Boundary tone processing during online comprehension

Sarah Bibyk[1], Willemijn Heeren[2], Aaron Albin[3], Wil Rankinen[4], Anne Pier Salverda[5], Christine Gunlogson[5], & Michael K. Tanenhaus[5]
[1]Vanderbilt University, [2]Utrecht University,[3] Kobe University, [4]Grand Valley State University, [5]University of Rochester

Introduction

An intonational contour includes a pitch accent, a phrase accent, and a boundary tone[1]. Boundary tones are thought to be meaningful on their own (e.g. distinguish question from statement), but their meanings are also context-dependent and interact with other prosodic, lexical, and syntactic cues.

- How are boundary tones incorporated during online sentence comprehension?
  - Previous research suggests that listeners need the full contour to come to an interpretation[2,3]
  - Structural confounds however may have obscured evidence of earlier online processing
- When do listeners distinguish rising from falling intonation?
  - At (**1**) the pitch accent, (**2**) the transition from pitch accent to boundary tone (the "turning point"), or (**3**) after the full contour?

Methods

Procedure: Participants played a card game against a computer

- Goal: get rid of one's playing cards by matching them with a "match" card
- Current player: announce a match or ask the opponent to make a match
- Opponent: attempt to block the match (if a "blocking card" also matches) or check own playing cards to make a match

Computer's utterances on critical trials

- Same sentence with statement vs. question indicated by intonation only
  - Scenario A: *Got an armadillo.* Scenario B: *Got an armadillo?*

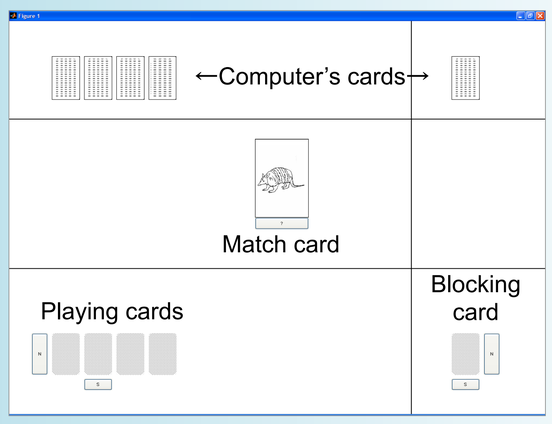
Eyetracking:

- Participants' fixations to playing cards vs. blocking card reflect their online interpretation of the computer's utterance as a question or a statement

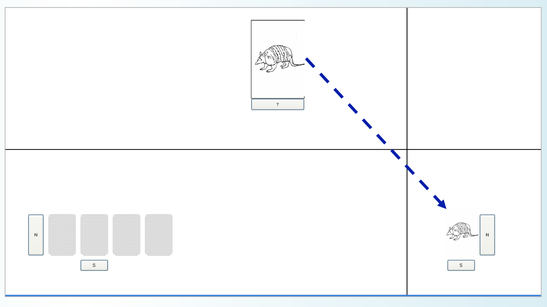
Demonstration:

On computer's turn

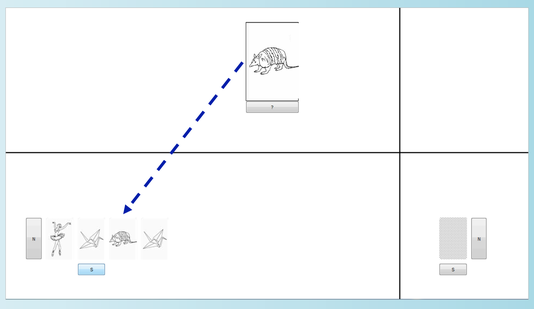
New match card appears in middle of screen



Scenario A: Computer says... "Got an armadillo."



Scenario B: Computer says... "Got an armadillo?"



Materials: Critical utterances end with 4-syllable word (penultimate stress)

- One of four words: *armadillo*, *ballerina*, *origami*, or *ravioli*
- Readily-identifiable 'turning points' in intonation contours:

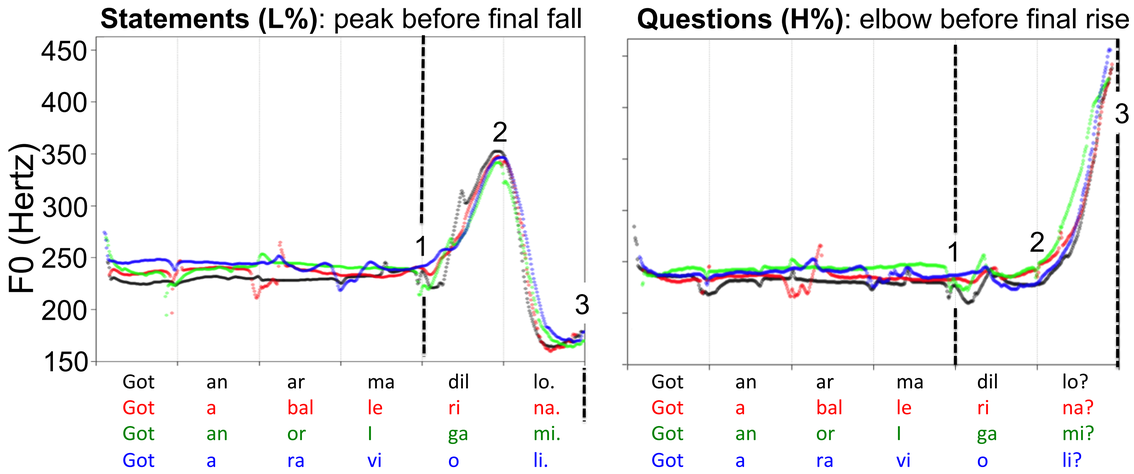


Figure 1: Time normalized F0 traces for target statements and questions

Filler sentences add variety:

- e.g. "I've got an armadillo." or "Do you have an armadillo?"

Table: acoustic characteristics of target utterances (All pre-recorded)

	Tone	Dur.	Mean F0 1st 4 σ's	F0 change final σ	Spect. balance final vowel
armadillo	L%	967 ms	230 Hz	-188 Hz	-23.4 dB
	H%	875 ms	233 Hz	+193 Hz	-19.0 dB
ballerina	L%	943 ms	237 Hz	-187 Hz	-16.7 dB
	H%	853 ms	234 Hz	+192 Hz	-14.4 dB
origami	L%	911 ms	240 Hz	-177 Hz	-29.8 dB
	H%	839 ms	244 Hz	+177 Hz	-23.7 dB
ravioli	L%	1000 ms	244 Hz	-177 Hz	-24.8 dB
	H%	878 ms	238 Hz	+171 Hz	-19.2 dB

Participants: Monolingual native speakers of American English (N=24)

Equipment: Head-mounted EyeLink II eyetracker (sampling at 250Hz)

Analysis and Results

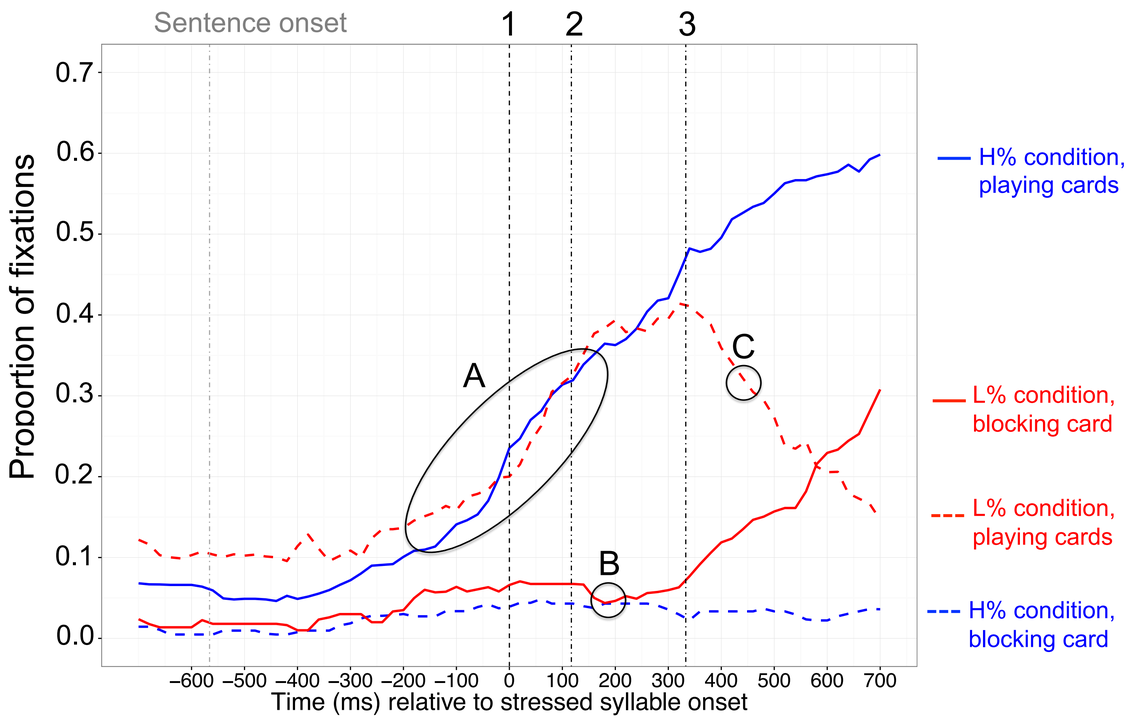


Figure 2: Proportion of fixations to playing and blocking cards by intonation

Condition of interest: Statements (L%, red)

- Breakpoint analysis: determine when changes in fixation proportions occur

Three major results:

Result A: Participants initially fixate playing cards

- "Got a X" construction is biased toward question interpretation[3]

Result B: Participants look *toward* target (blocking card, red solid)

- 180 ms after stressed syllable onset (**1**), 50 ms after F0 turning point (**2**)

Result C: Participants look *away* from competitor (playing cards, red dashed)

- 440 ms after stressed syllable onset (**1**), 320 ms after F0 turning point (**2**)

Interpretation: With **200 ms** delay for a change in fixations:

- Result B may be driven by the pitch accent at (**1**)
- Result C is most likely driven by the F0 turning point at (**2**)
- All changes are driven by information prior to sentence offset (**3**)

Conclusion

When do listeners infer a question vs. statement interpretation?

- Near turning point where pitch accent transitions to boundary tone:
  - Earliest acoustic evidence indicative of upcoming boundary tone
- Listeners need not wait and hear full contour (as previously suggested)

Future research using this paradigm:

- Explore when boundary tone information is integrated with other information, e.g. expectations based on prior lexical, syntactic, and prosodic information

[1] Beckman, M. E., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S. Jun (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing* (pp. 9-54). Oxford: Oxford University Press.
[2] Dennison, H. Y., & Schafer, A. J. (2010). Online construction of implicature through contrastive prosody. In *Proceedings from Speech Prosody 2010*. Chicago, Illinois, USA: International Speech Communication Association.
[3] Heeren, W. F. L., Bibyk, S. A., Gunlogson, C., & Tanenhaus, M. K. (2015). Asking or telling — Real-time processing of prosodically distinguished questions and statements. *Language and Speech, 58*(4), 474-501.