

# Prosodic Disambiguation in Ambiguous and Unambiguous Situations

Amy J. Schafer, UCLA  
Paul Warren, Victoria University of Wellington  
Shari R. Speer, University of Kansas  
S. David White, Ohio State University  
Shari Sokol, University of Kansas

## 1 Introduction

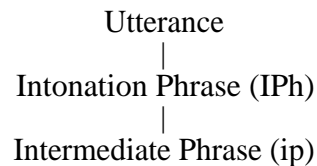
- What kind of variation is there in the prosodic structures that untrained speakers produce, and what accounts for this variation?
- In particular, do naïve, untrained speakers use prosody to disambiguate syntactic structure even when the sentence is otherwise disambiguated, or does the degree of prosodic disambiguation vary with the situational need to disambiguate?
  - Is prosodic disambiguation of syntax more likely in directive utterances than in utterances that confirm previous information?
  - Is prosodic disambiguation more likely when the pragmatic context is ambiguous about which of two syntactic structures is correct than when it includes biasing or disambiguating information?
  - Is prosodic disambiguation more likely when the two syntactic structures for an ambiguous string can each be associated with referents in the discourse context than when only one of the syntactic structures is supported by discourse referents?

We will present results from a production study and a comprehension study that show that untrained speakers generally produce distinct prosodic structures for the two interpretations of the PP-attachment ambiguity given in (1).

- (1) I want to change the position of the square with the triangle.
- a. High attachment: a triangle is used as an instrument to change the position of the square.
  - b. Low attachment: a combined square + triangle piece is moved.

We further show that the degree of prosodic disambiguation provided by these untrained speakers does not appear to vary in response to the level of ambiguity in the situation.

## 2 Prosodic Structure of Mainstream American English



- Each utterance is composed of one or more intonation phrases (IPhs).
- Each intonation phrase is composed of one or more intermediate phrases (ips).
- Each intermediate phrase must contain at least one pitch accent.
- Prosodic phrase edges are marked by:
  - Lengthening of the final syllable.
  - Silent intervals between prosodic phrases, particularly for intonation phrases.
  - Edge tones:           IPhs: boundary tones (H%, L%).  
                                  ips: phrase accents (H-, L-).
  - Segmental effects, pitch range reset.
- Pitch accents can be High (H\*), Low (L\*), or Bitonal (e.g., L+H\*, L\*+H).

## 3 Previous Research

Many comprehension studies have argued that prosody can resolve syntactic ambiguities, including PP-attachment ambiguities (e.g., Warren, 1985; Price et al, 1991; Schafer, 1997). These studies have generally found that High PP attachments are associated with stronger prosodic boundaries preceding the PP. However, this research has relied on productions from trained speakers.

Previous research using untrained speakers has suggested that the productions of trained speakers may be unrepresentative. These studies have found that whether speakers disambiguate syntactic structure depends on the level of ambiguity in the discourse situation. Untrained speakers disambiguate through prosody when two or more syntactic structures are plausible in a discourse situation, but do not prosodically disambiguate in unambiguous situations unless instructed to (Allbritton, McKoon & Ratcliff, 1996; Straub, 1997). However, this research has relied on oral reading tasks.

## 4 Concerns with Reading Tasks

- Readers may have different pragmatic goals from people producing spontaneous speech.
- Reading and speaking have different production demands. For example, readers are provided a word order and orthographic representations of words; speakers must generate a surface structure from a preverbal message.
- Disambiguating contexts might not be strong enough to establish the correct syntactic structure.
- Readers might not fully parse the material before producing it, or the time at which (parts of) semantic, syntactic, and prosodic structures are created may differ for read speech and spontaneously produced speech.
- Prosodic structures have been shown to differ for read speech and spontaneous speech (e.g., Ayers, 1994).

## 5 Experiment 1: Sentence Production

### 5.1 Outline of Task

- Cooperative game task. Naïve pairs of speakers played a boardgame, in which they had to work together to gain points.
- They used scripted sentences to negotiate gamepieces from start positions to goal positions.
- More realistic discourse setting than reading tasks:
  - Speakers had clear communicative goals.
  - They took turns speaking.
  - They interacted with each other by asking for further information or correcting misunderstandings.
- Speakers were seated on opposite sides of a divider in a sound-proofed room, wearing head-mounted microphones
- They played multiple rounds of the game, switching roles and gameboards between rounds, allowing us to collect matched pairs of sentences and multiple tokens of each syntactic structure of interest.

### 5.2 Roles of the Two Players

#### Player 1 ("Driver")

- Directed the Slider to move a particular piece.
- Knew locations of goal positions, but not of bonuses (cookies) and hazards (ravenous goats).
- Could correct the Slider if s/he moved incorrectly.
- Marked intended moves with a placemaker, before uttering instructions.

#### Player 2 ("Slider")

- Confirmed completed moves with the Driver.
- Choose the direction to move in.
- Knew locations of bonuses and hazards, but not goals.
- Asked the Driver for more detailed information when necessary.
- Marked intended moves with a placemaker, before sliding pieces and confirming the move.

### 5.3 Sample Dialogue

- Driver:** I want to change the position of the triangle.  
**Slider:** Which triangle do you want to change the position of this time?  
**Driver:** The blue one. When that moves it should land in a good spot.  
**Slider:** I am able to confirm the move of the triangle. It has moved 3 spaces up.  
**Driver:** Whoops, go back, there's another direction I want you to go in.  
**Slider:** OK, it's back where it was before.

## 5.4 Further Game Information

- The square had to be pushed by another gamepiece. It could not move on its own. Moving the square thus required the use of a High-attached PP sentence.
- Gamepieces moved by sliding in a straight path until they reached the end of the path or were blocked by another gamepiece. They slid over cookies and goats, picking up the cookies for future use, and losing points from the goats, if not protected by a cookie.
- Players had to plan moves carefully to succeed at the game. For example, they might have had to position a first piece somewhere to prevent a second piece from sliding too far. They also tried to pick up cookies before they were forced to encounter ravenous goats.
- Subjects had a short list of printed sentences. They were free to look at the list as they selected and uttered sentences, although they seemed to learn the sentences early in the game.
- Subjects were not informed of the syntactic ambiguities until the completion of the experiment.
- Subjects were not told to prosodically disambiguate, nor told to do anything special in their pronunciation of the sentences. The experimenters were careful to never produce the sentences in front of the subjects.
- Subjects played for about two hours. They first saw a demonstration board, on which an experimenter demonstrated moves while the subjects uttered an ordered subset of sentences drawn from the larger sentence list. Next they completed a simple practice round, using the regular, unordered list of sentences. They then played on experimental boards. Subjects played as many rounds as they could complete in the two hour session, switching roles and gameboards between rounds.

## 5.5 Materials

### High Attachment: A triangle is used as an instrument to push a square

- Driver:** (2) I want to change [the position of the square] [with the **triangle**].  
(3) I want to change [the position of the square] [with the **cylinder**].

- Slider:** (4) I am able to confirm the move [of the square] [with the **triangle**].  
(5) I am able to confirm the move [of the square] [with the **cylinder**].

### Low Attachment: A combined square + triangle piece is moved

- Driver:** (6) I want to change the position of [the square with the **triangle**].

- Slider:** (7) I am able to confirm the move of [the square with the **triangle**].

## 5.6 Experimental Conditions

1. Syntactic ambiguity between **High** and **Low attachment**.
2. Directive vs. Confirmatory comparison:
  - **Driver** identifies which piece to move.
  - **Slider** confirms moves requested by the Driver.
3. Three levels of pragmatic ambiguity, determined by the gameboard configuration:
  - **Unambiguous**: Gameboard configuration allows only one interpretation.
  - **Biased**: Both moves are possible, but one is favored. E.g., a triangle has just been moved adjacent to a square.
  - **Ambiguous**: Both moves are possible; neither is strongly favored.
4. Two referential contexts, determined by the set of gamepieces:
  - **Triangle** utterances: Both interpretations are supported by the gamepieces.
  - **Cylinder** utterances: Only the High attachment interpretation is supported by the gamepieces. A combined square + cylinder gamepiece is never introduced, in any of the games.

## 5.7 Predictions

1. If prosodic structure partially reflects syntactic structure, prosody should differ for High- vs. Low-attached sentences. The prosodic boundary before the PP is predicted to be stronger for High-attachment sentences than for Low-attachment sentences.
2. If speakers disambiguate only when the discourse situation requires it, differences between High- and Low-attached sentences should be modulated by the discourse situation:
  - The directive Driver should disambiguate more than the confirmatory Slider.
  - There should be more disambiguation in the Ambiguous condition than in the Biased condition, and more in the Biased condition than the Unambiguous condition.
  - There should be more disambiguation with *triangle* utterances than with *cylinder* utterances.
3. If speakers tend to disambiguate regardless of the discourse situation:
  - There should be reliable differences between the High- and Low-attached sentences for speakers in the Driver role and speakers in the Slider role.
  - There should be reliable differences between the High- and Low-attached sentences regardless of whether the gameboard configuration supports only one move, is biased, or is ambiguous.
  - The prosody for *cylinder* sentences should not differ from the prosody for High-attached *triangle* sentences, and both of these should differ from Low-attached *triangle* sentences.

## 5.8 Method

14 speakers of Midwestern American English from the University of Kansas participated. In this talk, we report results only from subsets of speakers that produced sufficient numbers of fluent tokens in the relevant conditions to allow reasonable comparisons. Results from the complete set of subjects closely mirror the results from the subsets, and results from contrasts shown (e.g., word + pause duration for 'square') are representative of contrasts not shown (e.g., word duration for 'square').

Duration and fundamental frequency measurements were taken for critical regions of the sentences. The prosodic and intonational structures of the sentences were transcribed, following the standard ToBI (Tone and Break Indices) transcription system for English. Two ToBI-trained transcribers, who were not informed of the intended syntactic structure, worked to consensus.

## 5.9 Results

The following results will be shown:

- Duration of 'square' + following pause for Drivers and Sliders.
- Variation in contours across speakers. High attachment tokens from 3 speakers.
- Variation in contours within a speaker. Three Low attachment tokens from one speaker.
- Durations for High-attached sentences in ambiguous, biased, and unambiguous situations.
- Durations for Low-attached sentences in ambiguous, biased, and unambiguous situations.
- Durations for *triangle* and *cylinder* sentences, Driver subjects.
- Durations for *triangle* and *cylinder* sentences, Slider subjects.
- Summaries of prosodic boundary patterns for *triangle* and *cylinder* sentences.

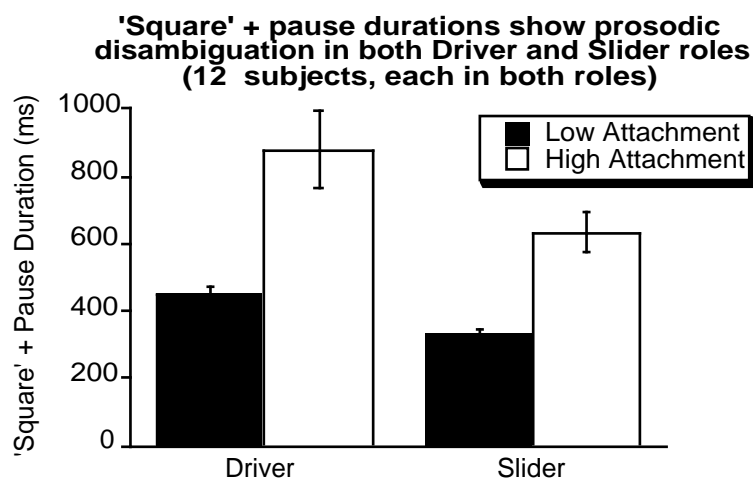
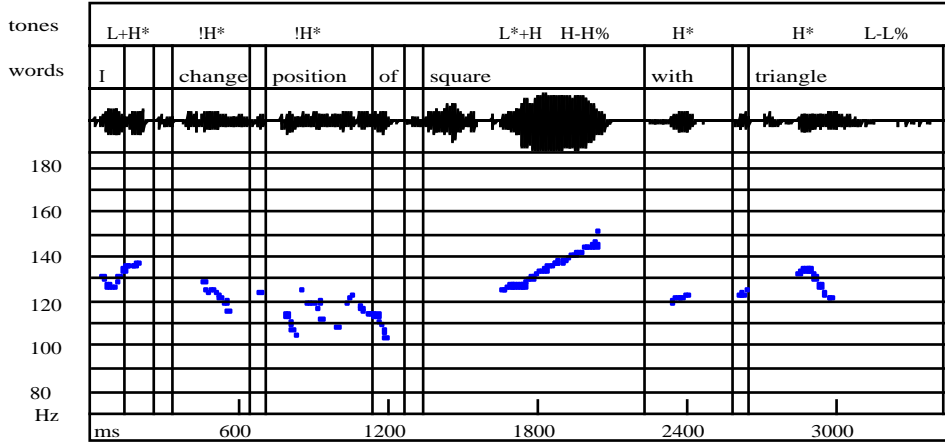
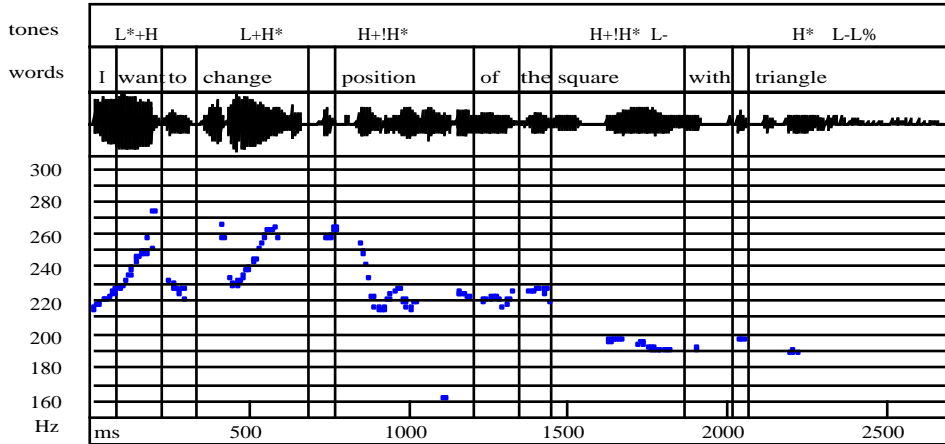


Figure 1. Duration of 'square' + following pause for Drivers and Sliders. Significant main effect of syntax. Significant main effect of role. Interaction was not significant. 12 subjects, each in both roles. Drivers: 71 High- and 95 Low-attached tokens. Sliders: 54 High- and 95 Low-attached tokens.

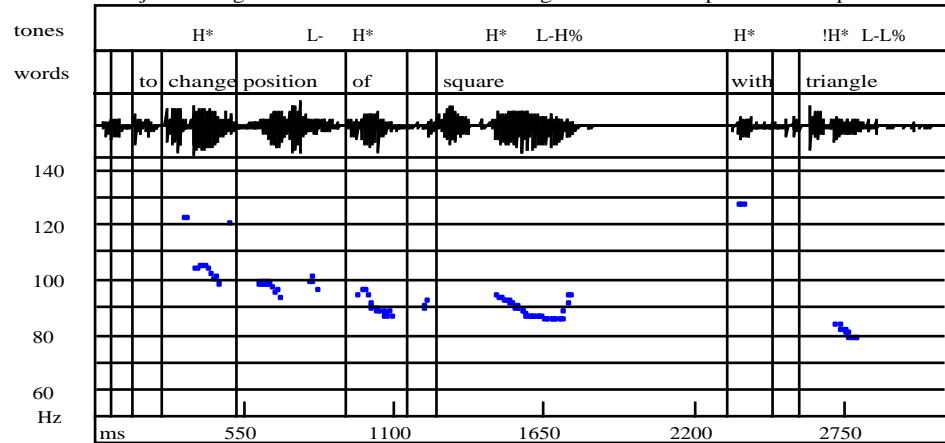
Subject 1. High attachment. 79% correct categorization in comprehension experiment.



Subject 2. High attachment. 74% correct categorization in comprehension experiment.

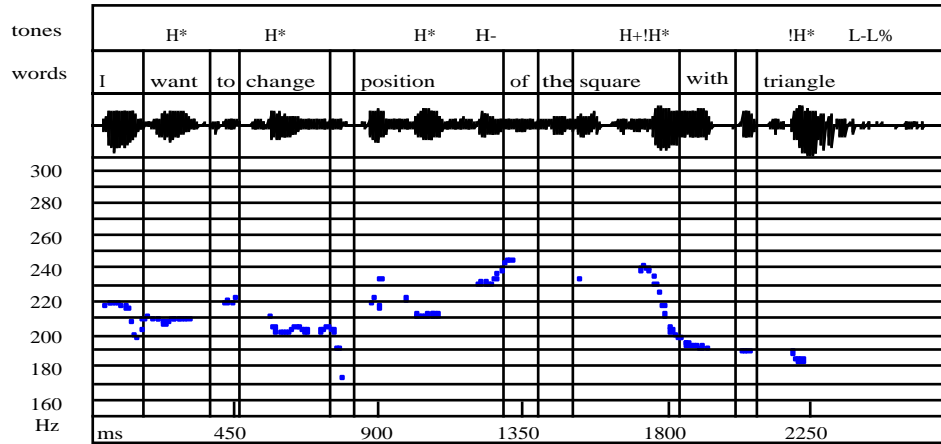


Subject 3. High attachment. 68% correct categorization in comprehension experiment.

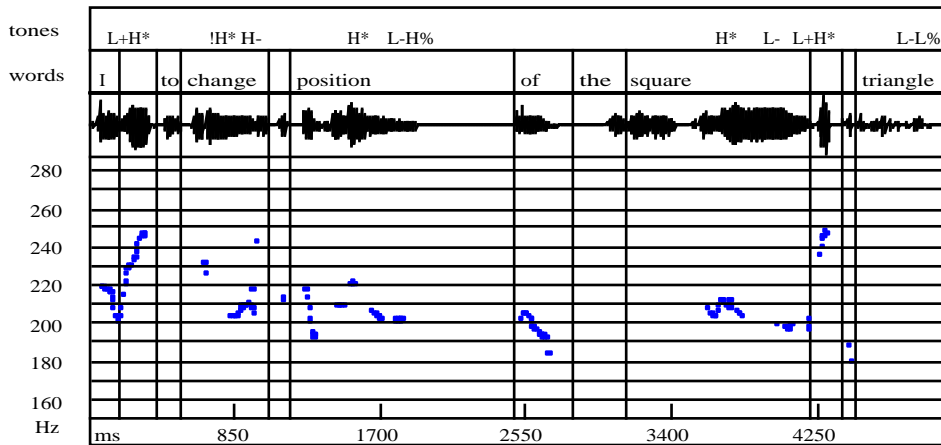


Figures 2 - 4. Variation in contours across speakers. High attachment tokens from 3 speakers.

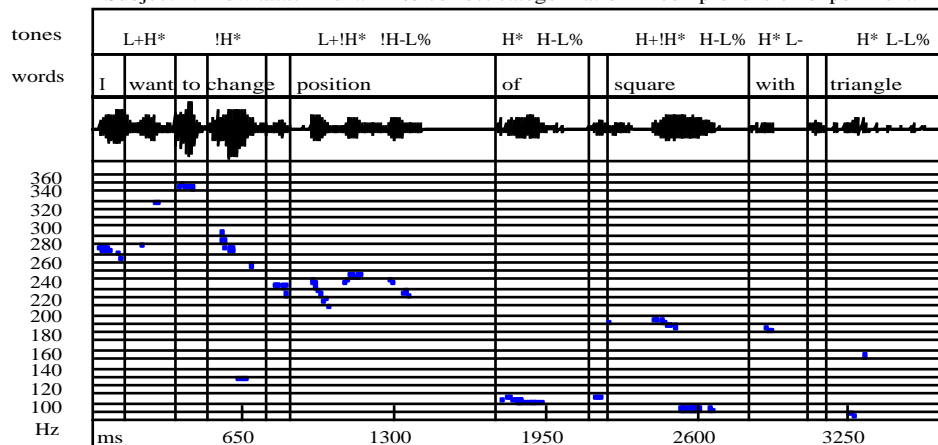
Subject 2. Low attachment. 89% correct categorization in comprehension experiment.



Subject 2. Low attachment. 47% correct categorization in comprehension experiment.



Subject 2. Low attachment. 21% correct categorization in comprehension experiment.



Figures 5 - 7. Variation within a speaker. Three Low attachment tokens from one speaker.



**Prosodic disambiguation of syntax for High-attached PPs in three discourse situations (8 Driver subjects)**

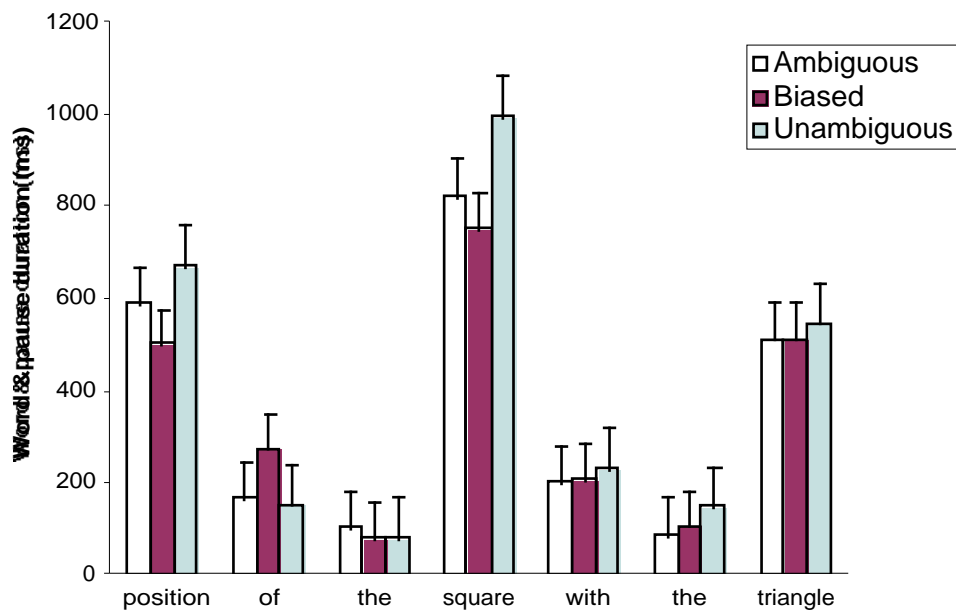


Figure 8. Durations for High-attached sentences in ambiguous, biased, and unambiguous situations. 8 Driver subjects with at least one fluent token in each situation. 18 ambiguous, 22 biased, and 20 unambiguous tokens.

**Prosodic disambiguation of syntax for Low-attached PPs in three discourse contexts (2 Driver subjects)**

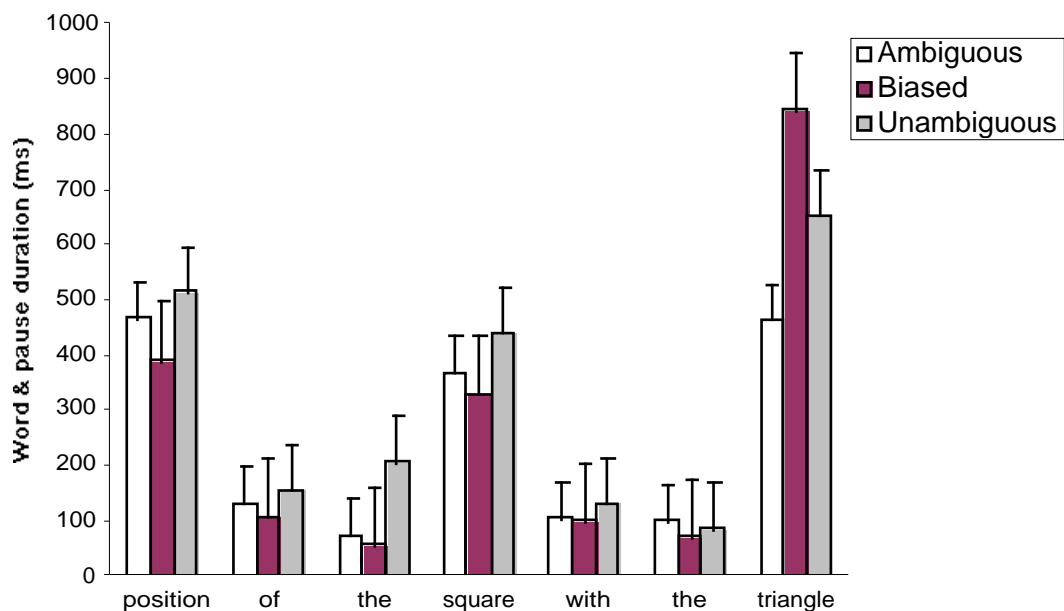


Figure 9. Durations for Low-attached sentences in ambiguous, biased, and unambiguous situations. 2 Driver subjects with at least one fluent token in each situation. 3 ambiguous, 10 biased, and 9 unambiguous tokens.

**Prosody disambiguates syntax regardless  
of interpretations available from discourse context  
(12 Driver subjects)**

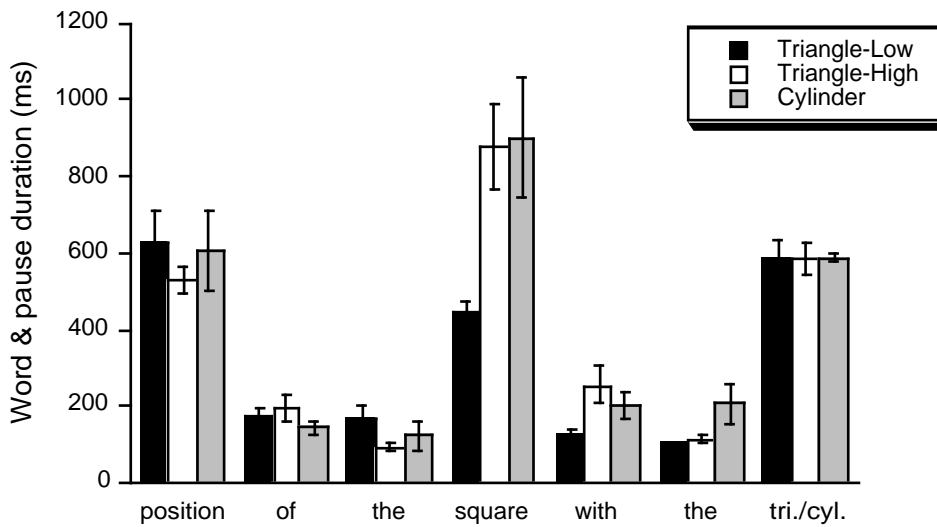


Figure 10. Durations for *triangle* and *cylinder* sentences. 12 Driver subjects with at least one fluent token in each situation. 71 High-attached *triangle* tokens, 95 Low-attached *triangle* tokens, 21 High-attached *cylinder* tokens.

**Prosody disambiguates syntax regardless  
of interpretations available from discourse context  
(12 Slider subjects)**

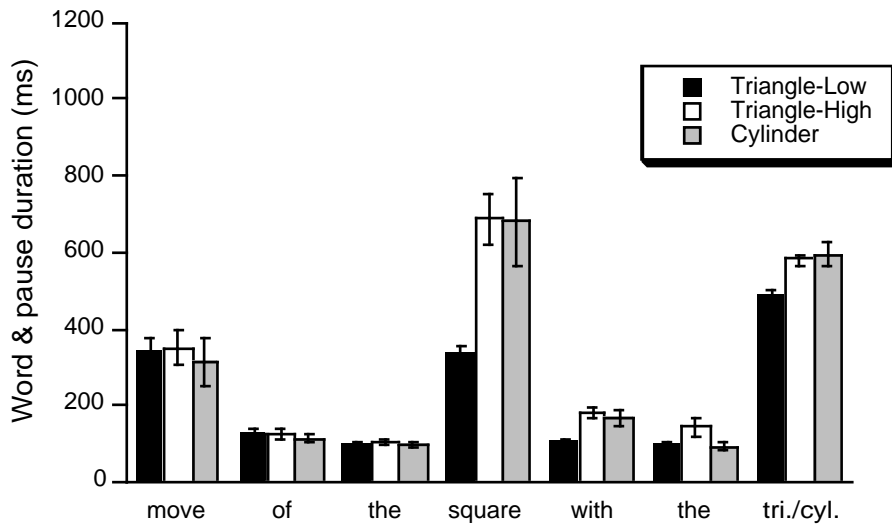


Figure 11. Durations for *triangle* and *cylinder* sentences. 12 Slider subjects with at least one fluent token in each situation. 55 High-attached *triangle* tokens, 96 Low-attached *triangle* tokens, 16 High-attached *cylinder* tokens.

**Comparable prosodic disambiguation when context provides multiple vs. single interpretations (4 Driver & 3 Slider subjects, 115 utterances)**

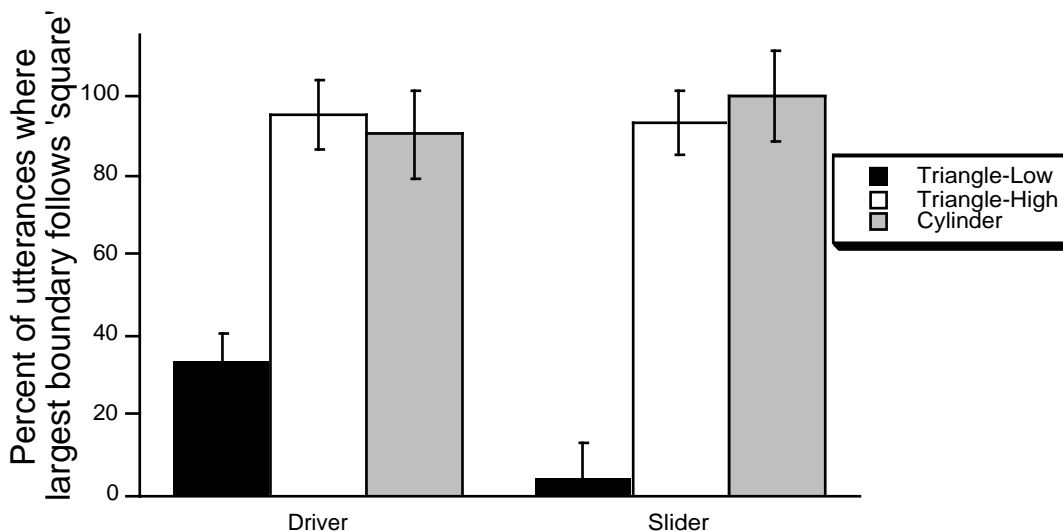


Figure 12. Percentages of *triangle* and *cylinder* sentences in which the largest prosodic boundary immediately follows 'square'. 4 Drivers subjects and 3 Slider subjects, each with at least one fluent token in each situation. Drivers: 20 High-attached *triangle* tokens, 36 Low-attached *triangle* tokens, 10 High-attached *cylinder* tokens. Sliders: 15 High-attached *triangle* tokens, 27 Low-attached *triangle* tokens, 7 High-attached *cylinder* tokens.

**Comparable prosodic disambiguation when context provides multiple vs. single interpretations (4 Driver and 3 Slider subjects, 115 utterances)**

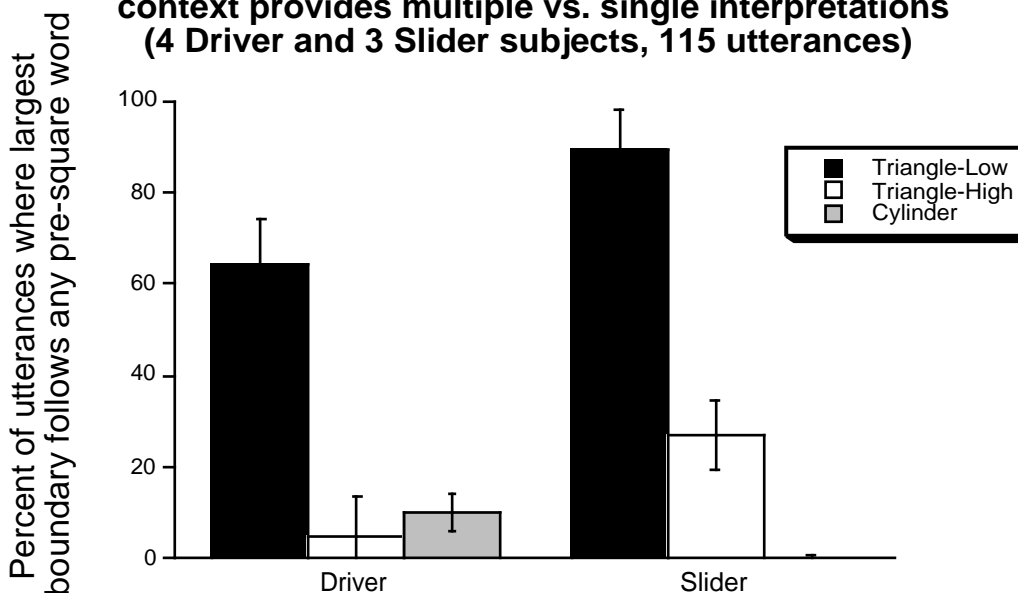


Figure 13. Percentages of *triangle* and *cylinder* sentences in which the largest prosodic boundary follows any word preceding 'square'. Same tokens as in Figure 12.

## 6 Comprehension Experiment

### 6.1 Method

- Auditory presentation of full sentences:
  - (8) I want to change the position of the square with the triangle.
- Forced-choice paraphrase selection.
- 19 listeners, all native speakers of Midwestern American English from the University of Kansas.

### 6.2 Predictions

- If speakers produce prosodic structures which partially reflect syntactic structure, the percentages of correct categorization should be above chance for both High- and Low-attached sentences.
- If speakers provide prosodic disambiguation only when the situation requires it, the percentages of correct categorization should be higher for tokens produced in the Ambiguous condition than in the Biased condition, and higher for tokens produced in the Biased condition than in the Unambiguous condition.
- If speakers provide prosodic disambiguation regardless of the situational ambiguity, the percentages of correct categorization should not differ systematically across the levels of situational ambiguity.

### 6.3 Results

The results are given in Figure 14.

- Results from a subset of tokens are presented: all fluent tokens from all speakers who produced at least one fluent token in each of the three conditions. Results from the complete set of tokens closely mirror the results of this subset.
- High-attachment sentences: results for all fluent tokens from 8 speakers (18 ambiguous, 22 biased, and 20 unambiguous tokens).
- Low-attachment sentences: results for all fluent tokens from 2 speakers (3 ambiguous, 10 biased, and 9 unambiguous tokens).

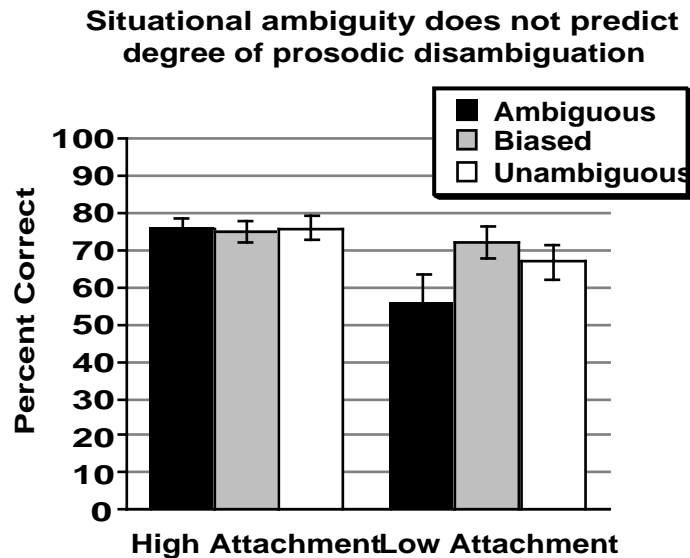


Figure 14. Percentages of correct categorization of tokens as High- or Low-attached sentences.

## 7 General Discussion

- Consistent with previous findings from this task, speakers reliably disambiguated the PP-attachment sentences.
- Speakers varied in the prosodic structures they produced, and in how strongly they disambiguated, but this variation is not well explained by the level of situational ambiguity.
- Some factors that we think may affect the prosodic resolution of syntactic ambiguity:
  - **Information structure:** Certain focal structures should lead to a prosodic structure which is ambiguous or misleading for the syntactic structure. E.g., in English, contrastive focus can result in the addition of a prosodic boundary, so a Low-attached PP that is contrastively focused may be more syntactically ambiguous than one which is not contrastively focused.
  - **Sentence type:** Some syntactic ambiguities should be disambiguated more often than others. For example, ambiguously attached arguments should be disambiguated more frequently than ambiguously attached adjuncts in English, since adjuncts are more likely to be set off in a separate prosodic phrase.
  - **Sentence token:** Within a syntactic structure, sentences with longer words and phrases may differ from ones with shorter words and phrases.
  - **Speech rate/style:** If a slower rate of speech induces more prosodic structure, patterns of disambiguation may differ across speech rates. Styles expected to be clearer (such as slower speech to children) may in fact lead to less prosodic disambiguation in certain cases. For example, in our third Low-attachment example, the speaker used a slow, deliberate speech style, and the token was correctly categorized only 21% of the time.
  - **Language:** Some prosodic systems may allow for more prosodic disambiguation than others. For example, a language which allows little variation in the size of prosodic phrases may show less prosodic disambiguation than a language which allows more flexibility.

## 8 Conclusions

- Consistent with previous findings from the production game task (e.g., Schafer et al., in press), speakers reliably use prosody to disambiguate PP-attachment sentences even when disambiguating information is available in the discourse:
  - Durational analyses showed prosodic boundaries consistent with the syntactic structure. High-attached PPs were marked by lengthening and silent intervals immediately preceding the PP. Low-attached PPs were less likely to have these indications of a prosodic phrase boundary.
  - Phonological analyses of the prosodic structure showed prosodic boundaries consistent with the syntactic structure. High-attached PP sentences contained prosodic phrase boundaries immediately preceding the PP. Low-attached PP sentences were less likely to contain a prosodic phrasal boundary in this location.
  - Comprehension results showed that listeners could correctly identify the syntactic structure of spoken fragments.
- None of our analyses indicated that prosodic structure was modulated to reflect the level of situational ambiguity:
  - The durational difference between High- and Low-attached PPs occurred with both directive and confirmatory utterances. There was no interaction of speaker role and word or pause durations.
  - The durations of the critical regions in the sentences did not differ for sentences produced in ambiguous, biased, and unambiguous pragmatic contexts. Sentences produced in these contexts showed similar percentages of correct categorization in a comprehension task.
  - The durations of the critical regions in High-attached PP sentences did not differ for sentences that could be mapped to only one set of discourse referents and sentences that could be mapped to two sets of discourse referents.
- The reliability of the prosodic disambiguation suggests that prosody is an important source of information for sentence comprehension in a range of discourse situations, and is not limited to situations with trained speakers or speakers who are consciously attempting to disambiguate.
- Nevertheless, there is substantial variability in prosodic form which remains to be explained.

## References

- Allbritton, D.W., McKoon, G., and Ratcliff, R. 1996. Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 22, 714-735.
- Ayers, G.M. 1994. Discourse functions of pitch range in spontaneous and read speech. *OSU Working Papers in Linguistics*, 44, 1-49.
- Price, P.J., Ostendorf, S., Shattuck-Huffnagel, S. & Fong, C. 1991. The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, 9, 2956-2970.
- Schafer, A.J. 1997. *Prosodic parsing: The role of prosody in sentence comprehension*. Unpublished doctoral dissertation, Amherst, MA: University of Massachusetts.
- Schafer, A.J., Speer, S.R., Warren, P. and White, S.D. In press. Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research*, 29, 169-182.
- Straub, K.A. 1997. *The production of prosodic cues and their role in the comprehension of syntactically ambiguous sentences*. Unpublished doctoral dissertation, Rochester, NY: University of Rochester.
- Warren, P. 1985. *The temporal organisation and perception of speech*. Unpublished doctoral dissertation, Cambridge, UK: University of Cambridge.

## Acknowledgments

This work was supported by NIH research grants DC-00029 and MH-51768, NZ/USA Cooperative Science Programme grant CSP95/01, and Marsden Fund grant VUW604. We thank Jenny Kneale and Cara Prall for assistance with running subjects and measuring data.

## Correspondence

Amy J. Schafer	schafer@ucla.edu
Paul Warren	paul.warren@vuw.ac.nz
Shari R. Speer	speer@ukans.edu
S. David White	white.773@osu.edu
Shari Sokol	ssokol@ukans.edu