Defining "high quality" tokens of tone in Mandarin Infant-Directed Speech

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Emily Moeng
moeng@live.unc.edu

Jen Boehm
jengriff@live.unc.edu

Amy Reynolds
amyrey@live.unc.edu

Overview

- Phoneme acquisition involves a **distributional learning mechanism** (Maye, Werker, and Gerken, 2002)
- However, even Infant-Directed Speech (IDS) exhibits a great amount of **phonetic variation**
- It's been claimed that **prosody** can indicate to the learner which tokens are **"high quality"** tokens, giving the learner a subset of less variable tokens to learn from, therefore aiding in distributional learning (e.g. Adriaans and Swingley, 2012)
- Prosodic characteristics suggested include:
  - ❑ Longer duration
  - ❑ Higher average pitch
  - ❑ Greater pitch range
  - … which have been shown to yield a less variable subset of vowel tokens (Adriaans and Swingley, 2012)

Question: Do these prosodic characteristics yield a "high quality" subset of tokens for lexical tone, a phonemic distinction which makes use of (some of) the same phonetic cues (i.e. F0)? Specifically, **do these prosodic characteristics expand the tonal space in infant-directed speech?**

Answer: Longer duration and higher pitch expand the tonal space, pitch range does not

Background: Prosodical cues indicate "high quality" tokens

- Adriaans and Swingley (2012) show that "high quality" tokens expand the vowel space, with "high quality" tokens defined as…
  - ❑ Being longer in duration,
  - ❑ Having a higher average pitch, *and/or*
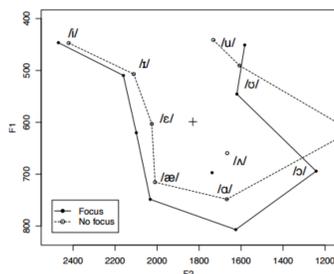 - ❑ Having a greater pitch range

Figure from Adriaans and Swingley (2012). Note that the subset of "focused" ("high quality") vowels expands the vowel space

- Adriaans and Swingley suggest that these prosodically-focused tokens are higher-quality examples of the categories to be learned, therefore aiding a learner in distinguishing categories from a high-variability signal

The Problem: Lexical tone and prosodic pitch both make use of F0

- Since pitch (F0) is the main acoustic correlate for lexical tone and an acoustic cue for prosody, it is not clear if this proposal would extend to tones.
- For example, Tone 1 is level, whereas Tone 4 is a falling tone. Since Adriaans and Swingley (2012) use pitch change as one criteria for determining whether a token is a "high quality" token or not, Tone 4 would yield more "high quality" tokens simply as a result of being a contour tone.

Research Question: Do all of these prosodic characteristics (greater duration, higher average pitch, and greater pitch range) expand the tonal space in infant-directed speech?
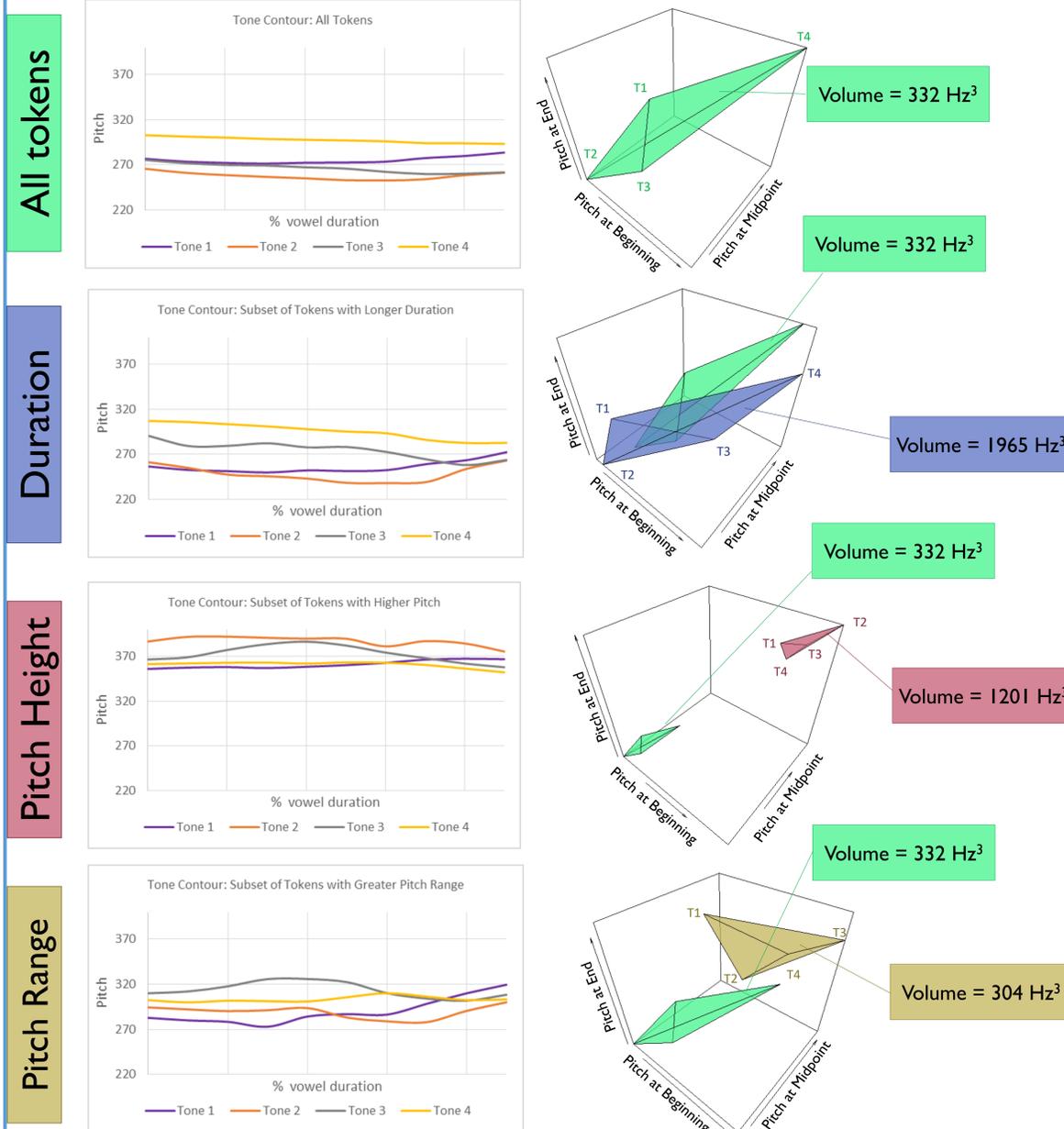
Methodology

- 3 transcribers marked in Praat the vowels spoken by a female Mandarin speaker interacting with her child from the Chinese Tong Corpus (Deng and Yip, 2015) in CHILDES (MacWhinney, 2000)
- For each vowel, measured vowel duration and F0 at 10 points along the contour using a Praat script
- Tokens were categorized into the following subcategories:
  - **Longer Duration Tokens**: Duration has a z-score greater than 0.5
  - **Higher Pitch Tokens**: Pitch has a z-score greater than 0.5
  - **Greater Pitch Range Tokens**: Pitch has a z=score greater than 0.5

Tone	Total Number of Tokens	Longer Duration Tokens	Higher Pitch Tokens	Greater Pitch Range Tokens
Tone 1	565	224	160	51
Tone 2	499	119	116	81
Tone 3	491	115	183	108
Tone 4	971	220	400	111

Results: Duration and Pitch Height expanded the tonal space, but Pitch Range did not

All tokens

Volume = 332 Hz³

Volume = 332 Hz³

Duration

Volume = 1965 Hz³

Volume = 332 Hz³

Pitch Height

Volume = 1201 Hz³

Volume = 332 Hz³

Pitch Range

Volume = 304 Hz³

Conclusion

- Longer vowel duration expanded the tonal space as predicted
  - Supports Adriaans and Swingley's suggestion that duration may indicate "high quality" tokens to a learner, even for a suprasegmental, like tone
  - Individual tones become more distinct from each other in the tonal space
  - Easier for a learner to perceive and acquire individual tones
- Pitch Height Tokens also expanded the tonal space
- Pitch Range Tokens had no effect on the tonal space
  - Mandarin tones have specific contours, thus prohibiting greater pitch ranges for level tones
  - Fewest tokens of level tones in the subset of tokens with greater pitch range (see token count for Tone 1 in Methodology section)
- The tonal contours created by the subset of Longer Duration Tokens was aligned with the contours created by all of the tokens more than any other subset of tokens
  - Likely because the Higher Pitch Tokens subset did not include "good" examples of low tones that were produced as low tones, and the Greater Pitch Range Tokens subset did not include "good" examples of level tones being produced as level tones

Discussion

- Data from a single speaker was used to avoid having to normalize each speaker's tonal space, but in the future we would like to include multiple speakers
- The current study focuses on a language which uses tone contrastively
  - If we were to analyze IDS in a language with contrastive vowel duration, would duration act like pitch range and have no effect on the durational space?
- Tone contours found differ greatly from traditional descriptions of Mandarin tone contours, suggesting connected speech greatly affects tone:

  - Liu et al. (2007) *did* find expected tone contours in Mandarin Chinese IDS, but only when analyzing certain target words (all were nouns)

- In the future, we may need to include information regarding sentence position, which learners may need to "undo" to learn the citation form
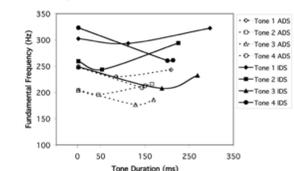
Selected References

- Adriaans, F. & Swingley, D. (2012). Distributional learning of vowel categories is supported by prosody in infant-directed speech. In Miyake, Peebles, & Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Deng, X. & Yip, V. (in press). A multimedia corpus of child Mandarin: the Tong corpus. Journal of Chinese Linguistics.
- Deng, X. & Yip, V. (2015). A corpus study of the acquisition of ba and bei constructions in Mandarin. Paper presented at The International Symposium on Psycholinguistics of Second Language Acquisition and Bilingualism, Chinese University of Hong Kong.
- Grieser, D.L. & Kuhl, P.K. (1988). Maternal Speech to Infants in a Tonal Language: Support for Universal Prosodic Features in Motherese. *Developmental Psychology*, 24, 14-20.
- Liu, H. et al. (2007). Acoustic Analysis of Lexical Tone in Mandarin Infant-Directed Speech. *Developmental Psychology*, 43, 912-917.
- MacWhinney, B. (2000). "The CHILDES project." Tools for Analyzing Talk, Part 1.
- Maye, J., J. Werker, L. Gerken. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*, B101-B111.