Interpreting Final Rises: Task and Role Factors

Catherine Lai
Centre for Speech Technology Research, University of Edinburgh

text

Introduction

- Studies generally examine how prosody affects the interpretation of the carrier utterance in the immediate context.
  - e.g. Whether a cue word is interpreted as a backchannel or not
  - e.g. Whether a speaker is uncertain about the content or requires ratification.

- While the local context clearly has a large effect on how prosody is interpreted, we would also like to know what impact higher-level features such as task and role have as well.

- Incongruence of findings based on varying dialogue types:
  - A correlation between rises and backchannels reported in map task dialogues in Bari Italian [Savino, 2011]; Swedish [Hedener et al., 2008], and Dutch [Caspers, 2000], as well as in other games corpora in English [Hockey, 1993; Elmes et al., 2007; Gravara, 2009].
  - Not found in studies of more free-form conversational dialogues in English [Steinberg et al., 1998; Ward, 2004; Truong and Heylen, 2010] and Hindi [Praud and Bal, 2010].

Questions:

- Do higher-level effects like task, role, and move type have an effect on whether an utterance was produced with rising features?

- Do differences in rise distributions backchannel studies extend to other sorts of dialogue moves?

Current Experiments

- Examine distribution of rises in IViE (International Variation in English) corpus [Grabe, 2004], free conversation versus map task in Standard Southern British English.
- Cambridge speakers are the most consistently ‘falling’ (read) dialect in IViE, so rises unlikely to just be a phonological boundary.
- Look at rises, turn-taking and affirmative responses.

IViE Corpus

- Designed to study differences across regions, speakers, and styles [Grabe, 2004]
- 6 male, 6 female speakers recorded between 1997-2000. The speakers were 16 years old at the time of recording and had been born in and grown up in the region.
- Read aloud isolated utterances and narratives.
- Map task (map): Instruction giver explains a pre-defined route around town on their map to the follower, who traced it out on their own (mismatched) map.
- Free conversation (conv): Participants discussed smoking, face-to-face. Speakers had the same role, which was simply a participant.

Sentence Types and Dialogue Moves

- Manually segmented into utterances into whole meaning units rather than phonological phrases (cf. Swerts and Geluykens [1994]).
- A conservative measure of the frequency of rises.
- Sentence type: Declarative (dec), Imperative (imp), Polar question (yes), Wh-question (whq), Tag (tag) question, Affirmative (afrm), Negative (neg), non-affirmative cue word (cue), If antecedent (ifa), NP or VP as an answer or a modifier (xp).
- Dialogue moves: Affirm, Neg, Contra (direct contradictions), CW (cue words), Inform, Instruct, Q (non-syntactically marked question), YMP (polar questions), WQH (wh-question). Tag (tag-question), sync (synchronize).
- No backchannel category. No map task affirmative could be clearly identified as a simple signal of attention.
- 430 conversational and 1287 map task utterances.

Boundary Pitch Features

- Target area: speech from the last prominence.
- Account for stress assignment in compounds (e.g. bowling alley) or pronoun deaccenting (e.g. about it).
- Utterances with speaker overlap at the target were excluded from the prosodic analysis [3% conv, 4% map].
- FO extracted using the Praat autocorrelation method.
- Parameter settings were automatically determined using the method described in Lu and Grabe [2010].
- Utterances with less than 5 FO points were discarded.
- Speaker normalized into semitones relative to the median F0 value.

- Higher level discourse factors, like task and role, have an effect on whether an utterance is produced with rising features or not.
- Content providing utterances in map-task dialogues had greater convexity than those from the conversational dialogue.
- No strong link between rising feature affirmative responses, speaker switches or stays.
- Need to take contextual conditions into account when collecting and analyzing intonational data.

Conclusion

- Higher level discourse factors, like task and role, have an effect on whether an utterance is produced with rising features or not.
- Content providing utterances in map-task dialogues had greater convexity than those from the conversational dialogue.
- No strong link between rising feature affirmative responses, speaker switches or stays.
- Need to take contextual conditions into account when collecting and analyzing intonational data.

Future Work

- Look at the relationship between frequency of rises and the overall quality of task-completion.
- Compare to other dialects, especially default rising ones such as Belfast English.
- Comparison of spontaneous monologue and dialogue.
Interpreting Final Rises: Task and Role Factors
Catherine Lai
Centre for Speech Technology Research, University of Edinburgh

References


Contact: clai@inf.ed.ac.uk