The effect of intonation on visual search:
An eye-tracking study

Kiwako Ito & Shari R. Speer

Ohio State University

Intonation & Pragmatic Status

English Pitch Accent

Distinctive tonal movement associated with stressed syllable.

→ perceivable prosodic prominence

Local targets L and H (Pierrehumbert 1980)

H*, H+!H*, L*, L+H*, L*+H

(a/c current ToBI guideline)

Pitch Accent Types & Pragmatics

Pierrehumbert & Hirschburg, 1990

Compositional tune meaning

PA type	meaning
H*	Added to mutual belief space.
L*	Already part of mutual beliefs. Salience-without-predication
L*+H	Uncertainty, lack of speaker commitment
L+H*	Contrast with some alternative related item

Pitch Accent Types & Pragmatics

Pierrehumbert & Hirschburg, 1990

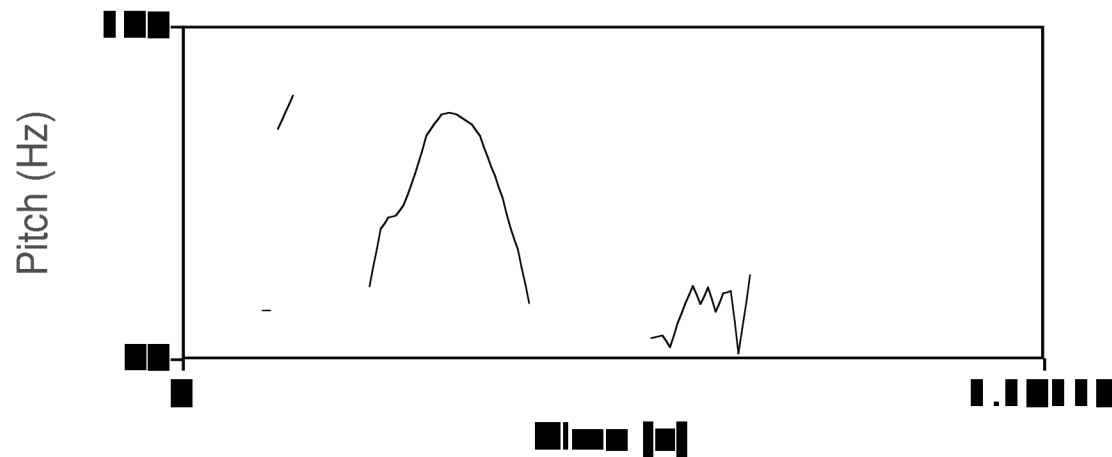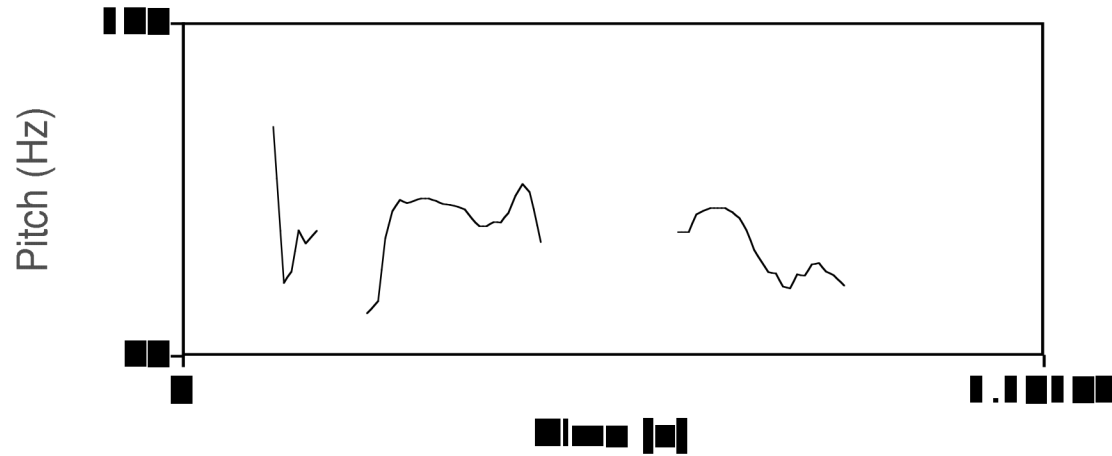Compositional tune meaning

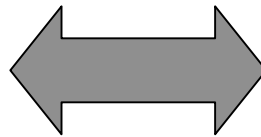| PA types	meaning
H*	**Added to mutual belief space.**
L*	Already part of mutual beliefs. Salience-without-predication
L*+H	Uncertainty, lack of speaker commitment
L+H*	**Contrast with some alternative related item**

Example pitch traces

Categorical distinction between H* and L+H*

Tonal shape

H*	→ ←	L+H*
flat or subtle rise & fall		clear rise & fall

Although listeners perform poorly in discrimination tasks along the accentual peak continuum, context identification tasks show categorical distinction between non-emphatic vs. emphatic interpretations. (Ladd & Morton, 1997)

Categorical distinction between H* and L+H*

<u>Testing tonal shape</u>

(Arvaniti & Garding, to appear)

- The height of peak and its alignment vary according to the degree of emphasis.

- Dialectal differences:

   Contrast between H* & L+H* preserves in Southern Californian but not in Minnesotan.

Unscripted speech production

- Task: Give instructions for tree decoration
- Display shows photo of an ornament and a photo of the tree, with a tag (e.g. blue ball) indicating the location of the ornament.
- Newness, Givenness & Contrastiveness manipulated by the order of mention.

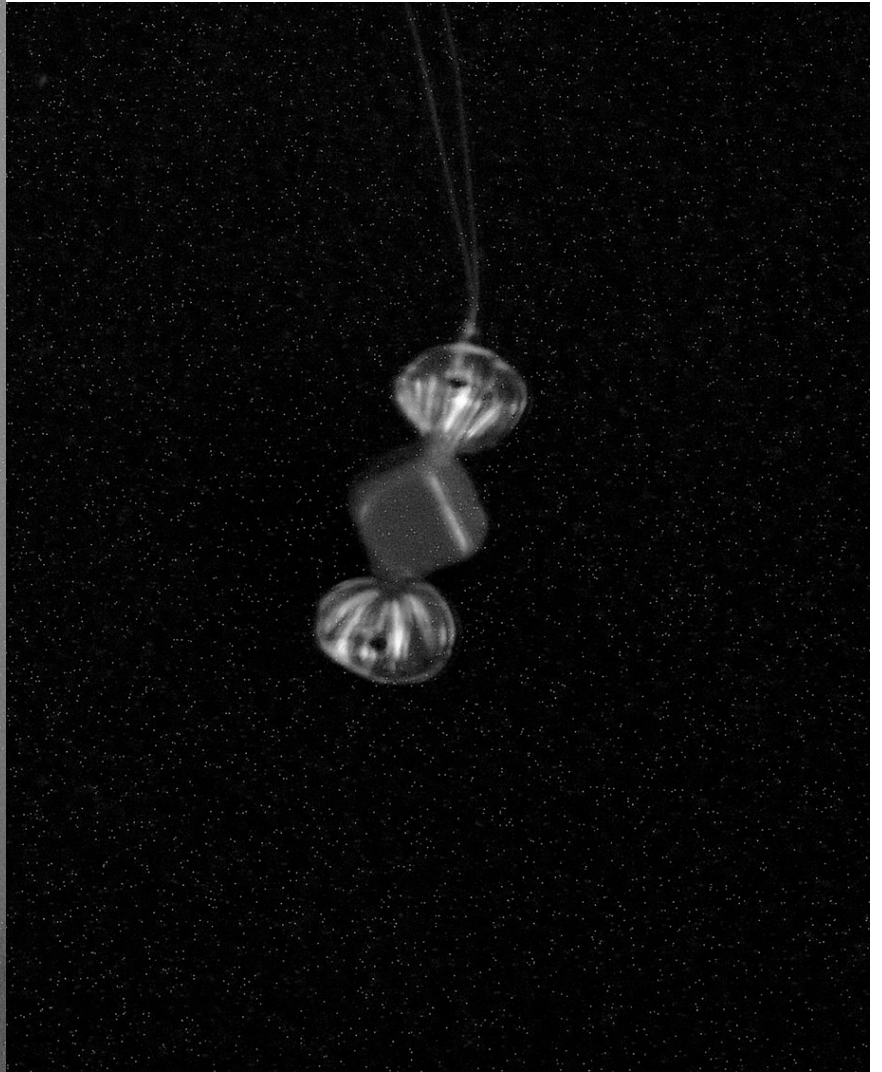 **New**:   The first mention of the discourse entity

  Given: Subsequent mentions of the discourse entity

  Contrastive: Immediate repetition of an entity.
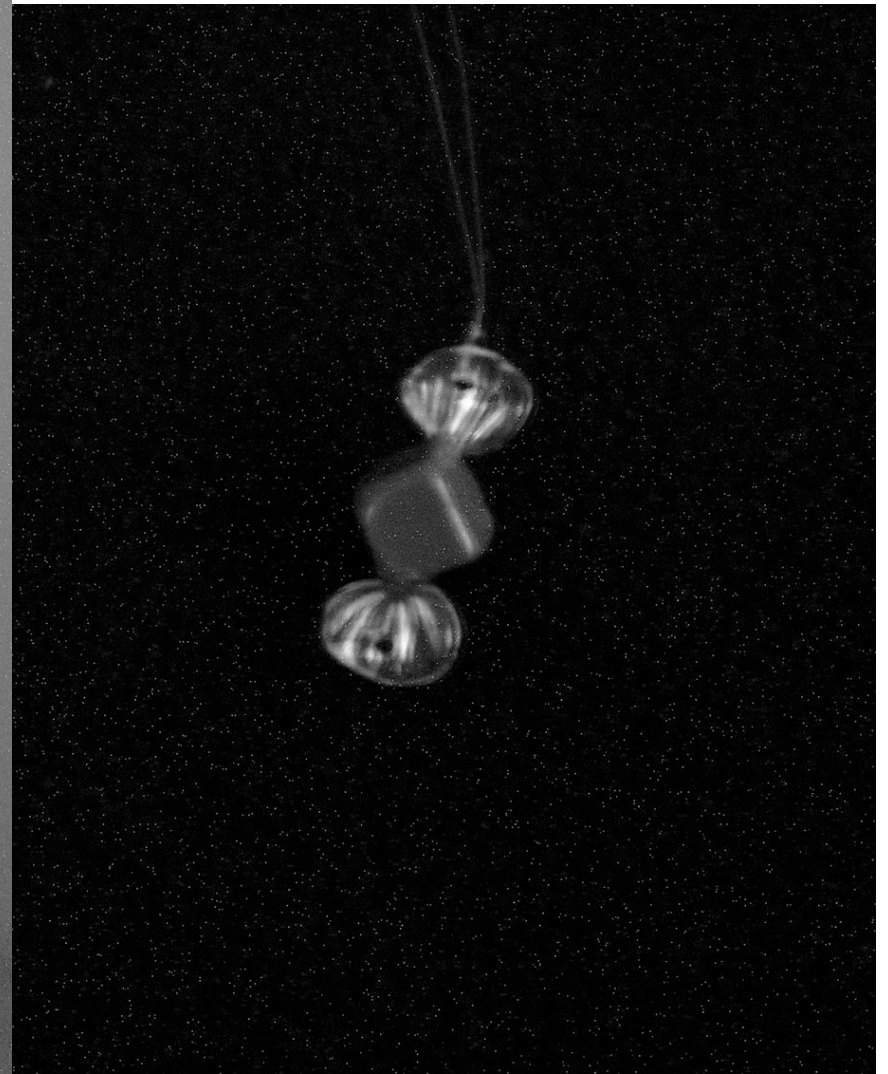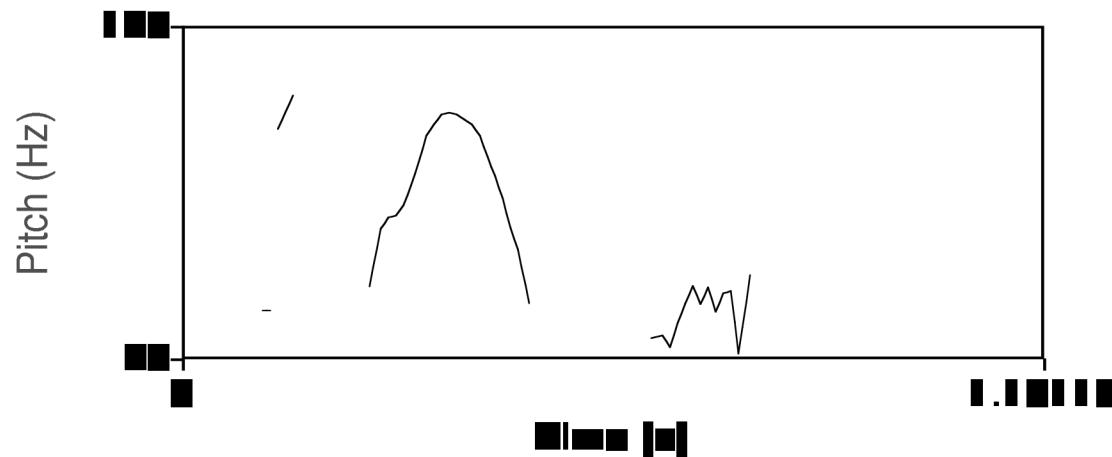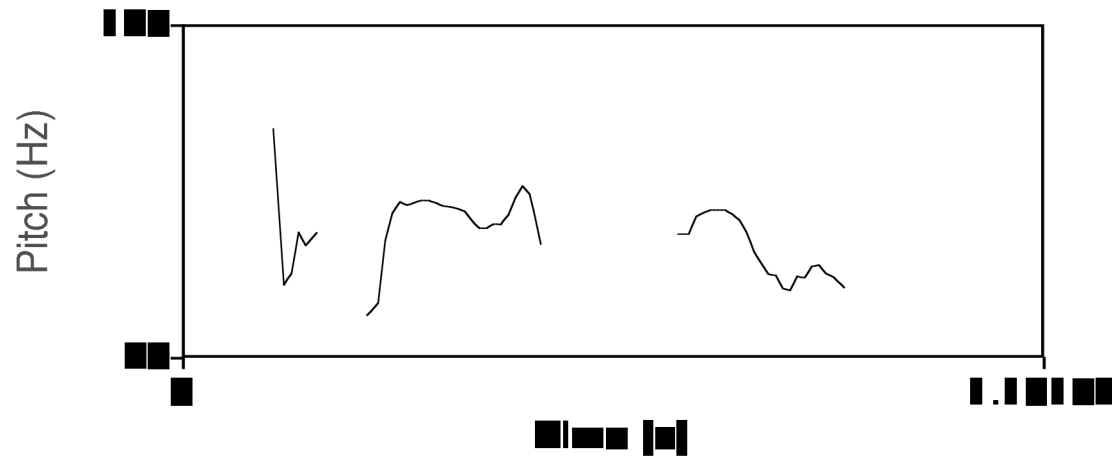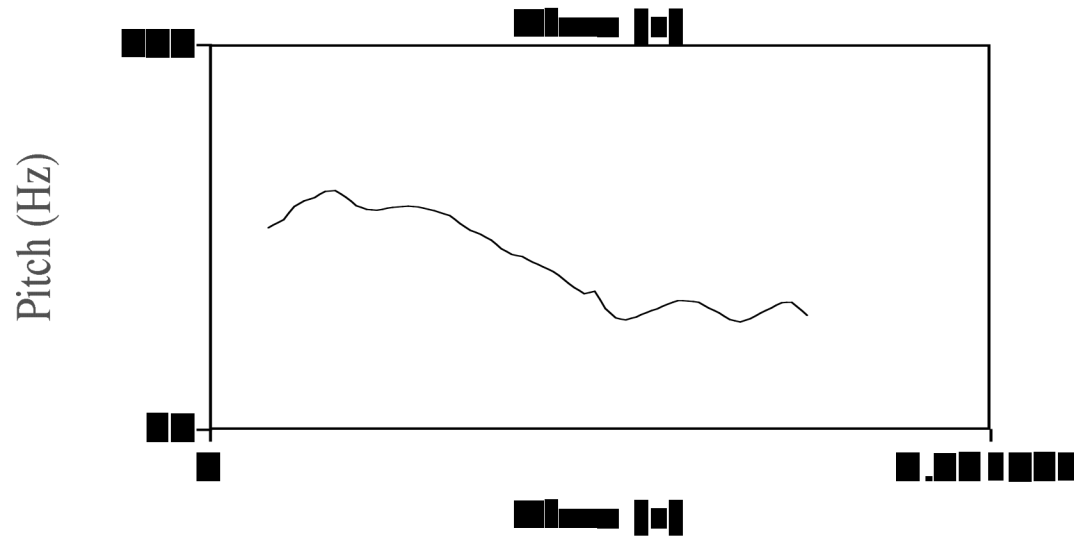
  E.g. blue ball -> blue DRUM / blue ball -> RED ball

orange
candy

blue candy 17

18 19 20 21 22

23 25 26 27

28 29 30 31

green
candy

18   19   20   21   22

23   25   26   27

28   29   30   31   32

orange candy

blue
candy 17

18 19 20 21 22

23 25 26 27

28 29 30 31

🔊 🔊

green
candy

* * * * * * * * * * * * *

18   19   20   21   22

23   25   26   27

28   29   30   31

Example pitch traces

Example pitch traces
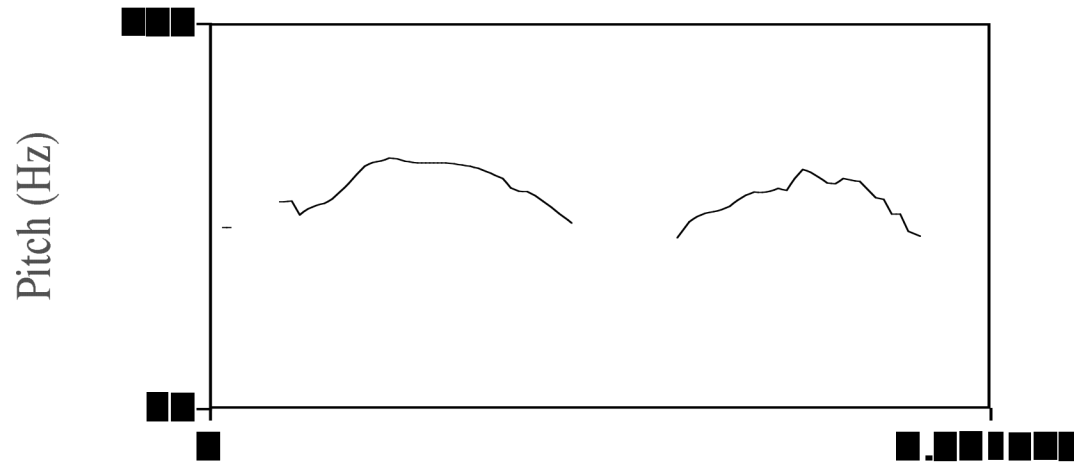
Accentuation Proportion

NEW/ GIVEN	# OF TRIALS	ADJECTIVE	NOUN
NewNew	31	**.94**	.84
NewGiven	32	**.84**	.50
GivenNew	32	**.91**	.81
GivenGiven	96	**.89**	.67

Accentuation Proportion

NEW/ GIVEN	# OF TRIALS	ADJECTIVE	NOUN
NewNew	31	.94	**.84**
NewGiven	32	.84	.50
GivenNew	32	.91	**.81**
GivenGiven	96	.89	.67

L+H* distribution

NEW/ GIVEN	PRAGMATIC STATUS	ADJECTVE		NOUN	
		# OF TRIALS	L+H*	# OF TRIALS	L+H*
New	**Contrastive**	26	**.46**	31	**.19**
	Non-contrastive	37	.03	33	.06
Given	**Contrastive**	37	**.51**	33	**.18**
	Non-contrastive	91	.00	95	.00

Is intonation informative?

Effect of Intonation on visual search

Listeners use pitch accent to single out the object referred to (Dahan et al. 2002)
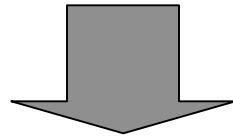
"Put the candle/candy below the triangle."

NOW, put the CANDLE above the square
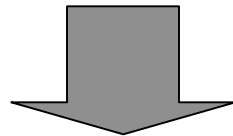NOW, put the candle ABOVE THE SQUARE.

candle → "CANDLE"   ⟹ fixation to **candy**
candle → "candy"   ⟹ fixation to **candle**

Question

Emphatic accent L+H*

⬇

Interpretation of CONTRAST

⬇

Does L+H* on a modifier (e.g. color adjective) facilitate the search for the modified object?

Experiment 1:Procedure

- Subjects followed audio instruction to decorate Christmas trees.

- Ornaments displayed on a grid with 11 cells (8 target + 3 filler)

- Eye-movements tracked by head-mounted eye-tracker (ASL e5000, magnetic head tracking) at 60Hz.

Experiment 1: Stimuli

- A trained female phonetician produced stimuli with the intended pitch accent patterns.

- Recorded 44.1KHz, 16bits

- F0 extracted with 10ms window

- All stimuli were ToBI transcribed by two native speakers of American English.

Experiment 1: Stimuli

Adjective	Noun		Accent Pattern	
New	New		H* !H*	
New	Given		H* !H*	
Given	Given		H* !H*	🔊
Given	New		H* H*	🔊
NewC	Given		L+H* no accent	🔊
Given	NewC		H* L+H*	🔊

Experiment 1: Stimuli

■ Felicitous Patterns

green onion ⟶ **ORANGE** **onion**

　　　　　　　　　L+H*　no accent

brown ball ⟶ **brown** **ANGEL**

　　　　　　　　　H*　　L+H*

■ Infelicitous Patterns

gray stocking ⟶ **brown** **STOCKING**

　　　　　　　　　H*　　　L+H*

orange candy ⟶ **ORANGE** **onion**

　　　　　　　　　L+H*　no accent

Experiment 1: Research Question

- Does felicitous intonation reliably facilitate visual search as compared to infelicitous intonation?

- Does felicitous L+H* reduce the time to locate and fixate the target?

- Are such effects general across word type and word position in the utterance?

Experiment 1: Results

Contrast on Color: [L+H* no accent] vs. [H* L+H*]

Experiment 1: Results

Contrast on Object: [H* L+H*] vs. [L+H* no accent]
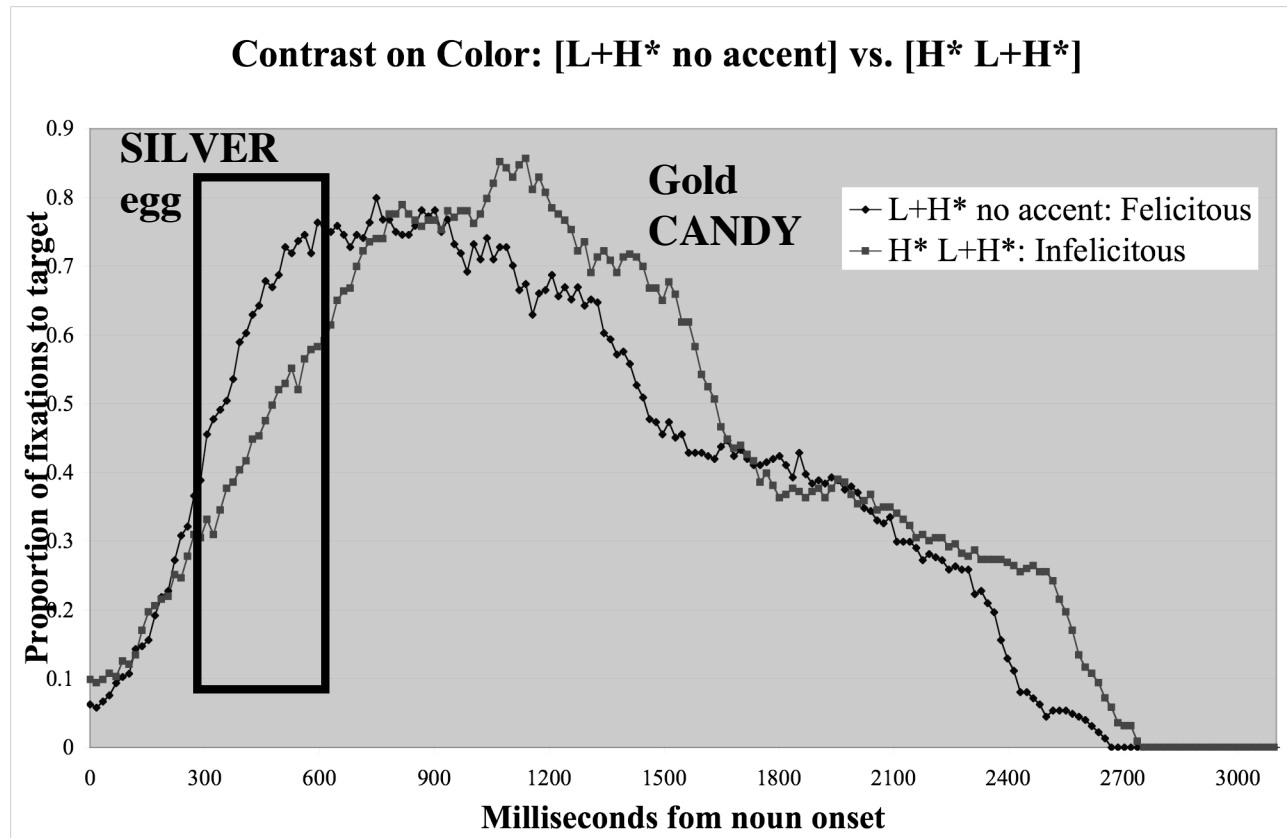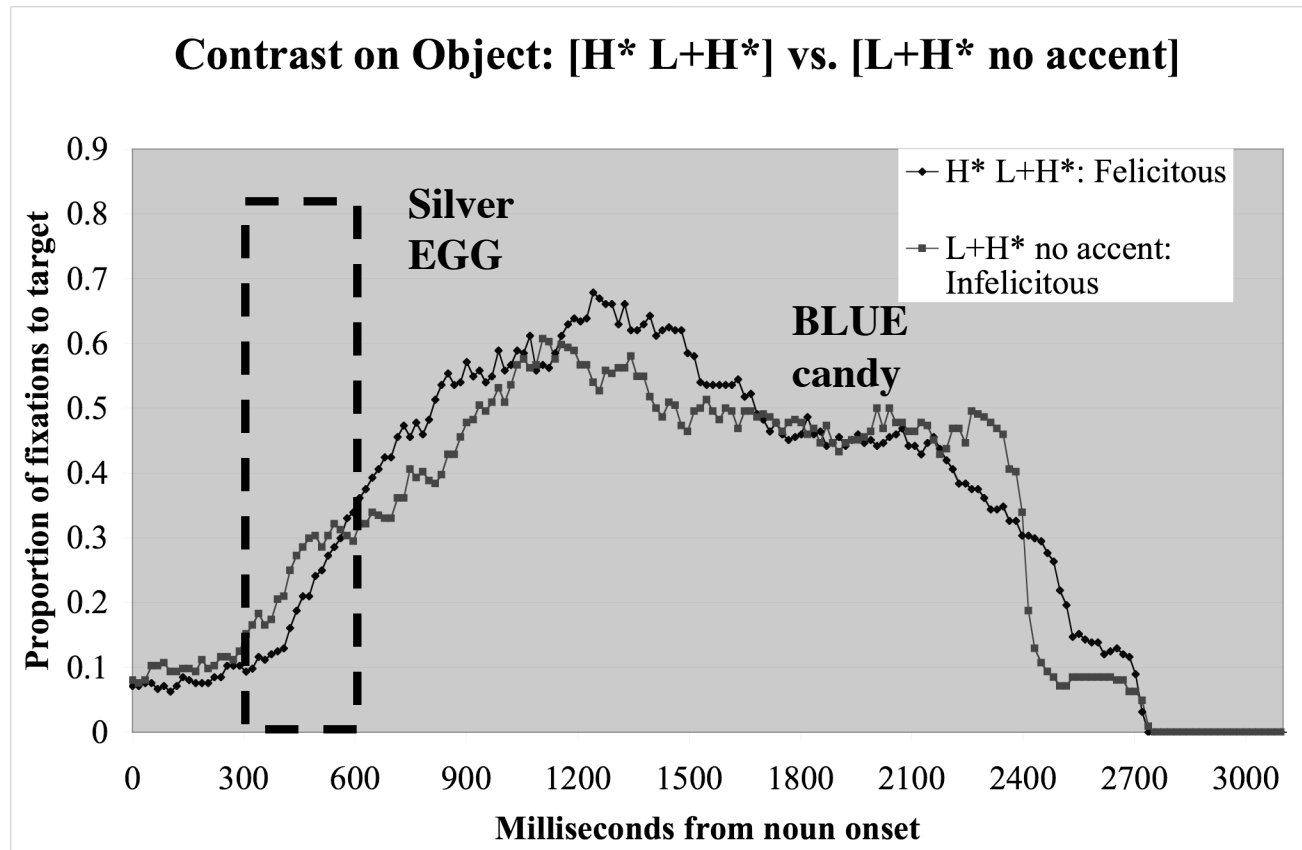
Experiment 1: Felicitous vs. Infelicitous

- Felicitous L+H* on the color term in contrastive environments (blue egg, SILver egg) facilitated visual search compared to infelicitous L+H* on non-contrastive noun (grey candy, gold CANdy).

- Felicitous L+H* on the object noun in contrastive environments (blue egg, blue CANdy) showed no advantage over infelictous L+H* on the adjective (blue egg, BLUE candy).

- Listeners 'tune' their sensitivity to contrastive accent on the basis of the visual task?

Experiment 1: Additional findings

- Overall, shorter latencies for trials with immediately repeated object nouns

    --> RECENCY/GIVENNESS

- Overall, shorter latencies for L+H* on noun vs. H* on noun trials

    --> PROMINENCE

Experiment 2: Questions

- Could the absence of L+H* on the contrastive adjective delay fixations on the repeated object?

  e.g. blue ball --> **red**/RED ball

- Does L+H* on the color adjective really lead to a strong expectation for the repetition of immediately preceding object?

  e.g. blue ball --> **RED**/red drum

- Does L+H* facilitate visual search only in felicitous contexts?

  e.g. blue ball --> **RED** ball

      blue ball --> **RED** drum
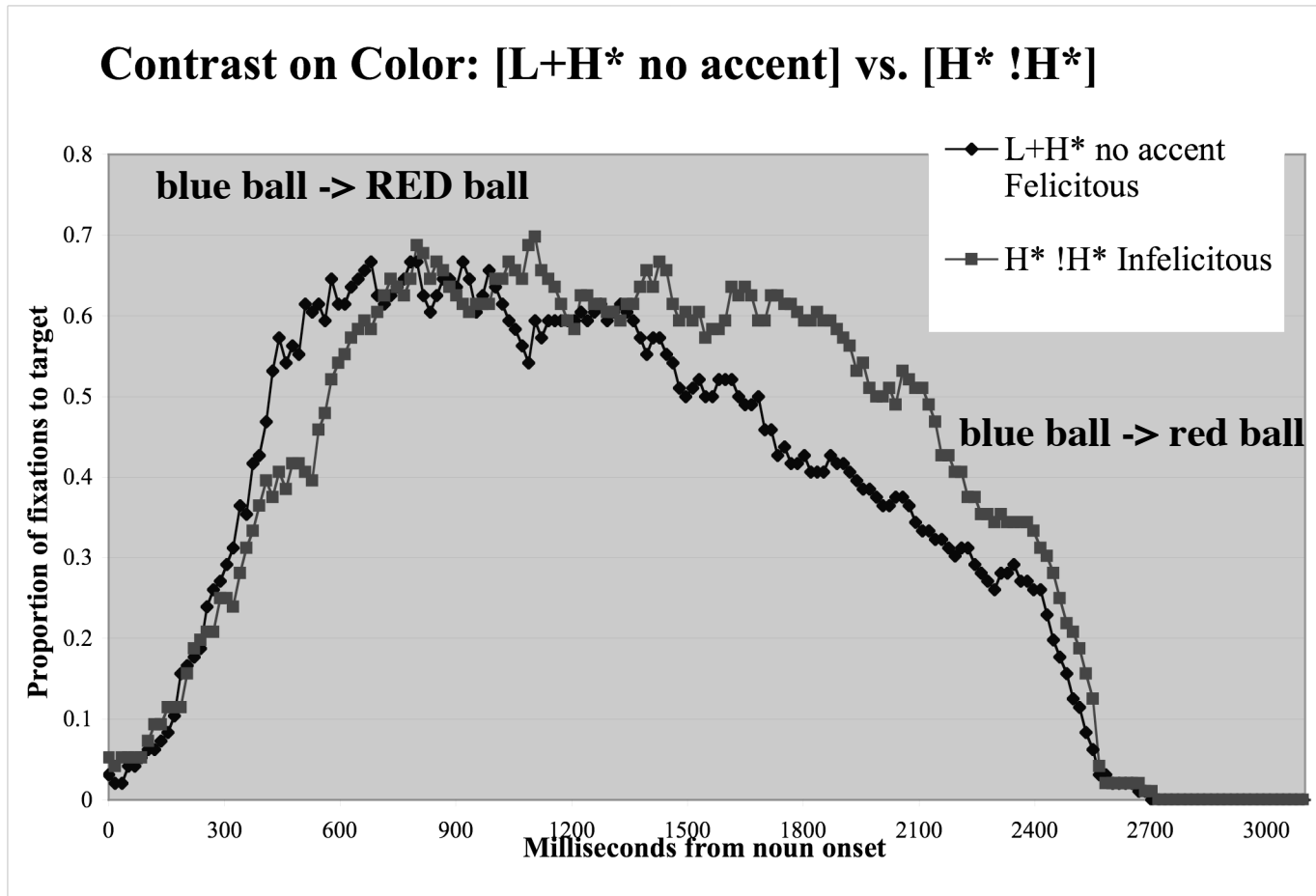
Experiment 2

Procedure Same as Exp 1

- Subjects followed audio instruction to decorate Christmas trees.

- Ornaments displayed on a grid with 8 target + 3 filler cells
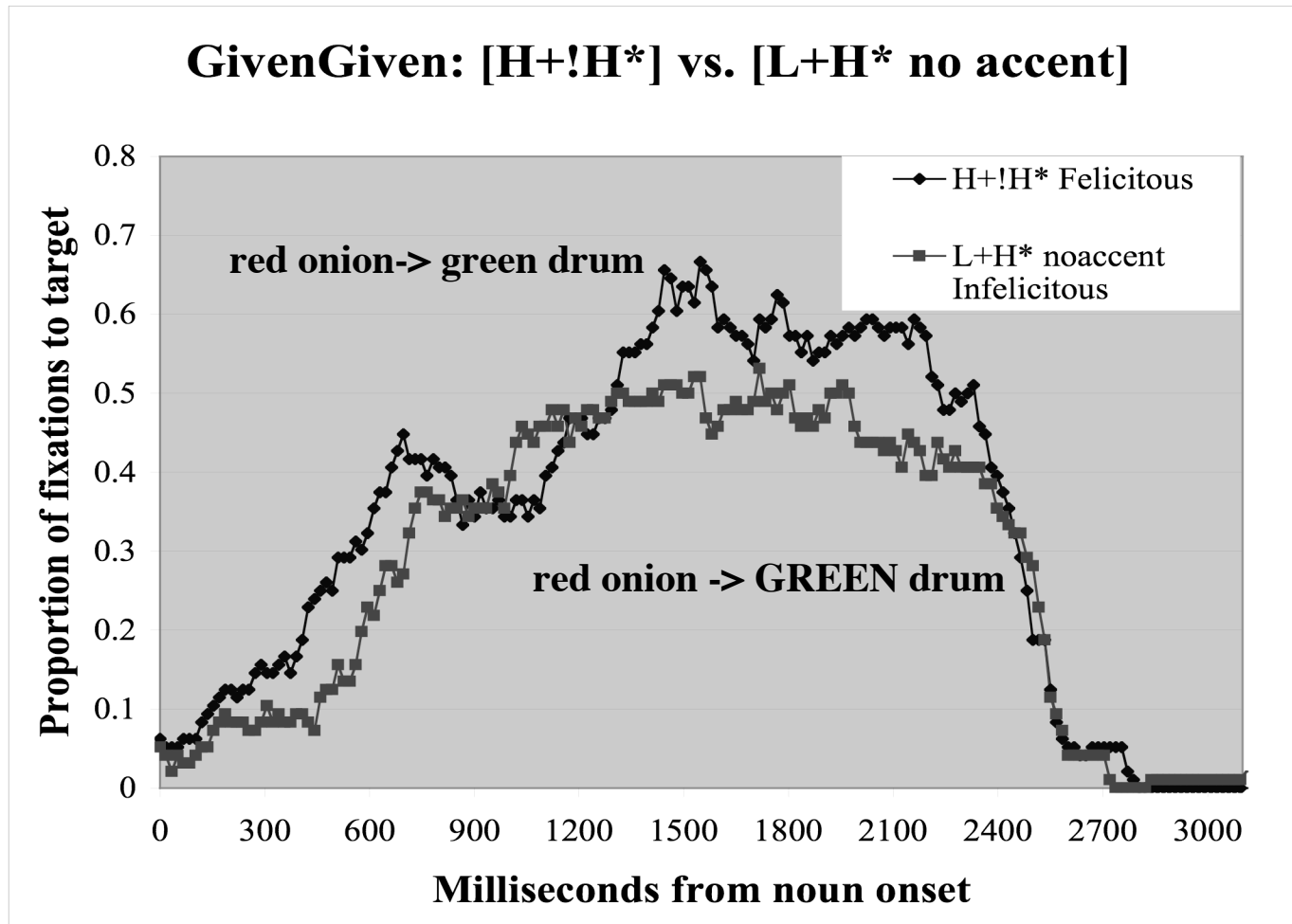
- Eye-movements monitoring at 60Hz.

Stimuli

- Prepared in the same way as Exp1

- Same speaker and ToBI transcribers

Experiment 2: Results

Contrast on Color: [L+H* no accent] vs. [H* !H*]

Experiment 2: Results

GivenGiven: [H+!H*] vs. [L+H* no accent]

Experiment 2: Results

Felicitous vs. Infelicitous L+H*

Experiment 2: Results

Infelicitous [L+H* no-accent]
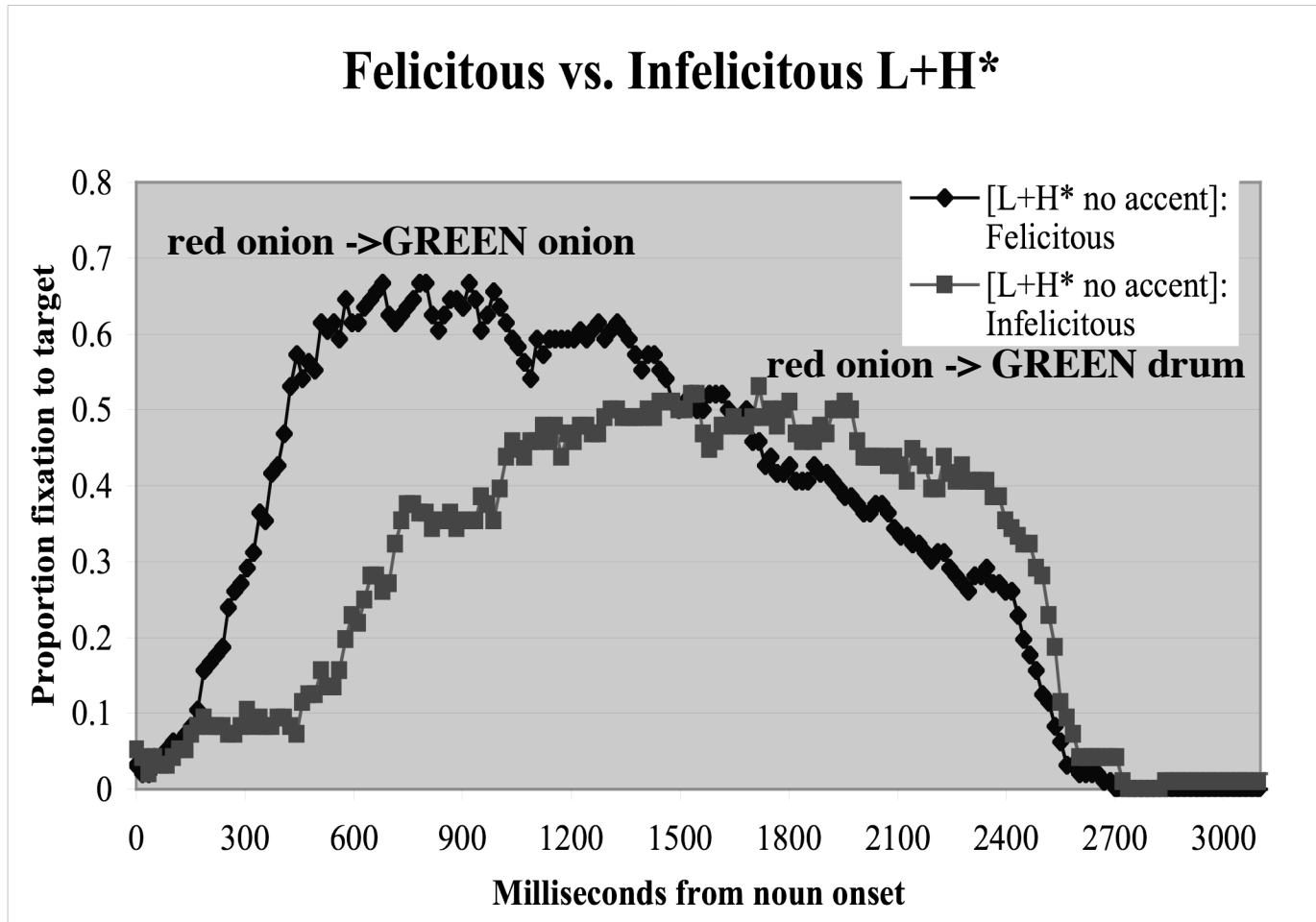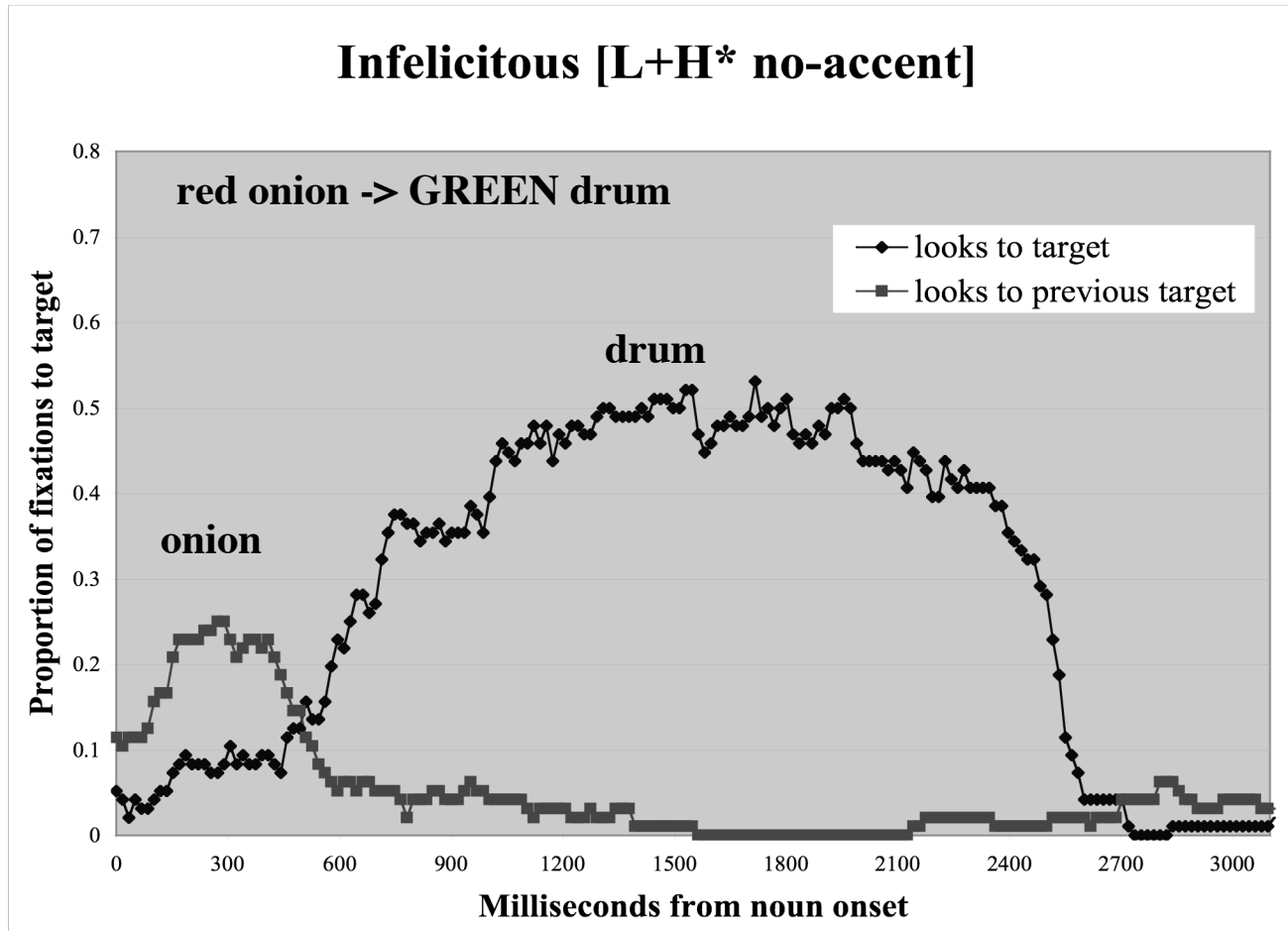
The effect of infelicitous L+H*
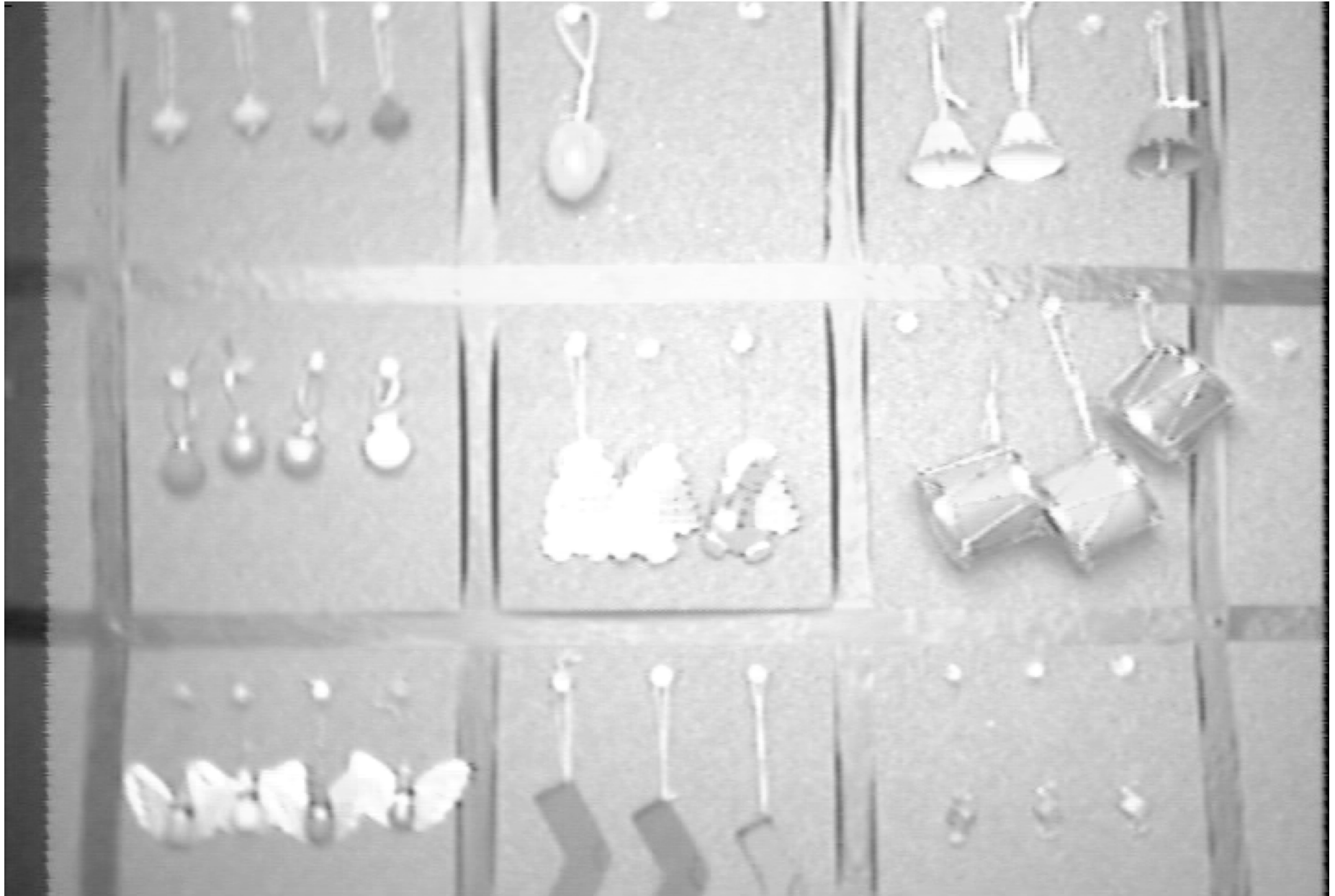
Experiment 2: Repeated Noun

- For cases with repetition of the immediately preceding noun, advantage for contrastive L+H* as compared to H* on adjective (blue ball --> RED/red ball)

- For cases without repetition of the noun, infelicitous L+H* on the adjective shows disadvantage; latency to fixate the target is delayed compared to the H* condition (red onion --> GREEN/green drum)

Experiment 2: Felicitous vs. Infelicitous L+H*

- Listeners develop expectations on the basis of L+H*. Comparison of felicitous to infelicitous L+H* on the adjective (Red onion, GREEN onion/drum) showed significant delay for the infelicitous condition.

- Analysis of fixations to target vs. previous noun in the non-repeated condition showed that infelicitous L+H* misled listeners to fixate the most recently mentioned object type.

Summary

- Behavioral differences shown for eyemovement latencies and probabilities on the basis of pitch accent type, role of the words in the the discourse structure, and structure of the visual task.

- L+H* accents led to a contrastive interpretation of the adjective in both experiments.

- These effects were not restricted to trials where the object noun was repeated.

Discussion

- Dahan et al results showed that L+H* accent led to the expectation of contrast. Contrastive accent led listeners to fixate the object that had *not* just been mentioned.
- Results here show effects on the modifier that precedes the noun; L+H* led listeners to fixate the object type that had been most recently mentioned.

This evidence suggests that L+H* evokes expectations on a contrast set that is generated based on the accented word, and the position of that word in the structure of the utterance. Here, contrastively accented adjectives lead to the expectation that the referent will be a member of the set of objects specified by an upcoming head noun.

Special Thanks to:

- NSF project 0418464
- Allison Blodgett
- Ping Bai
- Laurie Maynell
- Mary Beckman