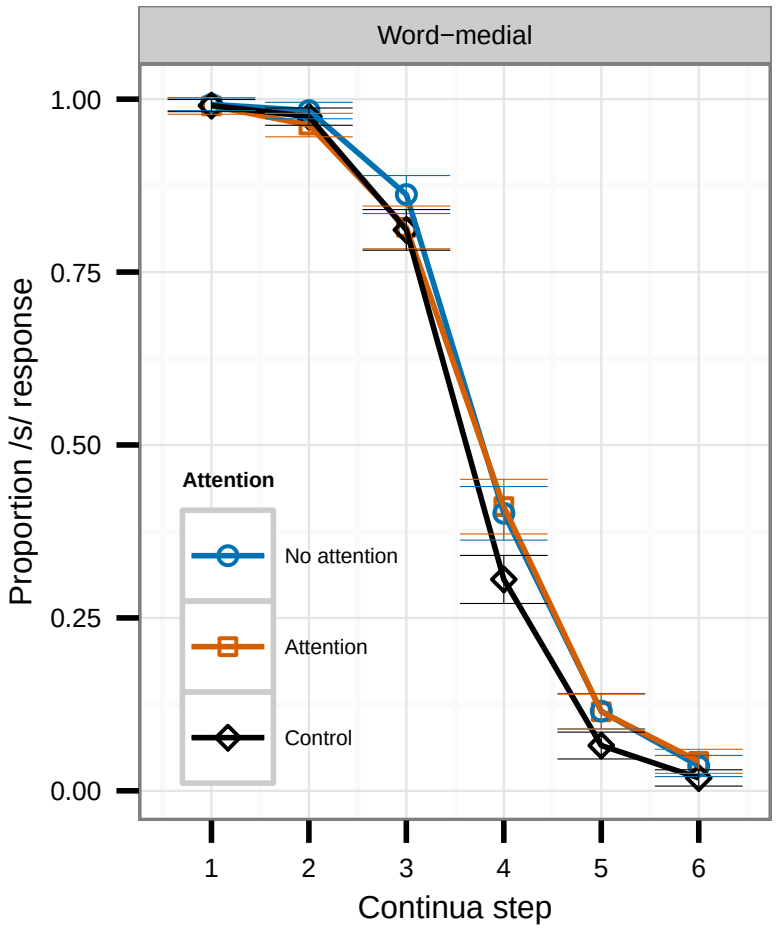
- More like /ʃ/ than /s/

Attention

- No /s/-oriented instructions
- Told /s/ would be ambiguous

Position of /s/

- Word initial
- *Word medial*



Summary

- Results align with attentional sets
- Conditions promoting a perception-oriented attentional set
  - Had smaller perceptual learning effects
  - Still showed perceptual learning
  - Did not differ from one another
- Fine-grained similarity did not appear to play a role
  - Word-medial exposure had the largest effect

Further promoting comprehension

- Lexical decision is comprehension-oriented
  - Word recognition
- Experiment 3 uses words in sentences
  - Attempt to further promote comprehension

Exposure task

- Predictable: The traffic cop alerted the driver by blowing her whistle (Audio)
- Unpredictable: The boy ran away when he heard the whistle(Audio)

Salience - Semantic predictability

- Listeners are more tolerant of acoustic reduction or noise in predictable sentences
- Modified /s/ should be more salient in unpredictable words
- Examples
  - Predictable: The cow gave birth to the calf.
  - Unpredictable: She is glad Jane called about the calf.

Lieberman (1963); Kalikow et al. (1977); Scarborough (2010)

Experiment 3

- 98 native English participants
- Cross-modal word identification
  - Auditory sentences
  - Identification of picture corresponding to final word in sentence
  - Same word-medial modified /s/ stimuli (Experiment 1)
  - Final targets were predictable or unpredictable (pre-test)
- Across subjects
  - Instructions (identical to Experiments 1 and 2)
  - Modified /s/ only in predictable or unpredictable words
- **Same categorization as Experiment 1 and 2**

Experiment 3 categorization predictions

Within sentences:

- Equal perceptual learning
- Bigger perceptual learning effect in predictable exposure
  - Less salient modification
- Smaller perceptual learning effect in predictable exposure
  - Attribution of variation to predictability

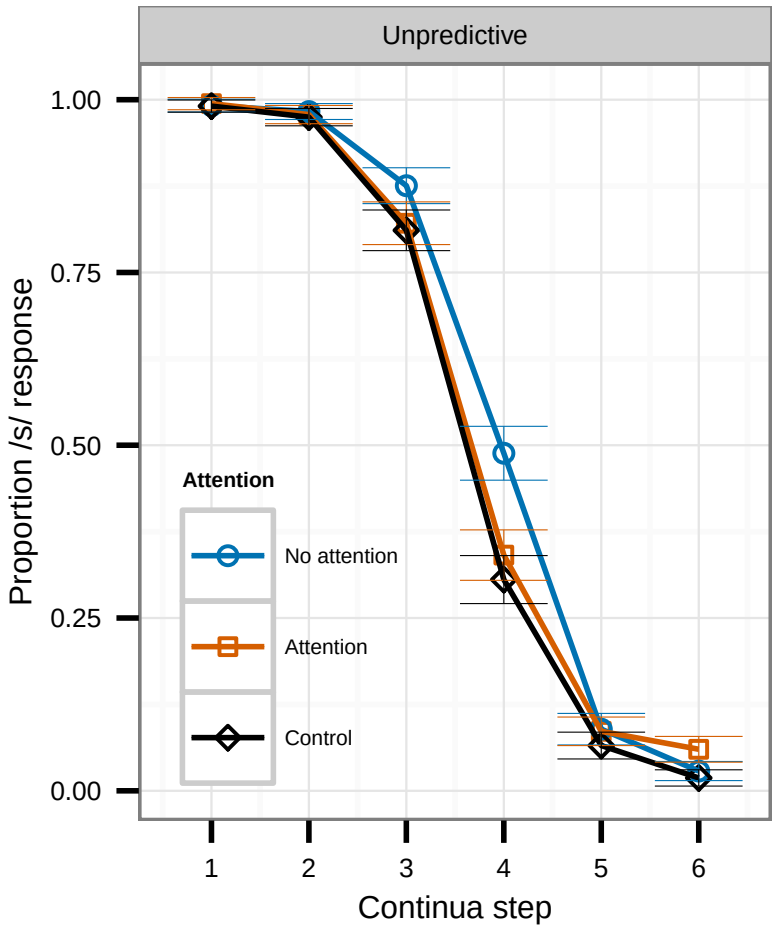
Experiment 3 categorization predictions

Words in isolation vs in sentences

- Equal perceptual learning
- Bigger perceptual learning effect in sentences
  - Comprehension of sentences rather than words
- Smaller perceptual learning effect in sentences
  - Less speaker attention in read sentences

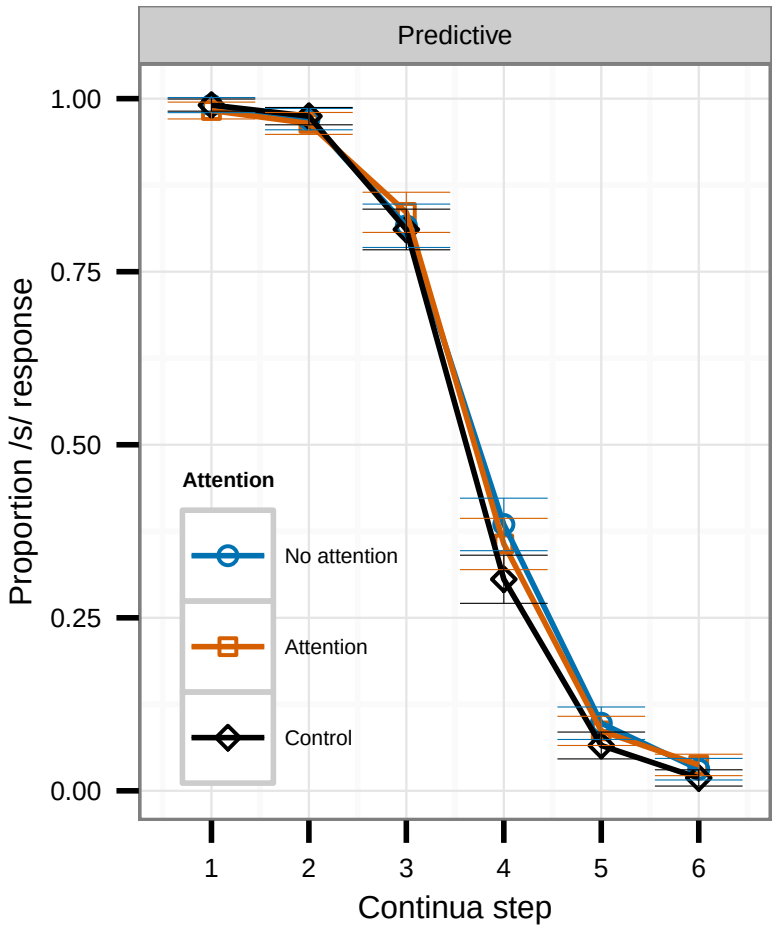
Experiment 3 - Unpredictable exposure

- **Ambiguous /s/**
  - Halfway between /s/ and /ʃ/
  - In sentences
- **Attention**
  - No /s/-oriented instructions
  - Told /s/ would be ambiguous
- **Predictability of final /s/ words**
  - *Unpredictable*
  - Predictable

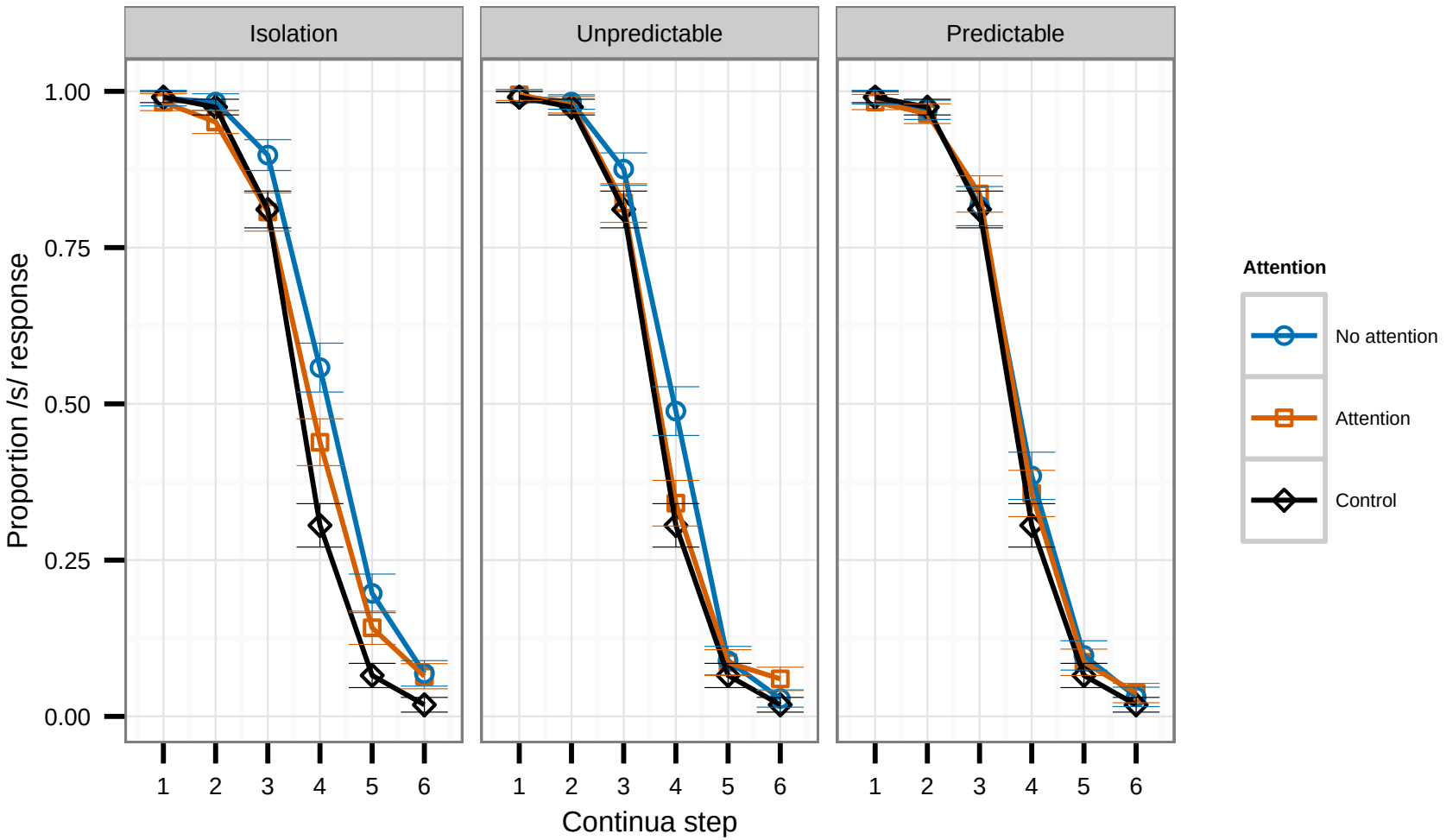


Experiment 3 - Predictable exposure

- **Ambiguous /s/**
  - Halfway between /s/ and /ʃ/
  - In sentences
- **Attention**
  - No /s/-oriented instructions
  - Told /s/ would be ambiguous
- **Predictability of final /s/ words**
  - Unpredictable
  - *Predictable*



Isolation vs Sentences



Summary

- Unpredictable exposure showed a similar pattern to words in isolation
- Predictable exposure showed no perceptual learning effect
  - Similar to studies using a coarticulation context (/stɹ/)
  - Predictable words are typically shorter and less clear
  - Listeners compensate for this predictability
  - Mean durations: Predictable words (0.53 s, SD = 0.06 s), Unpredictable words (0.53 s, SD = 0.07 s)

Clopper and Pierrehumbert (2008); Scarborough (2010); Kraljic et al. (2008)

Discussion

- Attentional sets affected perceptual learning
  - Conditions that did not promote perception-oriented attentional sets showed larger effects
- Predictability was not an effective attentional set manipulation
  - Instead, allowed for attribution of the modified category to predictability

Implications for theoretical models

- Supports hierarchical representations
- Attention to episodic representations or specific pronunciations inhibits learning in abstract categories
- Attention as a gain mechanism is not supported
  - Perception-oriented attentional sets would have larger effects
  - Valency of attention may play a role

Implications - Dialects

- Perceptual learning of salient dialectal features may be inhibited
  - New Zealand/Australian English: *fish and chips*
  - New Zealand/North American English: *Bret* vs *Brit*

Implications - Non-native accents

- Non-native listener / native speaker
  - Perception-oriented attentional set may be the default
- Native listener / non-native speaker
  - Perceptual learning inhibited when attending to pronunciation

Real world example:

- Students attending to the professor's message rather than pronunciation should perceptually adapt more
  - Timecourse of perceptual learning?
  - Size of perceptual learning?

Acknowledgements

Thank you to all those involved drafting and creating these experiments, particularly Molly Babel, Eric Vatikiotis-Bateson, Carla Hudson Kam, Graham Haber, Jamie Russell, Jobie Hui and Michelle Chan!

Feedback from various 530 helped tremendously in shaping the ideas and experiments in this thesis.

Thank you all for your attention!

References I

Ahissar, M. and Hochstein, S. (1993). Attentional control of early perceptual learning. *Proceedings of the National Academy of Sciences of the United States of America*, 90(12):5718–22.

Bradlow, A. R. and Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, 106(2):707–729.

Clopper, C. G. and Pierrehumbert, J. B. (2008). Effects of semantic predictability and regional dialect on vowel space reduction. *Journal of the Acoustical Society of America*, 124(3):1682–1688.

Kalikow, D., Stevens, K., and Elliott, L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *. . . Journal of the Acoustical Society of . . .*, 61(5).

Kraljic, T., Brennan, S. E., and Samuel, A. G. (2008). Accommodating variation: dialects, idiolects, and speech processing. *Cognition*, 107(1):54–81.

References II

Kraljic, T. and Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive psychology*, 51(2):141–78.

Kuhl, P. K. (1979). Speech perception in early infancy: perceptual constancy for spectrally dissimilar vowel categories. *The Journal of the Acoustical Society of America*, 66(6):1668–1679.

Labov, W. (1997). The Social Stratification of (r) in New York City Department Stores. In Coupland, N., editor, *Sociolinguistics: A Reader*, pages 168–178. St. Martin's Press.

Li, F., Munson, B., Edwards, J., Yoneyama, K., and Hall, K. (2011). Language specificity in the perception of voiceless sibilant fricatives in Japanese and English: implications for cross-language differences in speech-sound development. *The Journal of the Acoustical Society of America*, 129(2):999–1011.

Lieberman, P. (1963). Some Effects of Semantic and Grammatical Context on the Production and Perception of Speech. *Language and Speech*, 6(3):172 –187.

References III

Norris, D., McQueen, J. M., and Cutler, A. (2003). Perceptual learning in speech.

Pitt, M. and Szostak, C. (2012). A lexically biased attentional set compensates for variable speech quality caused by pronunciation variation. *Language and Cognitive Processes*, (April 2013):37–41.

Reinisch, E., Weber, A., and Mitterer, H. (2013). Listeners retune phoneme categories across languages. *Journal of experimental psychology. Human perception and performance*, 39(1):75–86.

Reinisch, E., Wozny, D. R., Mitterer, H., and Holt, L. L. (2014). Phonetic category recalibration: What are the categories? *Journal of Phonetics*, 45:91–105.

Scarborough, R. (2010). Lexical and contextual predictability: Confluent effects on the production of vowels. *Laboratory phonology*, pages 575–604.

References IV

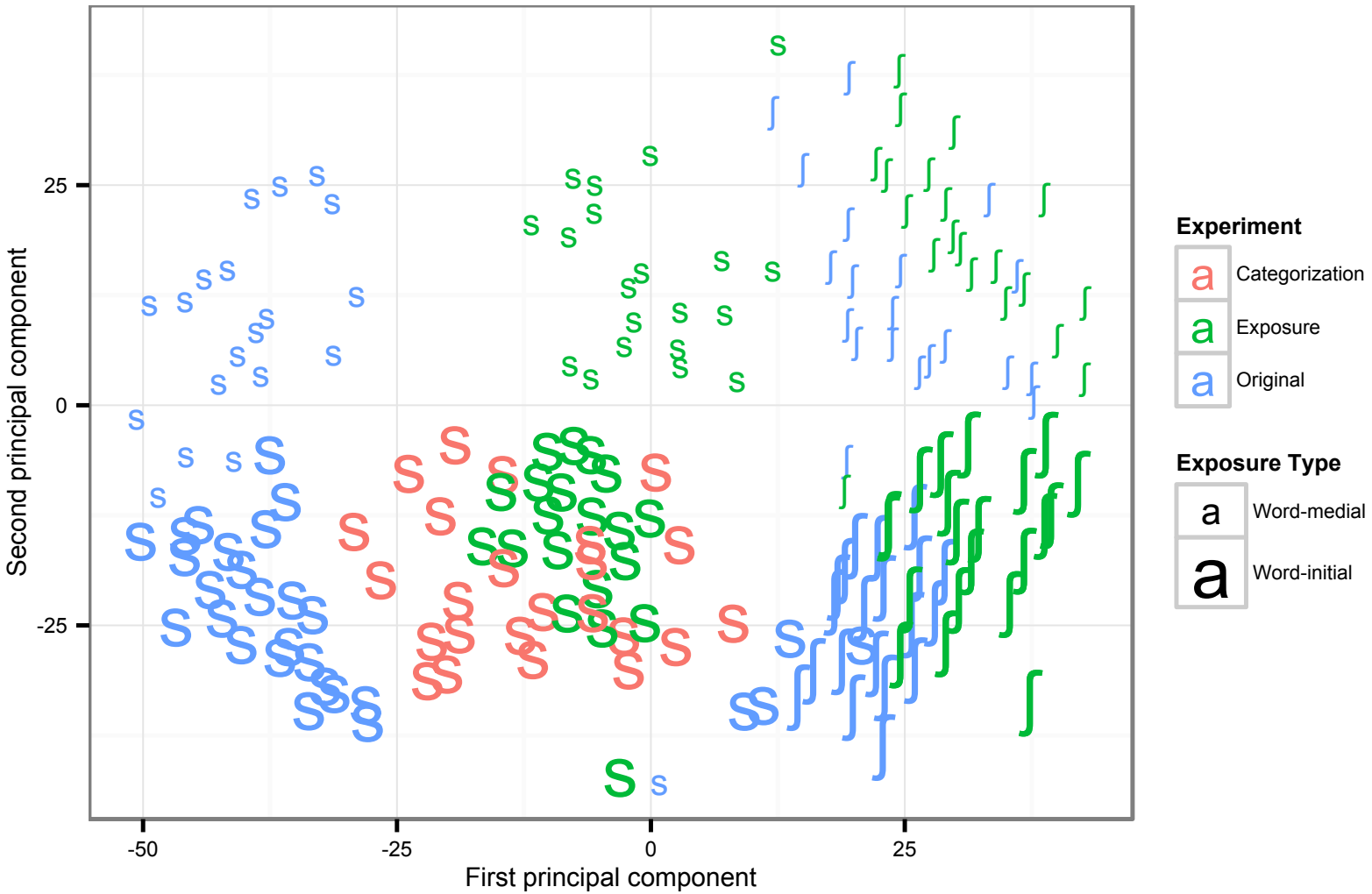
Shankweiler, D., Strange, W., and Verbrugge, R. R. (1977). Speech and the problem of perceptual constancy. *Perceiving, acting, and knowing: Toward an ecological psychology*, pages 315–345.

Strand, E. A. and Johnson, K. (1996). Gradient and visual speaker normalization in the perception of fricatives. In *KONVENS*, pages 14–26.

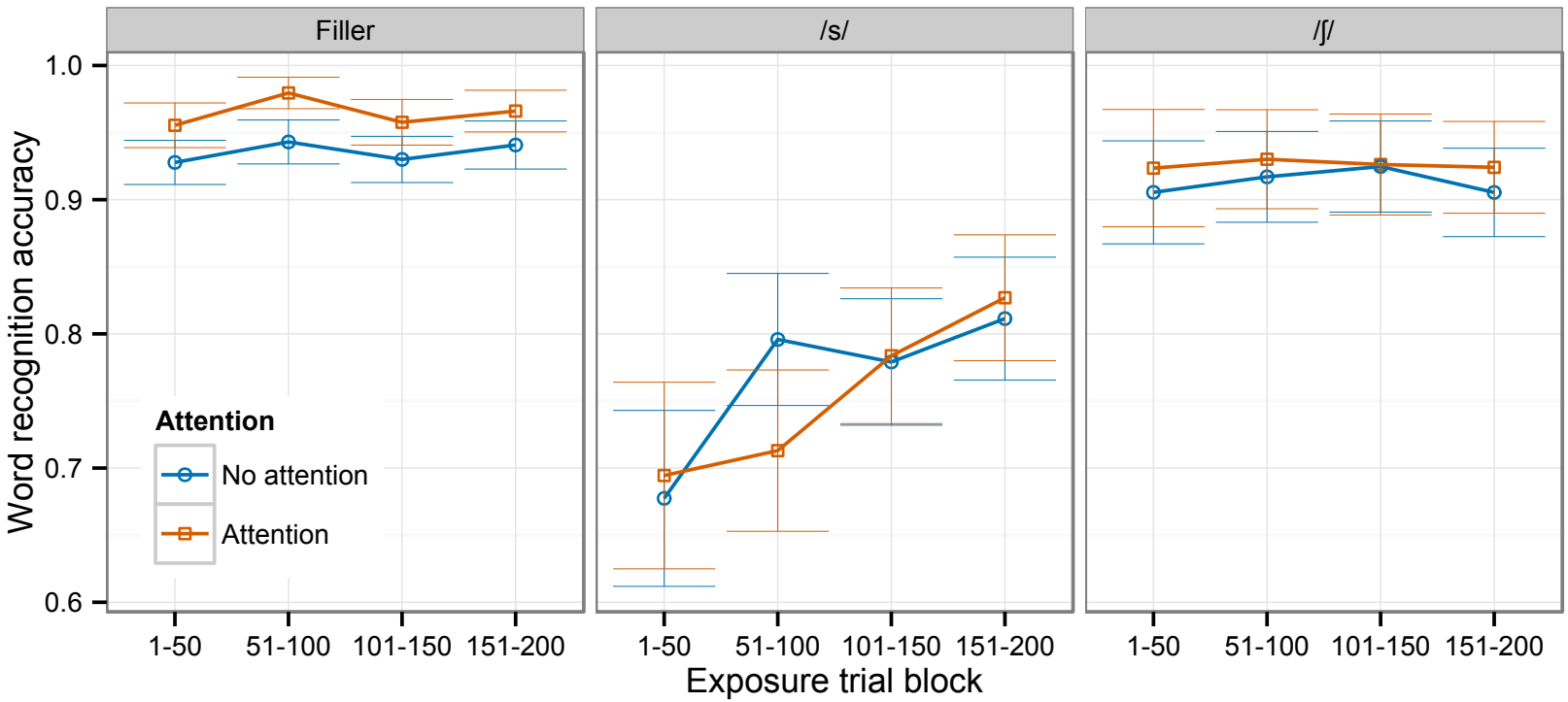
Sumner, M. and Kataoka, R. (2013). Effects of phonetically-cued talker variation on semantic encoding. *The Journal of the Acoustical Society of America*, 134(6):EL485.

Vroomen, J., van Linden, S., de Gelder, B., and Bertelson, P. (2007). Visual recalibration and selective adaptation in auditory-visual speech perception: Contrasting build-up courses. *Neuropsychologia*, 45(3):572–577.

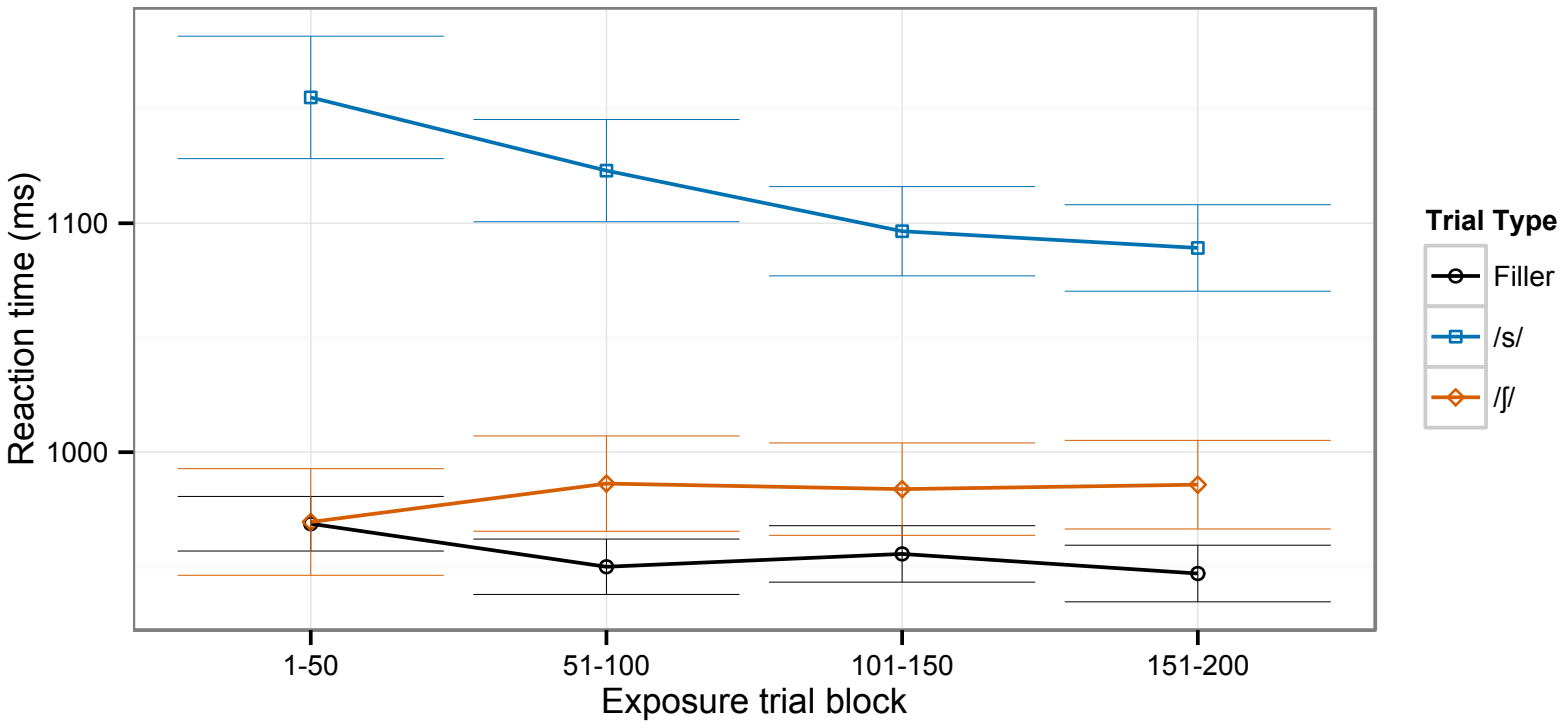
Experiment 1 - Acoustic analysis



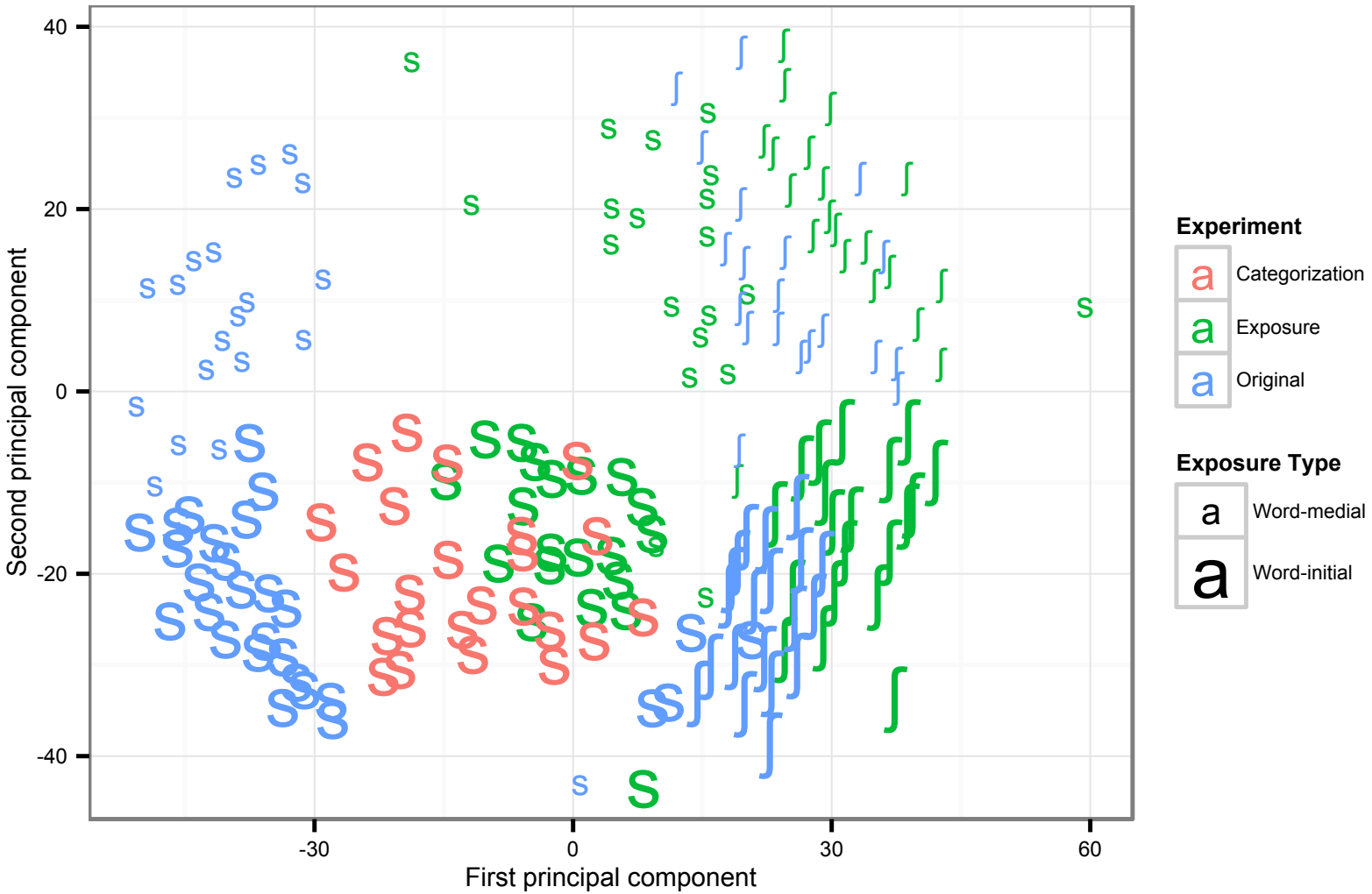
Experiment 1 - Exposure ACC



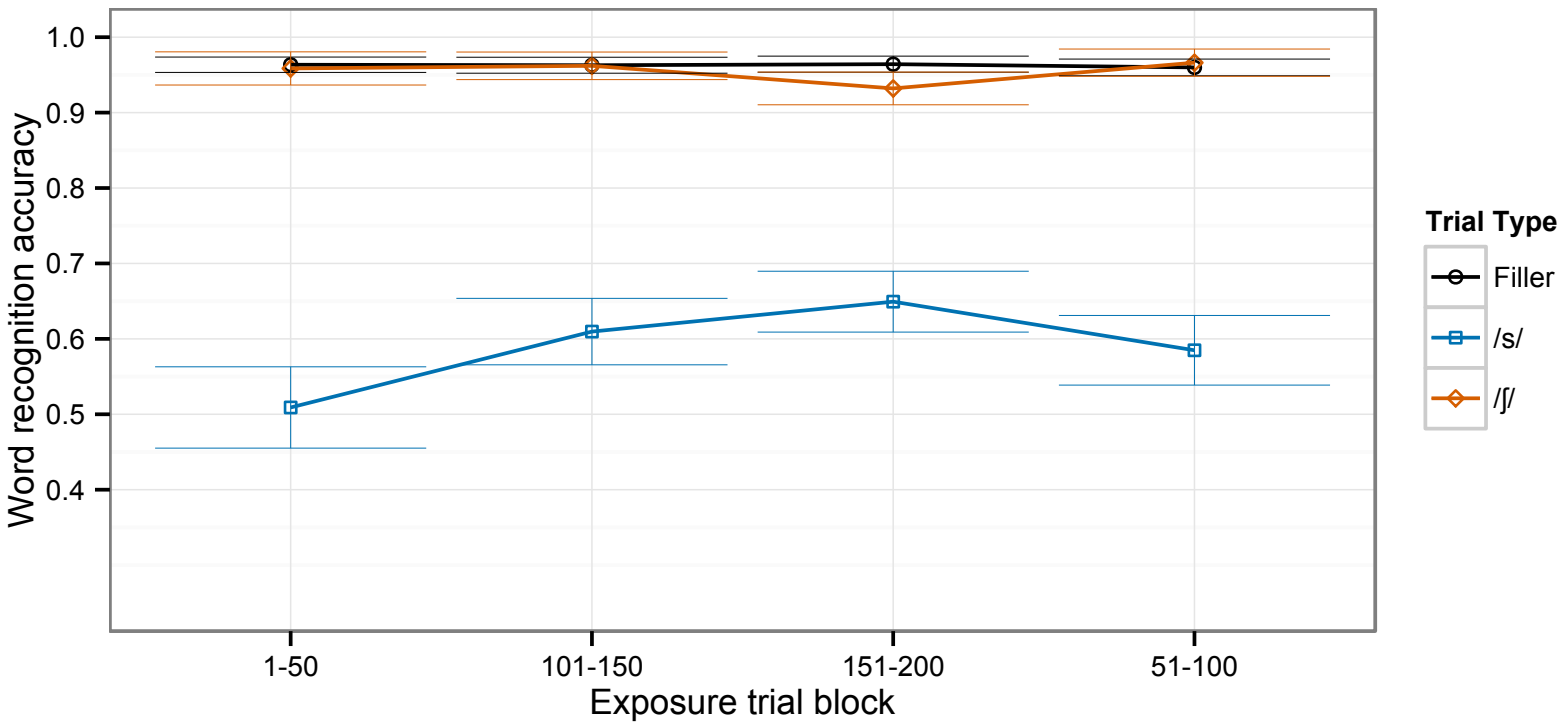
Experiment 1 - Exposure RT



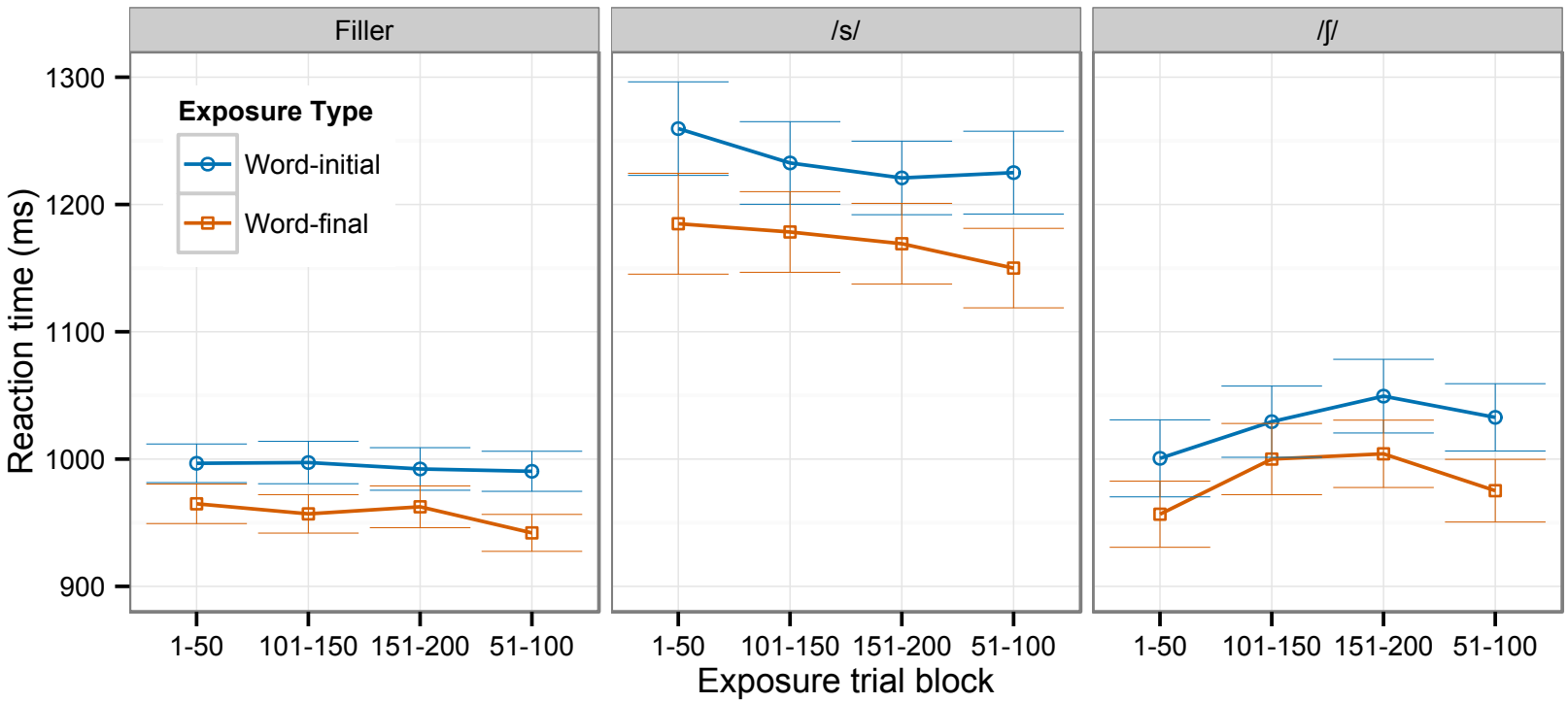
Experiment 2 - Acoustic analysis



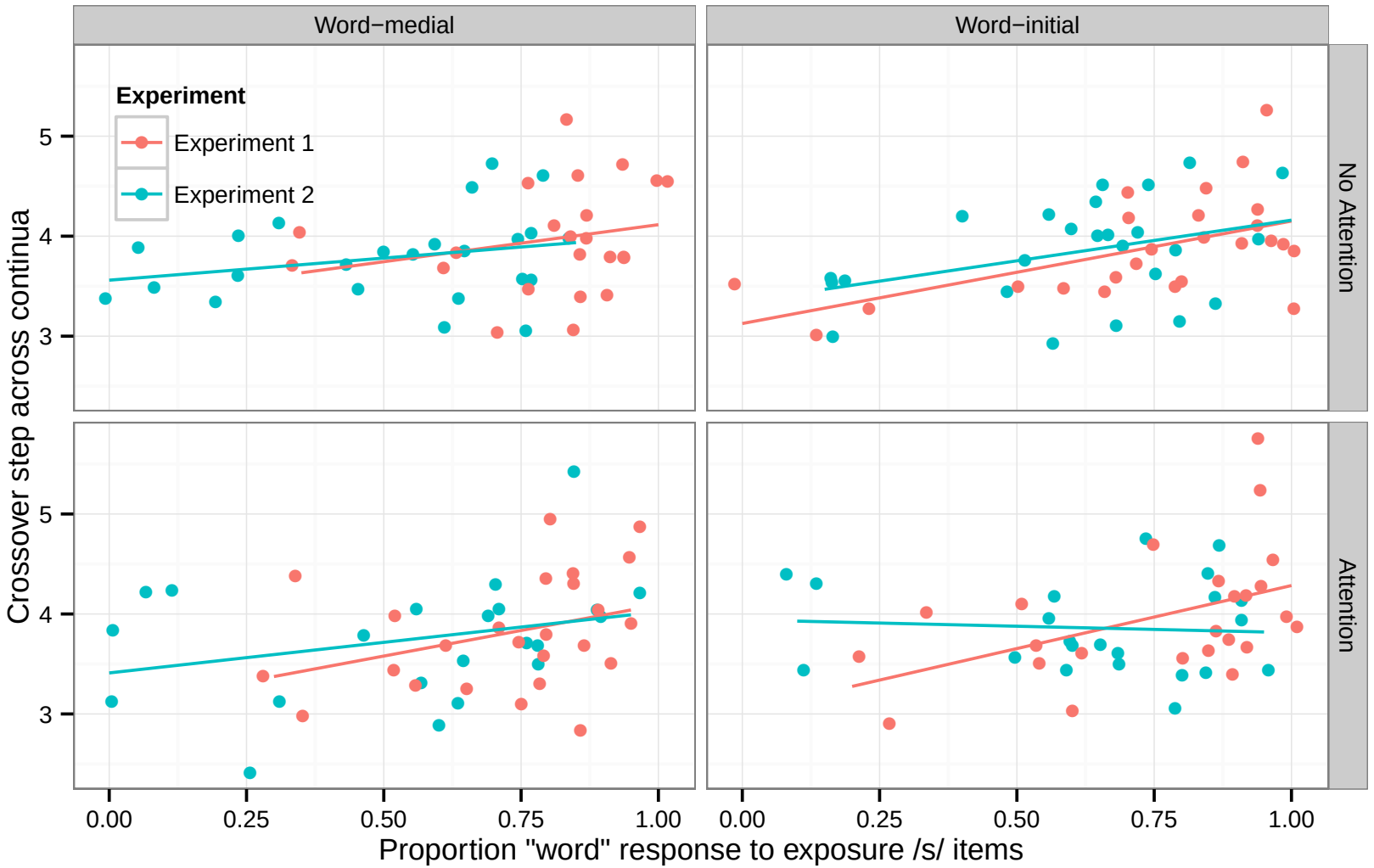
Experiment 2 - Exposure ACC



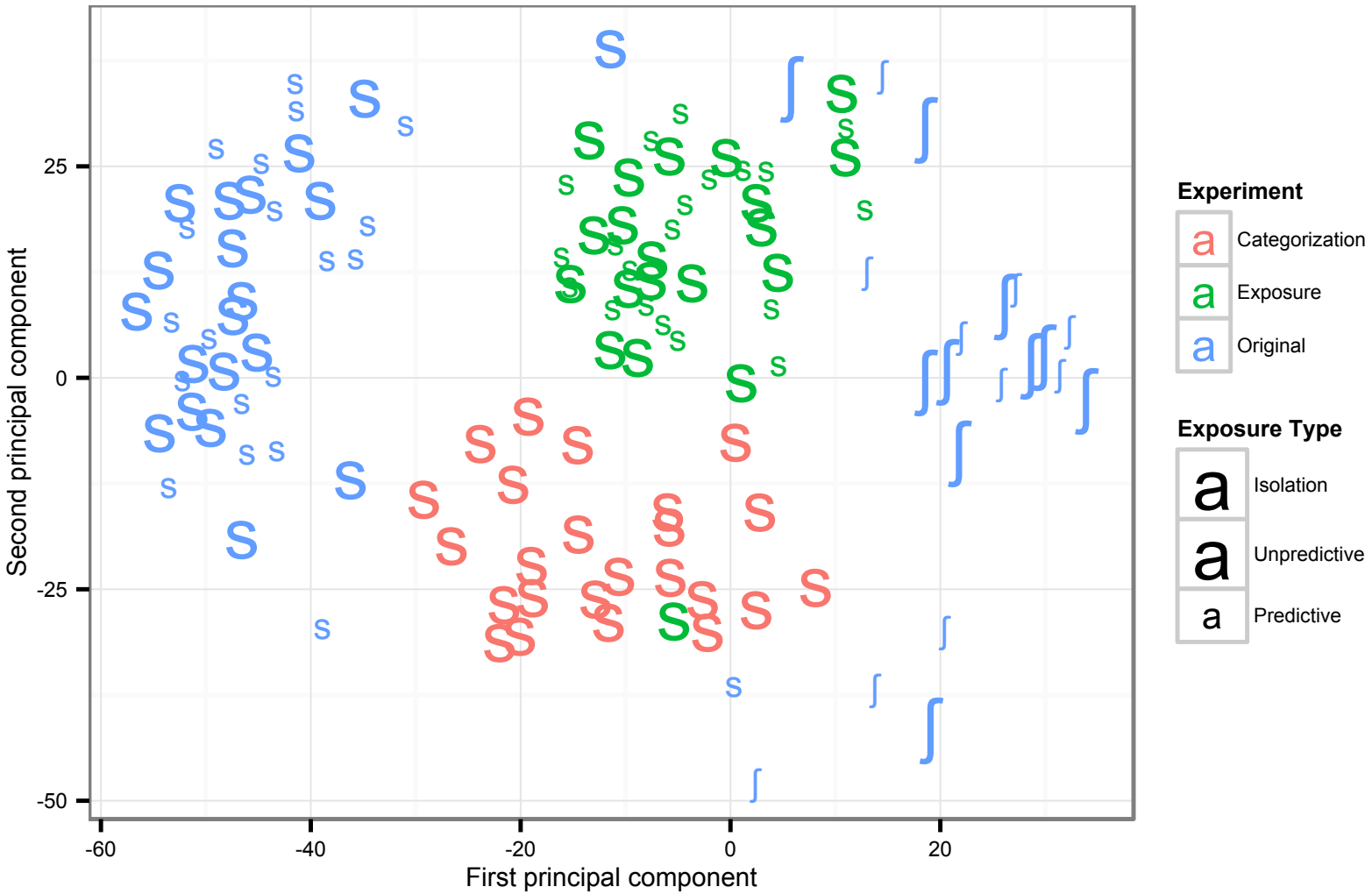
Experiment 2 - Exposure RT



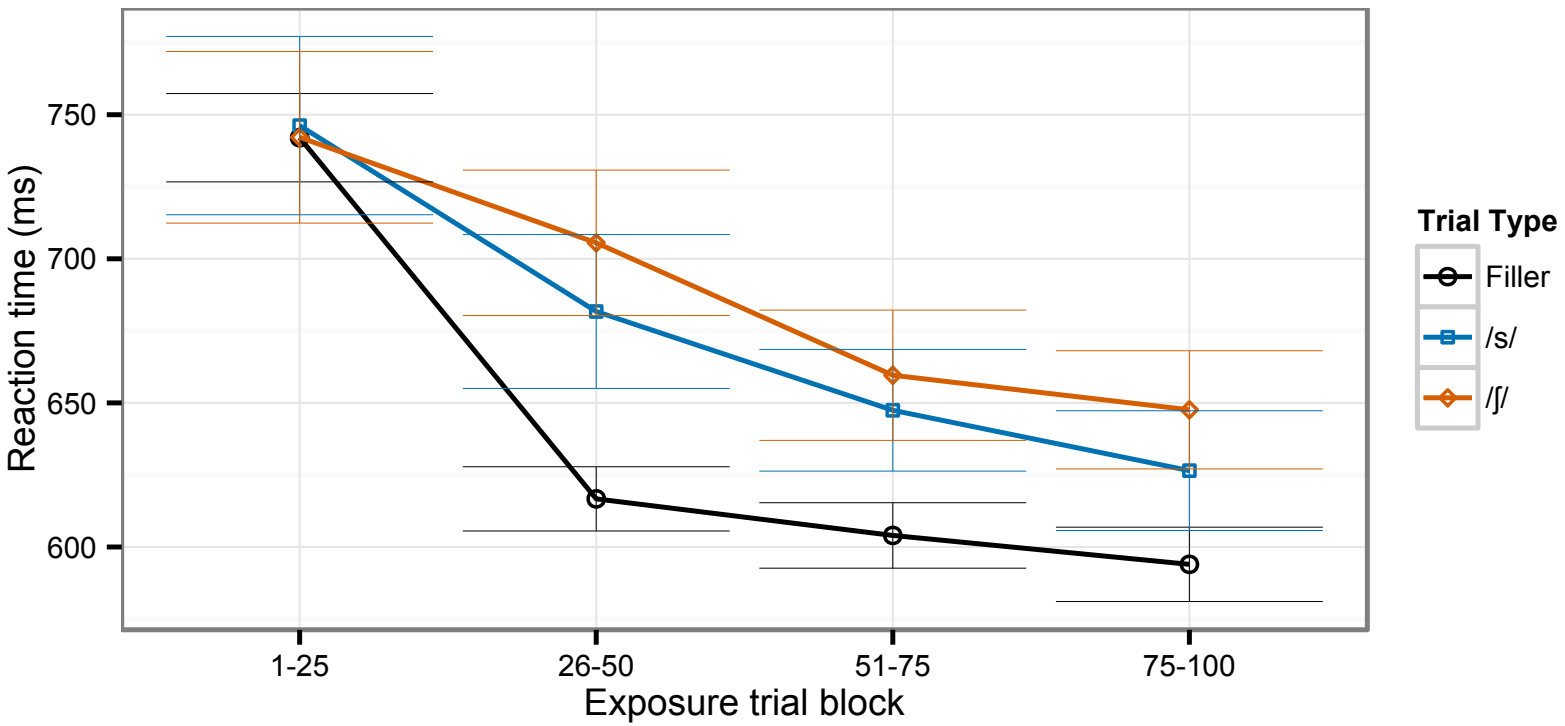
Experiments 1 and 2 - Cross-over and word responses



Experiment 3 - Acoustic analysis



Experiment 3 - Exposure RT



Experiment 1 and 3 - Cross-over distribution