A lexical locus for the integration of asynchronous cues to voicing:
An investigation with natural stimuli

Joseph C. Toscano and Bob McMurray
University of Iowa Dept. of Psychology

Multiple acoustic cues in speech
- Multiple cues contribute to phonetic categorization
- VOT contributes to both voicing and manner categorization
- VOT effects are often combined with other cues
- Synthesis of cues is a common linguistic strategy

Context effects in natural and synthetic speech
- VOT contributes to both voicing and manner categorization
- VOT effects are often combined with other cues
- Synthesis of cues is a common linguistic strategy

Issues
- Are synthetic and natural stimuli perceived differently?
- Can we see effects of multiple cues in both natural and synthetic speech?
- Are multiple cues used in natural speech?
- Are cues integrated in the same way in synthetic and natural speech?

Questions
- Under what circumstances are multiple cues used?
- When are multiple cues used?
- Are cues used in the same way in synthetic and natural speech?
- Are cues integrated in the same way in synthetic and natural speech?

Procedures
- Participants: 30 undergraduates at the University of Iowa participated in the experiment.
- Stimuli: Continuous speech from a particular speaker.
- Background noise: babble noise from multiple talkers.
- Procedure: Participants were asked to identify the voice of the speaker.

Procedure
- Participants: 30 undergraduates at the University of Iowa participated in the experiment.
- Stimuli: Continuous speech from a particular speaker.
- Background noise: babble noise from multiple talkers.
- Procedure: Participants were asked to identify the voice of the speaker.

References

Visual World Paradigm
- Target = "spoon"
- Competitor = "steak"
- Unrelated = "stick", "steak"

Vowel length
- VOT contributes to both voicing and manner categorization
- VOT effects are often combined with other cues
- Synthesis of cues is a common linguistic strategy

Eye movement results
- Expected pattern of eye movements, similar to previous visual world studies.
- Fixations to target, competitor, and unrelated items as a function of time for stimuli with a 0 ms VOT.

Conclusions
- VOT is a cue to voicing in natural speech.
- VOT effects are often combined with other cues
- Synthesis of cues is a common linguistic strategy

Locus of cue integration
- With synthesized stimuli, cues are integrated independently in time.
- With natural speech, cues are integrated simultaneously in time.

Predictions
- With synthesized stimuli, cues are integrated independently in time.
- With natural speech, cues are integrated simultaneously in time.
- Eye movements reflect which referents are considered during online word recognition.

Effect of VOT near category boundary
- Significant vowel effect near category boundary (where VOT is ambiguous), but not at endpoints (where VOT is more reliable)
- Proportion of looks to competitor
- Increased looks to /p/-competitor with short VOT, and to /b/-competitor with long VOT.

Conclusion
- VOT is a cue to voicing in natural speech.
- VOT effects are often combined with other cues
- Synthesis of cues is a common linguistic strategy

Work in progress: Effect of carrier phrase
- Cue weighting may differ in running speech.
- Cue weighting may differ in running speech.
- Cue weighting may differ in running speech.

Responses with Carrier Phrases
- Target word only
- Carrier phrase
- Carrier phrase creates larger VOT effect.
- Suggests that multiple cues are used in more natural contexts when individual cues may be less reliable than in isolation.

General Conclusions
- Gradual lexical activation in natural speech.
- Multiple cues to voicing are used in natural speech.
- Cues are used by the hearer.
- Demands of running speech may change utility.

Acknowledgments
- We would like to thank Meghan Clayards for theoretical discussions of these issues, Molly Robinson for assistance with stimulus preparation, and the students in the MACLab at the University of Iowa for helping to run subjects in the experiment.

Eye tracking will reveal influence of VL in natural speech, as differences in fixations to competitors (activation).

Natural vs. synthetic speech
- Natural speech: VOT is usually unambiguous
- Synthetic speech: VOT is immediately ambiguous
- Fixations to target, competitor, and unrelated items as a function of time for stimuli with a 0 ms VOT.

Summary
- Previous work: VL may not be used as a cue to voicing in natural speech.
- Previous work: VL is a cue to voicing in natural speech.
- Previous work: VL is a cue to voicing in natural speech.

Immediacy:
- Immediate evaluation.
- Immediate evaluation.
- Immediate evaluation.

Visual speech
- VOT is a cue to voicing in natural speech.
- Previous work: VL is a cue to voicing in natural speech.
- Previous work: VL is a cue to voicing in natural speech.

Online word recognition (Tanenhaus et al., 1995)
- Online word recognition (Tanenhaus et al., 1995)
- Online word recognition (Tanenhaus et al., 1995)
- Online word recognition (Tanenhaus et al., 1995)