

Attention and salience in lexically-guided perceptual learning

Michael McAuliffe

PhD Defense

Outline

- 1 Background
 - Sources of variation
 - Perceptual learning
 - Attentional sets
- 2 Experiments 1 and 2
 - Set up
 - Results
 - Summary
- 3 Experiment 3
 - Set up
 - Results
 - Summary
- 4 Discussion

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
 - Style
 - Speaking rate
 - Coarticulation (/stʌ/)
 - Position in word
 - Predictability

Listener

- Contextual
 - Speaking rate
 - Coarticulation (/stʌ/)
 - Position in word
 - Predictability

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

Sources of variation

Example: /s/

Speaker

- Attention
 - Tongue twisters

Listener

- Attention
 - Comprehension
 - Perception

Nozari and Dell (2012); Pitt and Szostak (2012)

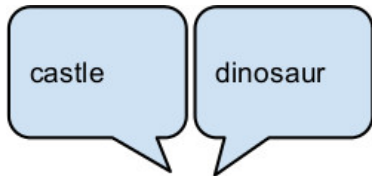
Perceptual constancy

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

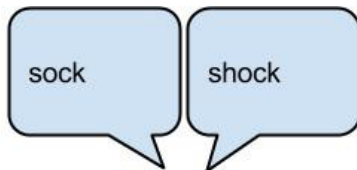


Shankweiler et al. (1977); Kuhl (1979); Sumner and Kataoka (2013)

Perceptual learning

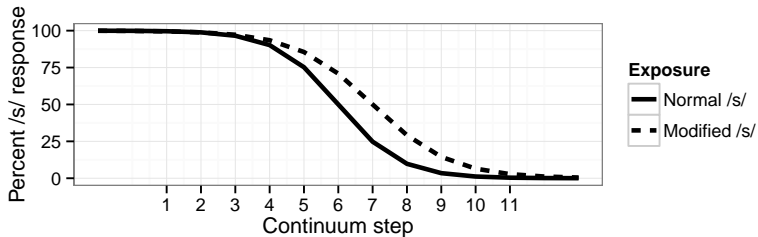
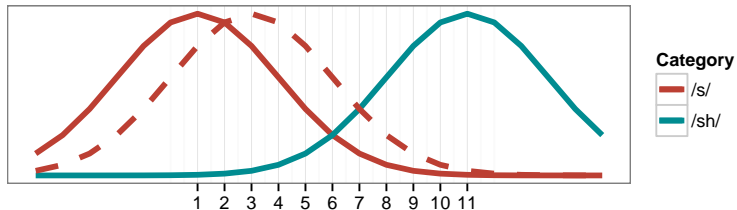


Exposure



Categorization

Categorization



Research question

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

Sources of variation

Example: /s/

Speaker

- Contextual
 - Style
 - Speaking rate
 - Coarticulation (/stʌ/)
 - **Position in word**
 - **Predictability**

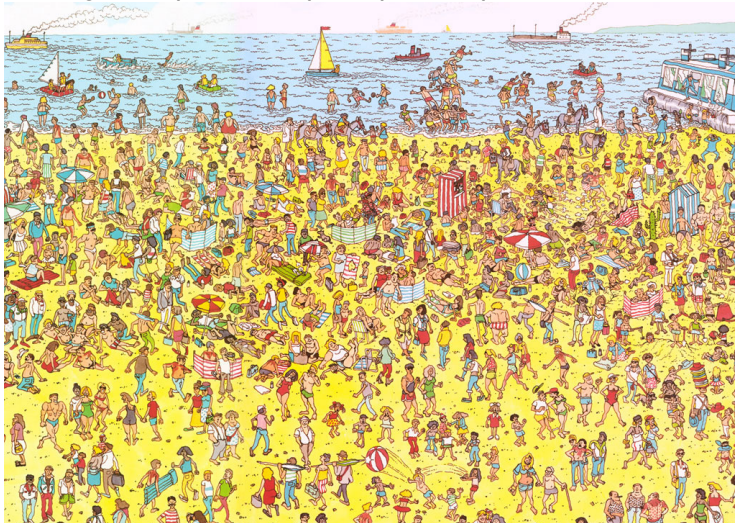
Listener

- Contextual
 - Speaking rate
 - Coarticulation (/stʌ/)
 - Position in word
 - Predictability
- **Attention**
 - Comprehension
 - Perception



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 an accent or affectation
 - 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
 - Audio-visual lipreading (nonwords)
 - Psychophysics perceptual learning

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

Generalization in attentional sets

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.

Attentional set manipulation

Explicit instructions

- “This speaker’s ‘s’ sounds are ambiguous”
- Promote perception-oriented attentional set

Pitt and Szostak (2012)

Attentional set manipulation

Perceptual salience of modified /s/

- The less predictable an element, the higher its salience
- Increase the likelihood of listener noticing modification
- Promote perception-oriented attentional set
- Assumption: similar to increasing the number of /s/ trials relative to filler trials

Position in word

- Listeners are more tolerant of variation later in the word
- Word-initial modified /s/ should be more salient

Category typicality

- Productions that are unexpected for a category are more likely to be noticed (salient)

Pitt and Szostak (2012)

Experiments 1 and 2

Experiment 1

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

Experiment 2

- Same design and materials as Experiment 1
- 30% word response rate in the pre-test (more atypical /s/; Word-initial vs word-medial)

Sample trials

Exposure

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

Categorization

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

Experiment 1 and 2 predictions

- Hypothesis 1:
 - Perceptual learning is affected by attentional sets
 - Perceptual learning should be less where perception-oriented attentional sets are promoted
- Hypothesis 2:
 - Perceptual learning is wholly automatic and consistent
 - Equal perceptual learning effects across all conditions
- Hypothesis 3:
 - Perceptual learning effects are dependent on similarity
 - Word-initial exposure > Word-medial exposure

Experiment 1 - Word-initial exposure

Exposed to ambiguous /s/

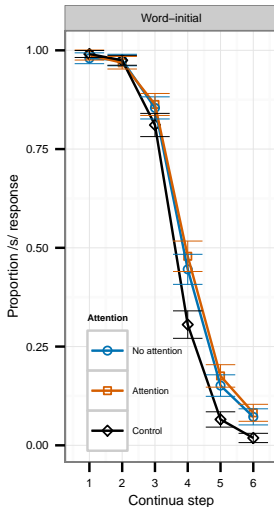
- 50% between /s/ and /ʃ/

Attention

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

Position of /s/

- *Word initial*
- Word medial



Experiment 1 - Word-medial exposure

Exposed to ambiguous /s/

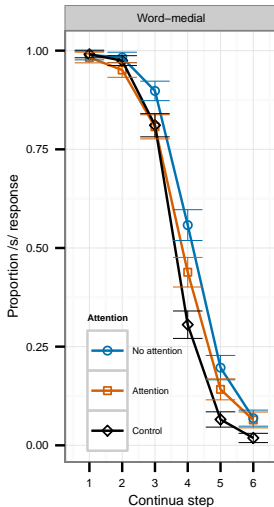
- 50% between /s/ and /ʃ/

Attention

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

Position of /s/

- Word initial
- *Word medial*



Experiment 2 - Word-initial exposure

Exposed to ambiguous /s/

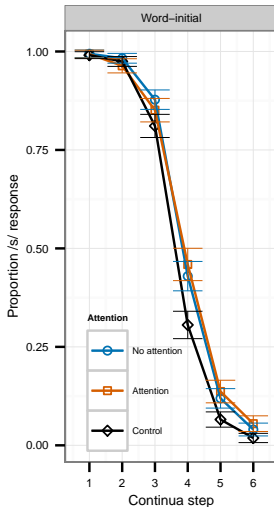
- More like /ʃ/ than /s/

Attention

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

Position of /s/

- *Word initial*
- *Word medial*



Experiment 2 - Word-medial exposure

Exposed to ambiguous /s/

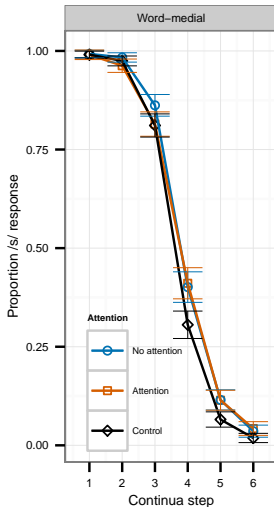
- More like /ʃ/ than /s/

Attention

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

Position of /s/

- Word initial
- *Word medial*



Summary

- Results align with attentional sets
- Fine-grained similarity did not appear to play a role
 - Word-initial exposure \leq word-medial exposure
- Conditions promoting a perception-oriented attentional set
 - Still showed perceptual learning
 - Had smaller perceptual learning effects
 - Did not differ from one another
- Task was comprehension-oriented (identifying word)
- Experiment 3 attempts to further promote comprehension-oriented attentional sets

Experiment 3

- Novel cross-modal paradigm
 - Auditory sentences
 - Identification of picture corresponding to final word in sentence
 - Same word-medial modified /s/ stimuli
 - Final targets were predictable or unpredictable
- Across subjects
 - Instructions (identical to Experiments 1 and 2)
 - Modified /s/ only in predictable or unpredictable words
- Predictable words are predicted to have lower salience than unpredictable words

Sample trials

Exposure

- Hear:
 - Predictable: The traffic cop alerted the driver by blowing her whistle (Audio)
 - Unpredictable: The boy ran away when he heard the whistle(Audio)
 - Picture 1 vs picture 2 FIXME

Categorization

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

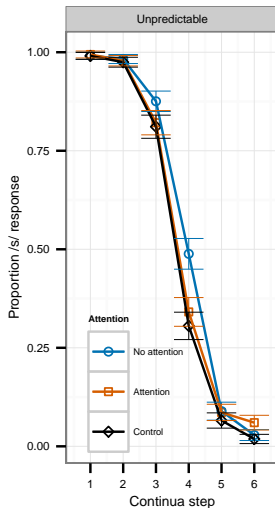
Experiment 3 predictions

- Hypothesis 1:
 - Words in isolation == words at the end of sentences
 - Perceptual learning effects == Experiment 1's Word-medial conditions
- Hypothesis 2:
 - Words in isolation \neq words at the end of sentences
 - Perceptual learning effect < Experiment 1's Word-medial conditions
- Hypothesis 3:
 - High predictability is associated with less distinct acoustics
 - Perceptual learning is not found in coarticulation contexts (/stu/)
 - No perceptual learning effect in predictable condition

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

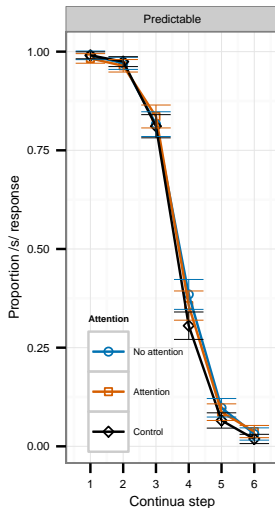
Experiment 3 - Unpredictable exposure

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

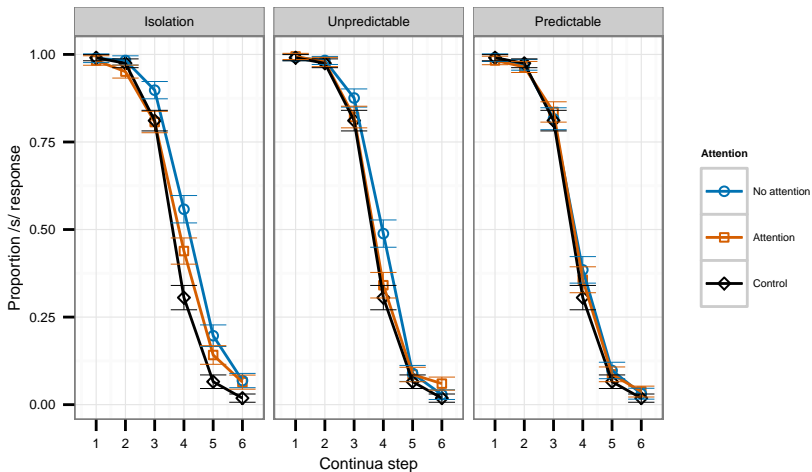


Experiment 3 - Predictable exposure

- **Exposed to 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 (/stu/)
 - Despite consistent durations for words and sibilants across the two sentence types

Discussion

- Attentional sets affected perceptual learning
 - Conditions that did not promote perception-oriented attentional sets showed larger effects
- Predictability was likely not an 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

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.
- Kraljic, T., Brennan, S. E., and Samuel, A. G. (2008). Accommodating variation: dialects, idiolects, and speech processing. *Cognition*, 107(1):54–81.
- Kraljic, T. and Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive psychology*, 51(2):141–78.

References II

- 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.
- 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.
- Norris, D., McQueen, J. M., and Cutler, A. (2003). Perceptual learning in speech.
- Nozari, N. and Dell, G. S. (2012). Feature migration in time: Reflection of selective attention on speech errors.

References III

- 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.
- 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.

References IV

- 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.