Prosody and Intonation

9.59 / 24.905

April 14, 2005

Ted Gibson

Prosody and Intonation

- Definition of prosody
  - Intonation: pitch and boundaries
  - Timing within intonation phrases
  - Stress
- Intonational phrasing in production: Two factors: the size of the left-hand constituent, and the size of the right: The Left / Right Boundary (LRB) hypothesis
- Intonational phrasing in comprehension: The anti-attachment hypothesis.

Prosody

Prosody: The aspect of a sentence's sound that is not specific to the properties of the words in the sentence.

Prosodic features include pitch (intonation), timing and stress.

E.g.: Different words and syntactic categories, but same prosody:

Homer likes to eat.

Barney drank his beer.

Declarative contour:

Pitch falls over the sentence.

The words "to" and "his" get clipped; "eat" and "beer" get lengthened.

- Minor terminological note:

  Some researchers use the term “intonation” as a synonym for “prosody”

  Other researchers use the term “intonation” as a subset of prosody: pitch and pause information. (e.g., Ferreira, 2000)

Pitch / Intonation

Acoustic Properties

The frequency of sound vibrations produced by the vocal chord. This is also called the fundamental frequency or the f_0. (Speech Range .25 - 30,000Hz)

Perceptual Properties

The tune of an utterance. What the listener hears as “high” or “low” speech.

The perceptual properties of pitch are correlated with but not equivalent to the acoustic properties.

- Doubling the sounds frequency does not mean the sound sounds twice as high (nonlinear).
- A “high” frequency for one speaker can be “low” for another speaker (200HZ is high for an adult male, low for an child).

Pitch / Intonation

The ToBI annotation system (Tones and break indices) is one popular system used to annotate pitch and boundary information.

Attempt: to convert the continuous information in the speech stream having to do with pitch,length, loudness and breaks, and convert this information into a set of categories, which may have some interpretation in the context.

*	pitch accent (one per intonational phrase)
H, L	pitch level
%	boundary tone

Ketchup is a vegetable?	Ketchup is a vegetable.
L* L* H%	H* H* L%

Graph removed for copyright reasons.

H*

L

L%

Pitch contour with ToBI transcriptions for
“It was the terrorist that kidnapped the ambassador.”

Intonational Phrasing

Speech is segmented into auditory chunks called “intonational phrases”

People pause for a variety of reasons:

- When having trouble formulating an utterance;
- External factors: interruptions etc.
- Natural production: **Planning**

Within **planning**: intonational phrasing:
An intonational phrase is a perceptually defined unit that is correlated with several quantitative measures (Nespor & Vogel, 1987; Selkirk, 1984), none of which are obligatory:

- Ending with pause
- Tends to have a pitch accent on the final word of the phrase
- Lengthening / Stretching of the final word.

Intonational Phrasing

“Bill wants to walk but Jane prefers to take a taxi.”

Usually produced as two intonational phrases (IPs):
“Bill wants to walk” and “but Jane prefers to take a taxi.”

Within an intonational phrase: One accented syllable (e.g., “Bill” in “Bill wants to walk”).

Different tunes to the two intonational phrases:

The first IP starts high and descends, but then ends with a rising tone.
The second IP ends with a falling tone: characteristic of an end of a complete utterance.

Pitch reset between IPs: Start again at high, and descend. This allows speakers to stay within their speaking range of pitch.

Pitch on intonational phrases

Different tunes for different meanings

1. Start high, end low: Declarative.
2. Start low, end high: Interrogative.
3. Negating presupposition: Flatter pitch throughout, then sharp rise in pitch at negated element:

   Homer drinks Budweiser beer?

   No, Homer drinks Duff beer.

Intonational Phrasing

The same syntactic structure can have multiple intonational phrasings:

a) John // gave the book to Mary.

b) John gave the book // to Mary.

c) John //gave the book // to Mary

Intonational Phrasing

Some structures seem to require their own intonational phrasing for pragmatic or semantic reasons:

Asides
Homer, *as you know*, can't take care of himself.

Non-restrictive modifiers
Homer, *who works at a power plant*, caused an accident.

Vocatives
Homer, did you take out the trash?

Timing

Concerns the overall timing of the segments in a sentence, specifically, the duration of words and pauses.

Intonational phrase final lengthening:

Compare “walk” in:

(1) “Bill wants to walk.”

(2) “Bill wants to walk to the store.”

(3) “Bill wants to walk, but Mary wants to drive.”

Long in (1) and (3); shorter in (2).

Timing

Function words: words that have a grammatical function (e.g. of, and, but, the,….)

Content words: words that convey meaning in a sentence (e.g. blue, dog, run,……)

Content words tend to be long and function words tend to be short (shorter then they would be in isolation).

I want to leave
I want two leaves

Stress

Stress = loudness.

Syllable by syllable:

“chimpanzee”: “zee” largest stress; “chimp” next largest; “pan” least stress.

Note: Duration and loudness are highly correlated. Loud sounds also get lengthened; Quiet sounds get shortened.

Information structure

- Focus: New information
- Background information: Old information

- Focused elements are stressed, lengthened and spoken with a pitch accent

  Person A: Homer likes Budweiser beer.
  Person B: No, Homer likes DUFF beer.

  “Duff” is the focus: stressed, accented.

Intonational Phrasing

Some factors affecting intonational phrasing seem to be due to performance:

- rate of speech
- sentence length

Question: What role does syntactic structure play in intonational phrasing?

Intonational Phrasing

The senator who the reporter attacked fired the secretary.

The senator who the reporter attacked // fired the secretary.

The goal of the Watson & Gibson research program is to find out where prosodic boundaries are preferred in production and comprehension.

Sentence Production: Two Factors (Watson & Gibson, 2004)

1) Resting: The size of the syntactic constituent that has just been completed predicts the likelihood of an intonational boundary. (A constituent is completed if it has no rightward dependents.)

2) Planning: The size of the upcoming syntactic constituent also predicts the likelihood of an intonational boundary.

Motivation for factor 1: the size of the constituent just completed

PP's modifying the verb in a V-NP-PP sequence are more likely to have a boundary preceding them than those modifying the NP:

Jeffrey hit the cop with a stick

The constituent just completed:

none for NP attachment
[the cop] for VP attachment

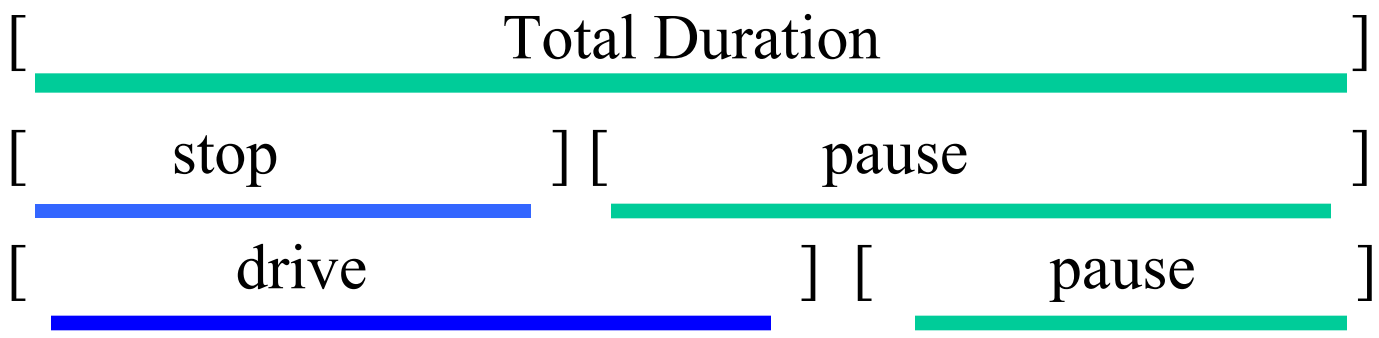
Thus there is a greater tendency to place a boundary after "the cop" when VP attachment is desired.

Aside: Syntactic structure predicts the **likelihood** of a pause, not the **size** of a pause

The segmental properties of a word (length of the vowels etc) along with the syntactic structure determine the prosodic phrasing: (Ferreira, 1993).

Even though the chauffeur thought he could **stop**, the passengers were worried (short vowel, phrase final)

Even though the chauffeur thought he could **drive**, the passengers were worried (long vowel, phrase final)



Upcoming Constituent Size

Intonational boundaries tend to occur before large syntactic constituents (Lehiste, 1973; Gee & Grosjean, 1983; Ferreira, 1991).

….also constrained by the Sense Unit Condition (Selkirk, 1984)

People tend not to place boundaries between a head and its adjacent argument.

A sense unit is defined as a constituent formed by a head and, optionally, the head of any number of its modifiers and/or arguments:

*1) [Three mathematicians] [in ten derive a lemma]

2) [Three mathematicians in ten] [derive a lemma]

The Sense Unit Condition of Intonational Phrasing: The immediate constituents of an intonational phrase must together form a sense unit. Two constituents Ci, Cj form a sense unit if Ci depends on Cj.

Observation: Obligatory arguments are planned with their heads.

Updated Planning Hypothesis 2:
Intonational boundaries tend to occur before large upcoming syntactic constituents **that are not obligatory arguments** of the current constituent.

Measure of Size

Initial proposal: Phonological phrases: all the words up to and including a content word (Nespor & Vogel, 1984).

For our purposes: Count only main verbs and content nouns. (No adjectives, adverbs)

The-judge who-ignored the-reporter fired the-secretary

Phonological phrases act as an additional constraint on intonational phrasing. (Gee & Grosjean, 1983)

LRB Hypothesis (Watson & Gibson, 2004)

Left Constituent / Right Constituent Boundary (LRB) Hypothesis: 2 independent factors are summed to predict the likelihood of an intonational boundary at a phonological phrase boundary in production.

1) The number of phonological phrases in the largest syntactic constituent that has just been completed. A constituent is completed if it has no rightward dependents.

2) The number of phonological phrases in the upcoming syntactic constituent if it is not an obligatory argument of the most recently processed constituent.

Positive features of the LRB

It is consistent with all grammars that rely on dependency relationships.

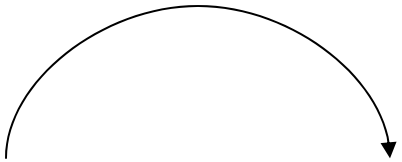
Fewer parameters and steps than previous production models:

Cooper & Paccia-Cooper (1980)

Gee & Grosjean (1983)

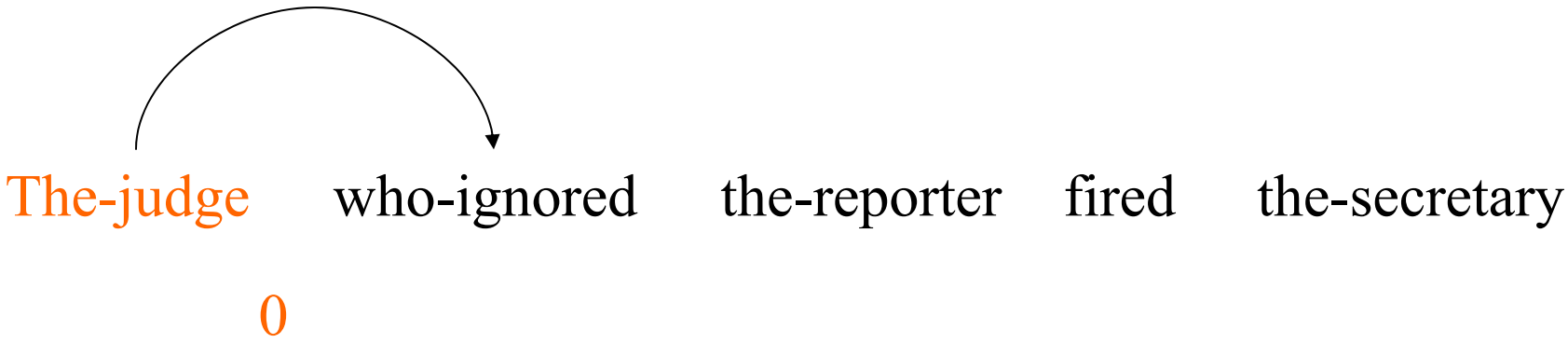
Ferreira (1988)

Example



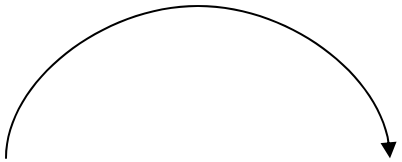
The-judge who-ignored the-reporter fired the-secretary

Example



Example

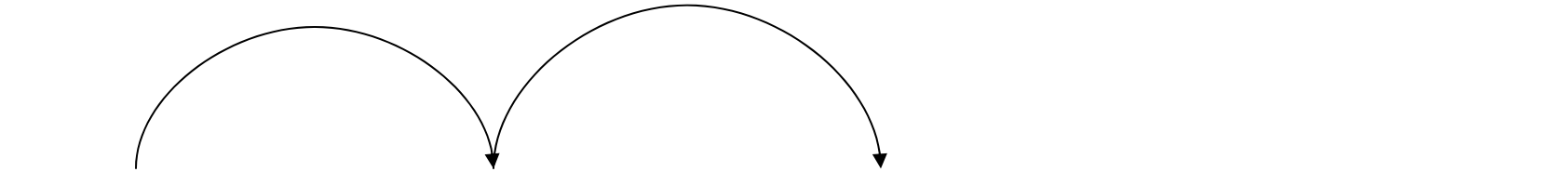
Example



The-judge who-ignored the-reporter fired the-secretary

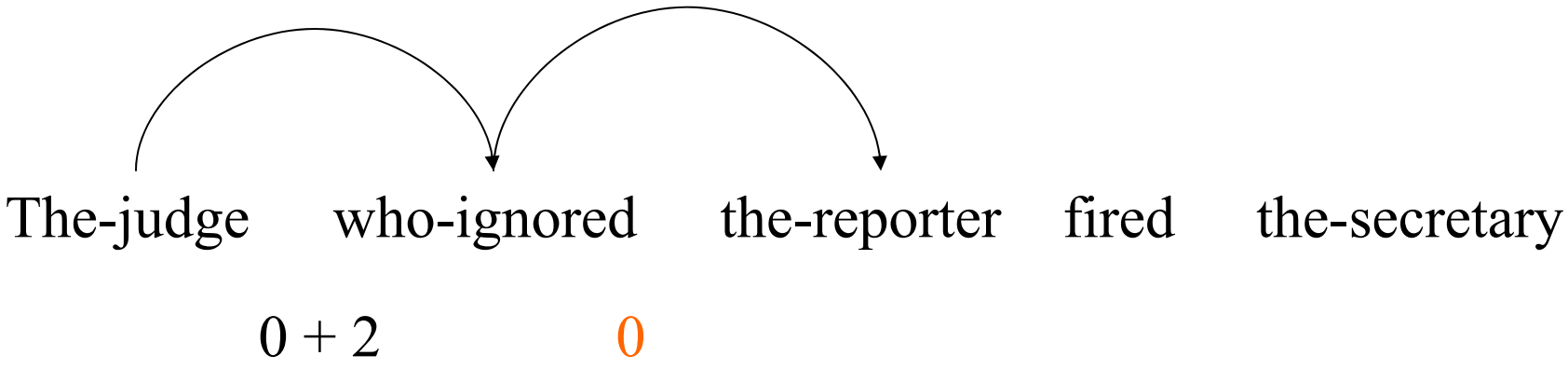
0 + 2

Example

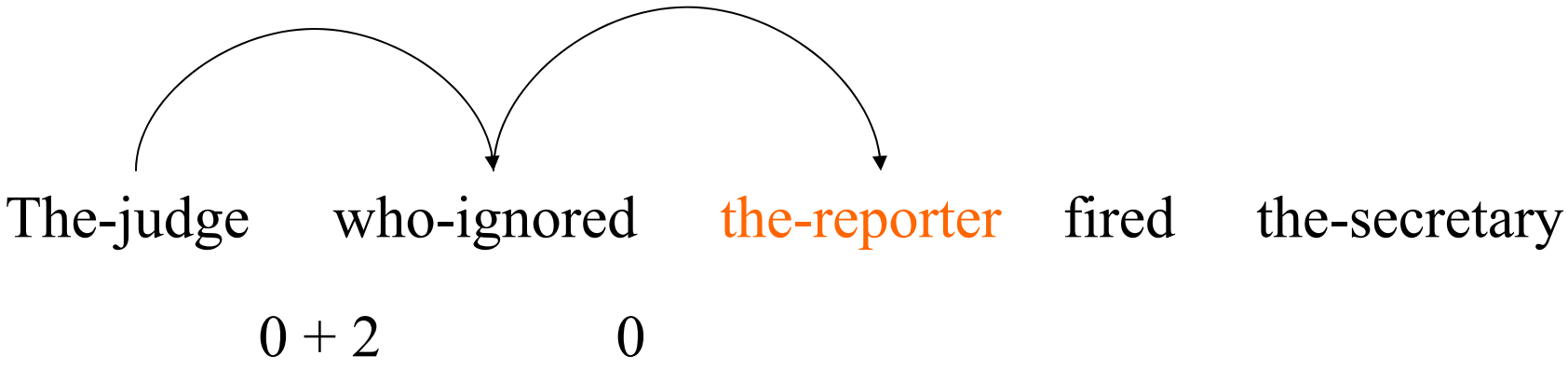


0 + 2

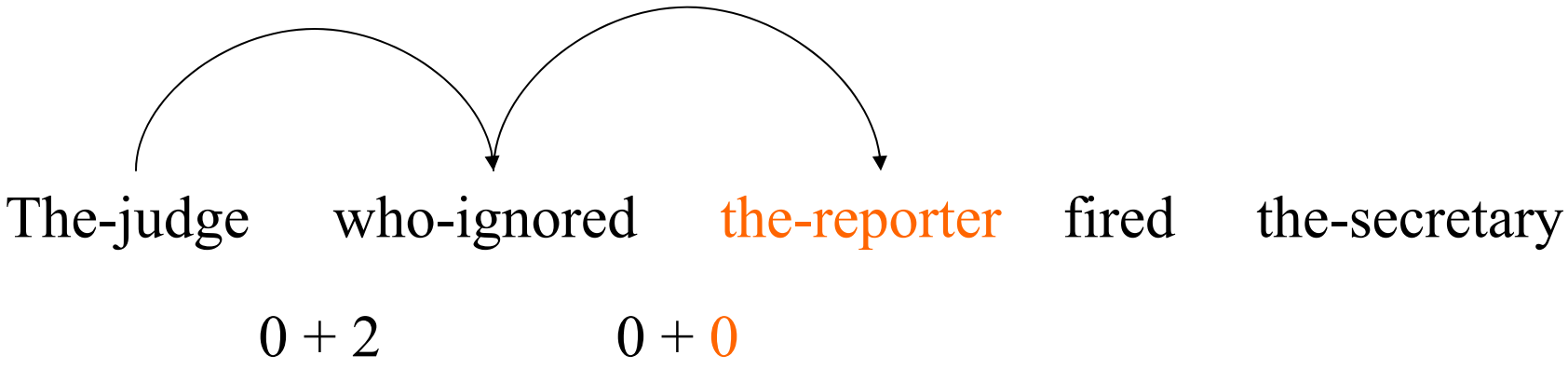
Example



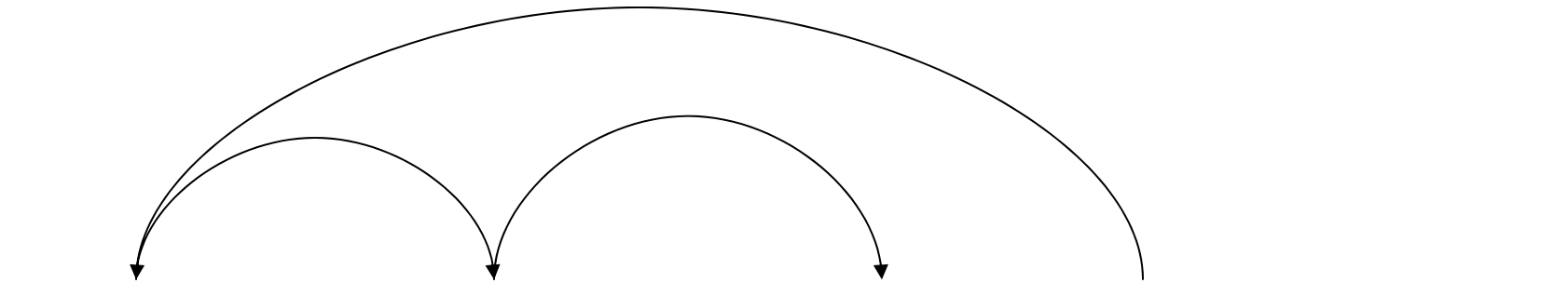
Example



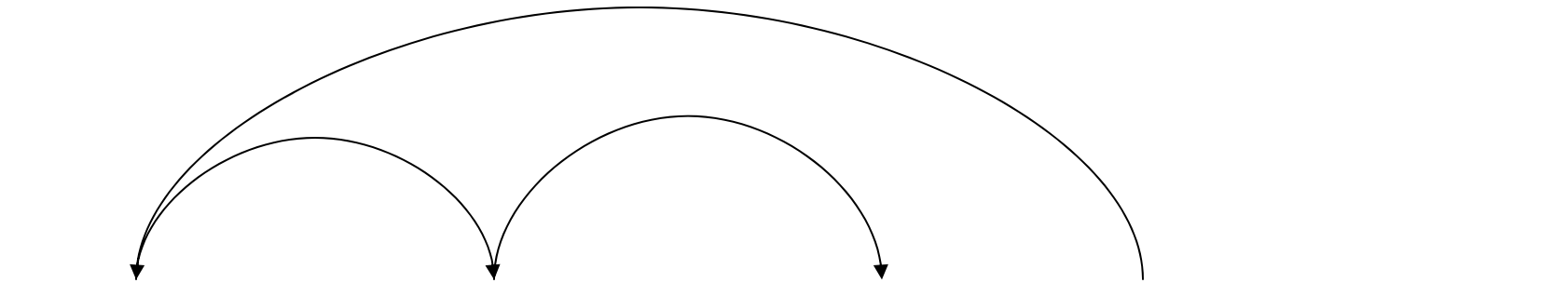
Example


The-judge
who-ignored
the-reporter
fired
the-secretary
0 + 2
0 + 0

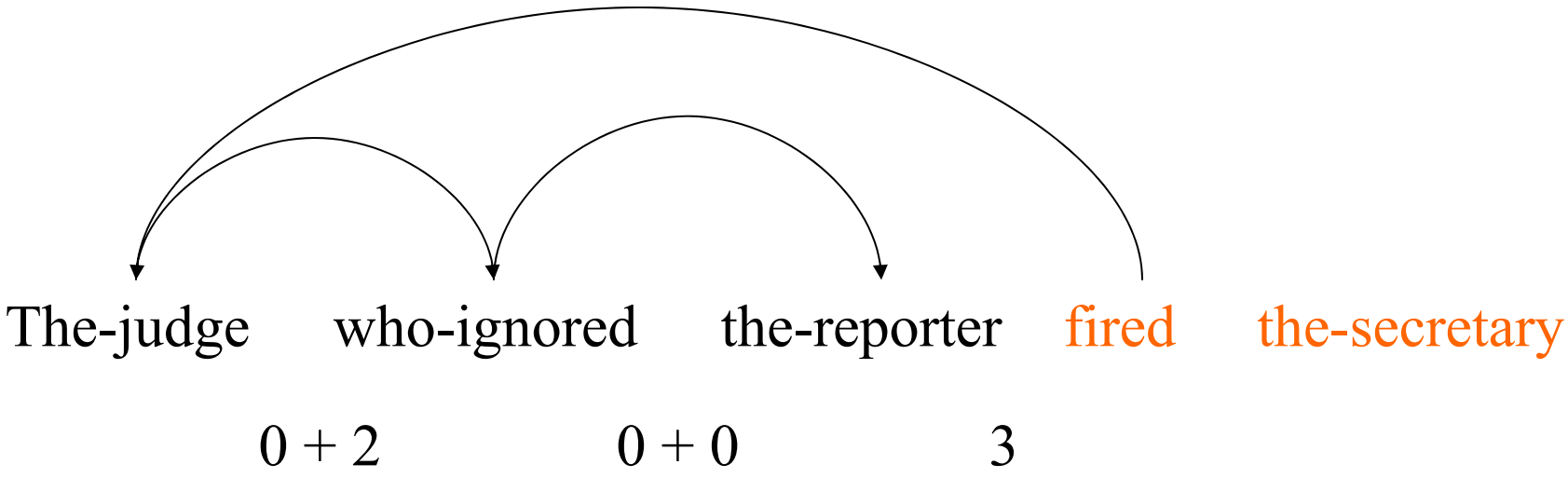
Example


The-judge
who-ignored
the-reporter
fired
the-secretary
0 + 2
0 + 0

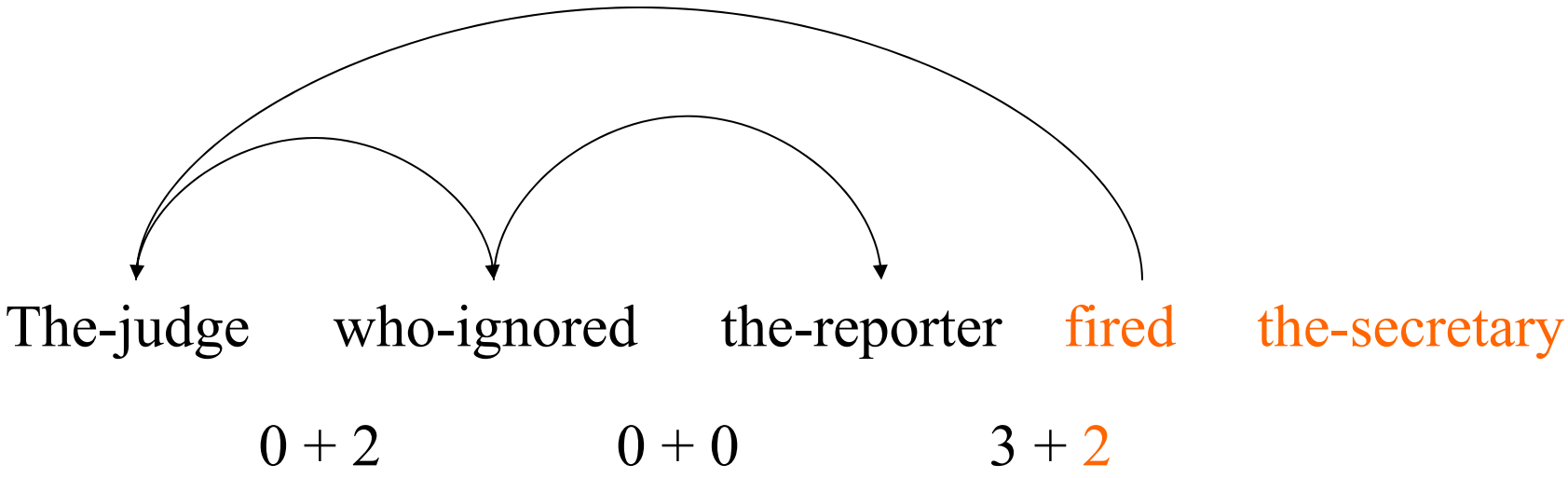
Example


The-judge
who-ignored
the-reporter
fired
the-secretary
0 + 2
0 + 0
3

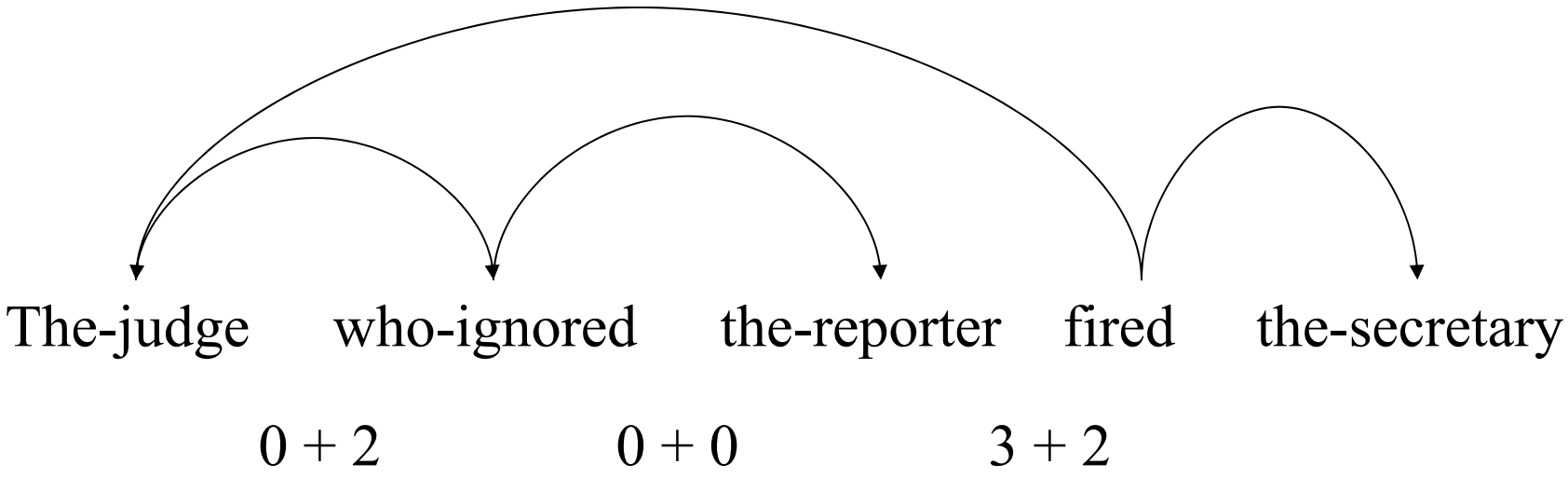
Example


The-judge
who-ignored
the-reporter
fired
the-secretary
0 + 2
0 + 0
3

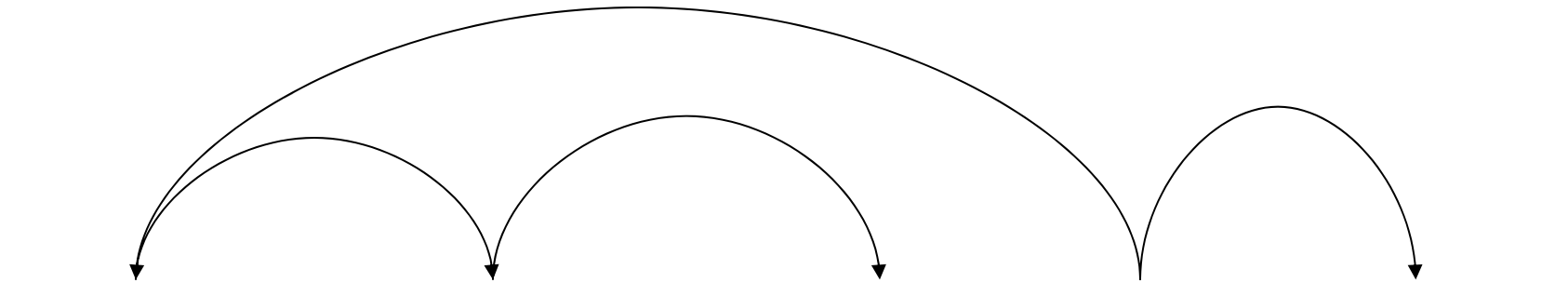
Example


The-judge
who-ignored
the-reporter
fired
the-secretary
0 + 2
0 + 0
3 + 2

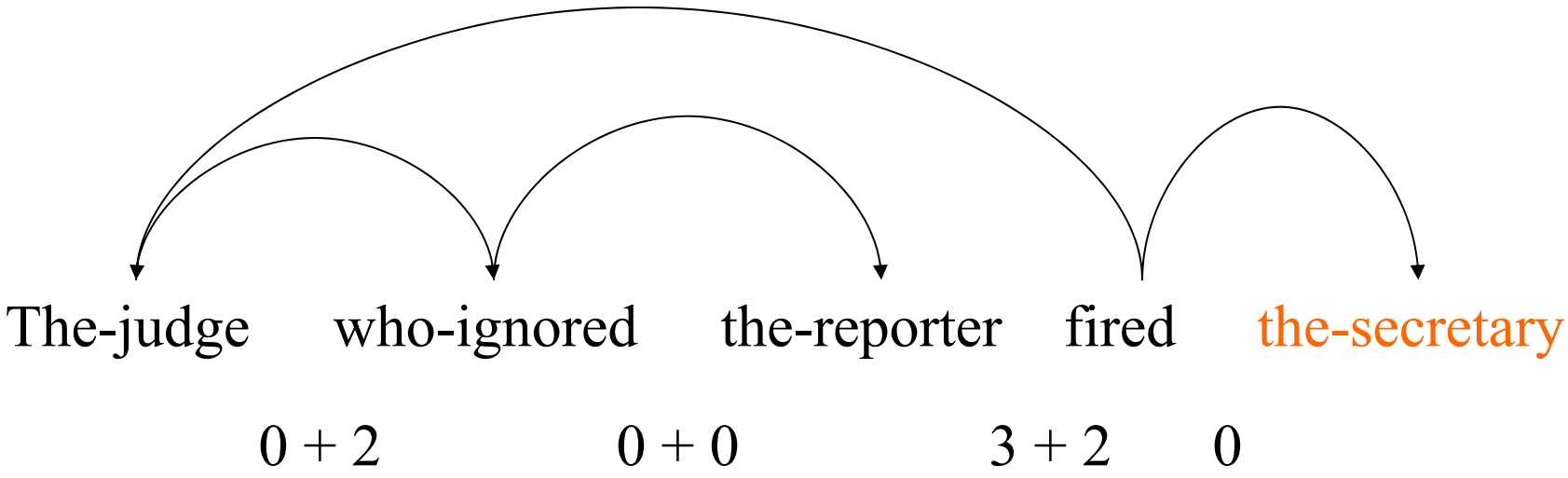
Example


The-judge
who-ignored
the-reporter
fired
the-secretary
0 + 2
0 + 0
3 + 2

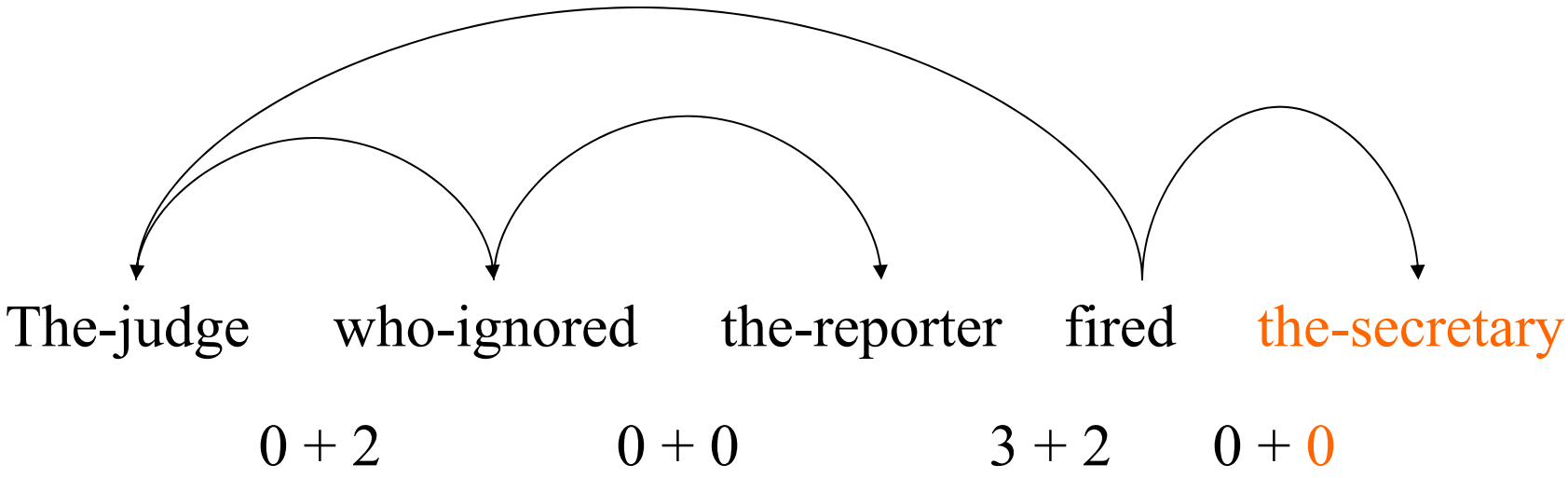
Example


The-judge
who-ignored
the-reporter
fired
the-secretary
0 + 2
0 + 0
3 + 2
0

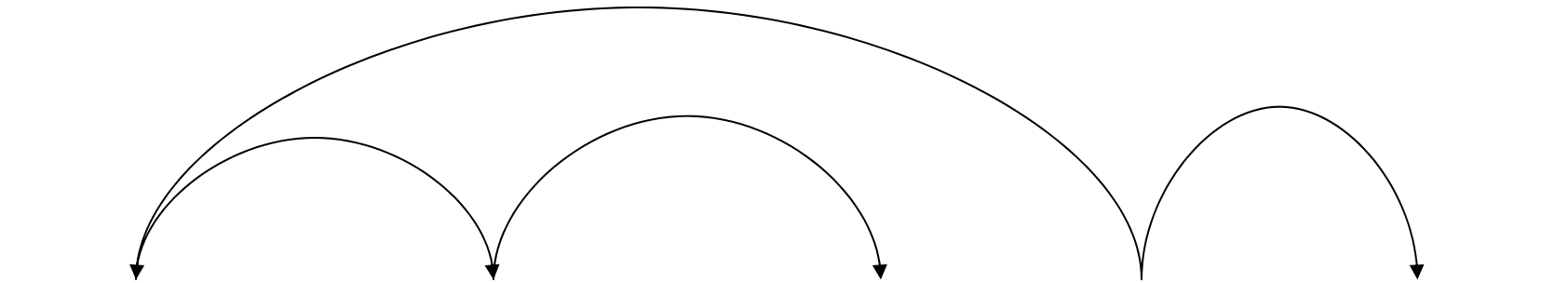
Example


The-judge
who-ignored
the-reporter
fired
the-secretary
0 + 2
0 + 0
3 + 2
0

Example


The-judge
who-ignored
the-reporter
fired
the-secretary
0 + 2
0 + 0
3 + 2
0 + 0

Example


The-judge
who-ignored
the-reporter
fired
the-secretary
0 + 2
0 + 0
3 + 2
0 + 0

Watson & Gibson (2002a) Experiment 1

Earlier studies only tested one token for each syntactic type. (Gee & Grosjean, 1983; Cooper & Paccia-Cooper, 1980; Ferreira, 1988).

Watson & Gibson Expt 1: 8 different syntactic structures.

32 subjects

32 items

3 coders coded prosodic breaks with 95% reliability between coders

Watson & Gibson (2002a) Experiment 1

Task

There were 2 participants: a reader and a listener

The reader was recorded while reading sentences to the listener.

The reader was instructed to read the sentences over in their head.

The listener answered yes/no questions regarding the sentences on a computer.

Watson & Gibson (2002a) Experiment 1

Subject-extracted relative clause

1)The judge who ignored the reporter fired the secretary.

Object-extracted relative clause

2) The judge who the reporter ignored fired the secretary.

Object-extracted relative clause + 1 PP

3) The judge who the reporter for the newspaper ignored fired the secretary.

Object-extracted relative clause + 2 PP

4) The judge who the reporter for the newspaper in the capitol ignored fired the secretary.

Watson & Gibson (2002a) Experiment 1

Doubly embedded object/subject relative clauses

5) The judge who the reporter who attacked the senator ignored fired the secretary.

Doubly embedded object/object relative clauses

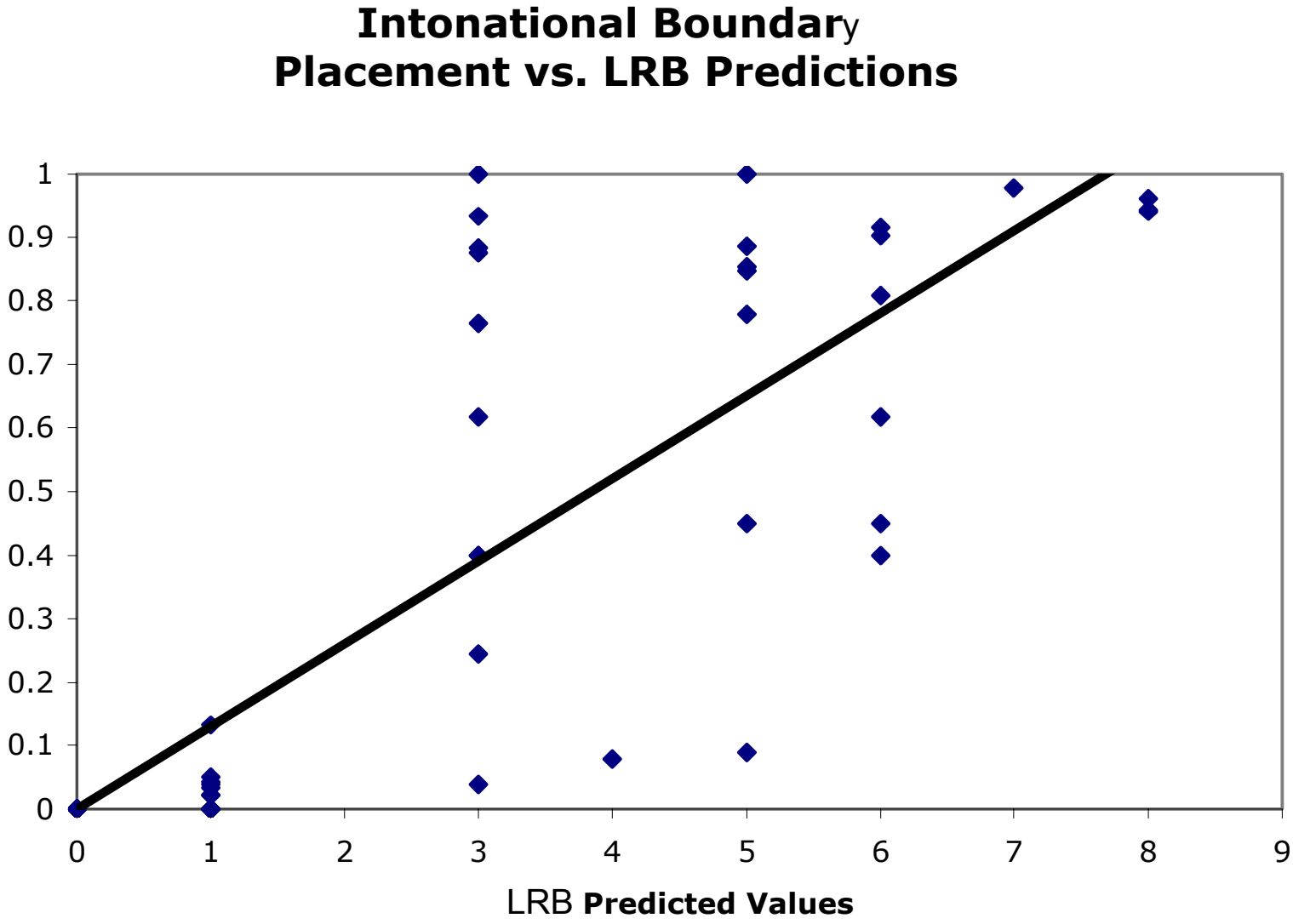
6) The judge who the reporter who the senator attacked ignored fired the secretary.

Two Clause right branching

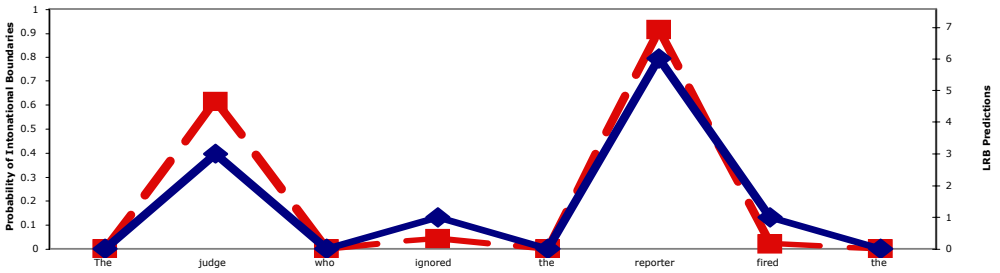
7) The reporter ignored the judge who fired the secretary

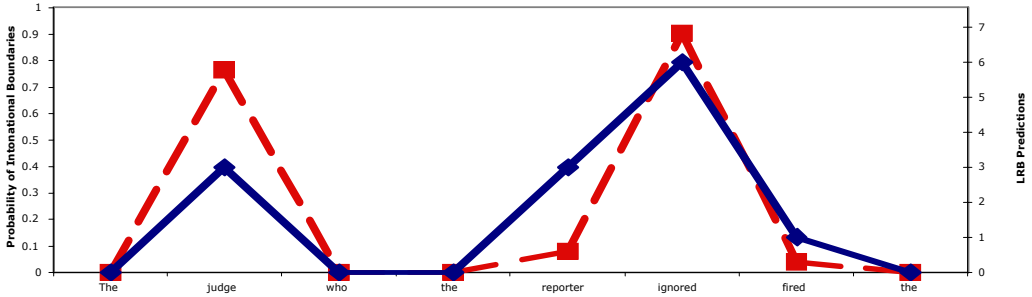
Three Clause right branching

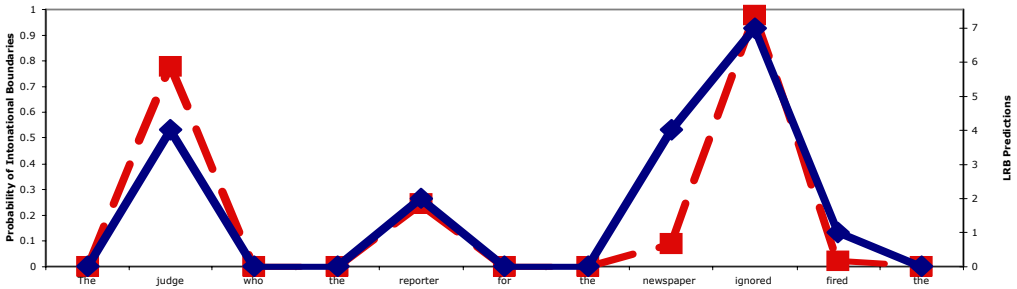
8) The senator attacked the reporter who ignored the judge who fired the secretary

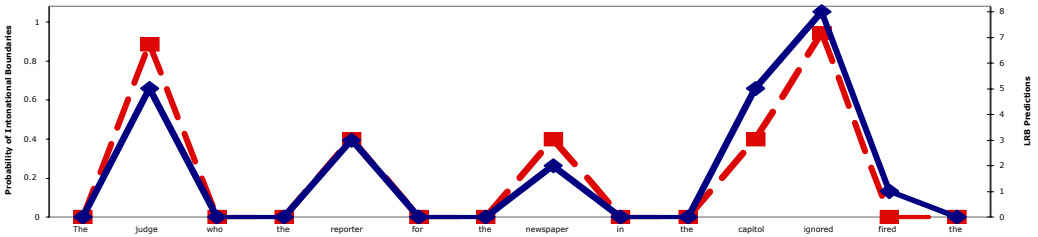

Intonational Boundary
Placement vs. LRB Predictions
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6
7
8
9
LRB Predicted Values

Model Predictions vs Proportion I-Boundary for each condition





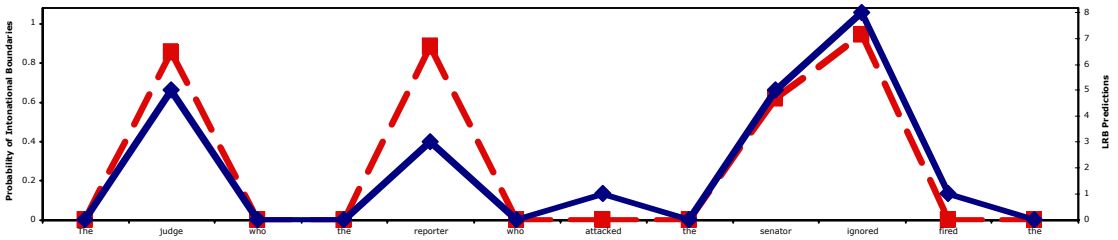


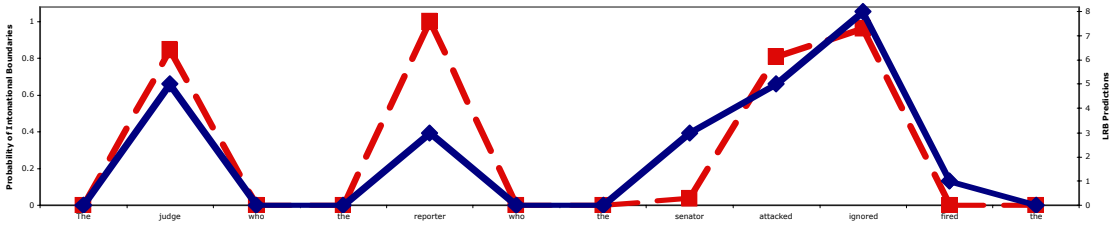


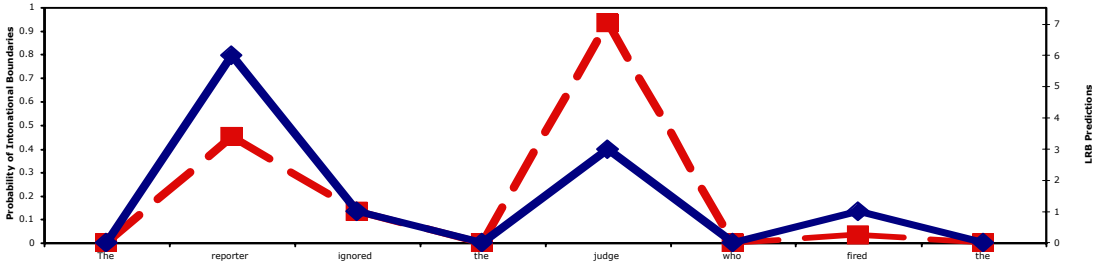
LRB Predictions

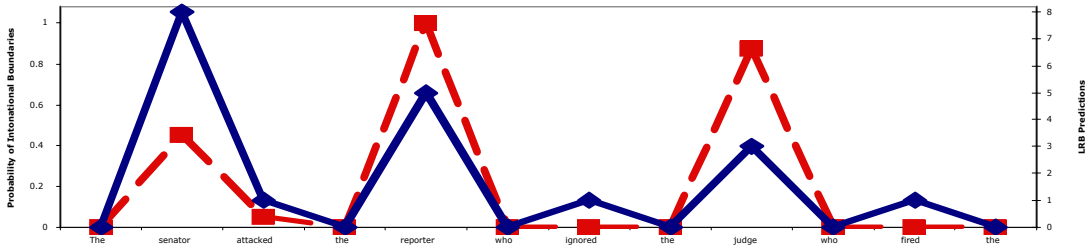
Intonational Boundaries

Model Predictions vs Proportion I-Boundary for each condition









LRB Predictions

Intonational Boundaries

Watson & Gibson (2004) Experiment 1

Model	R^2
LRB	.74
Gee & Grosjean (1983)	.76
Ferreira (1991)	.71

These models make roughly the same predictions.

Ex. The judge who ignored the reporter fired the secretary.

Watson & Gibson (2004) Experiment 2

The teacher planned a visit of the children of the class to the zoo

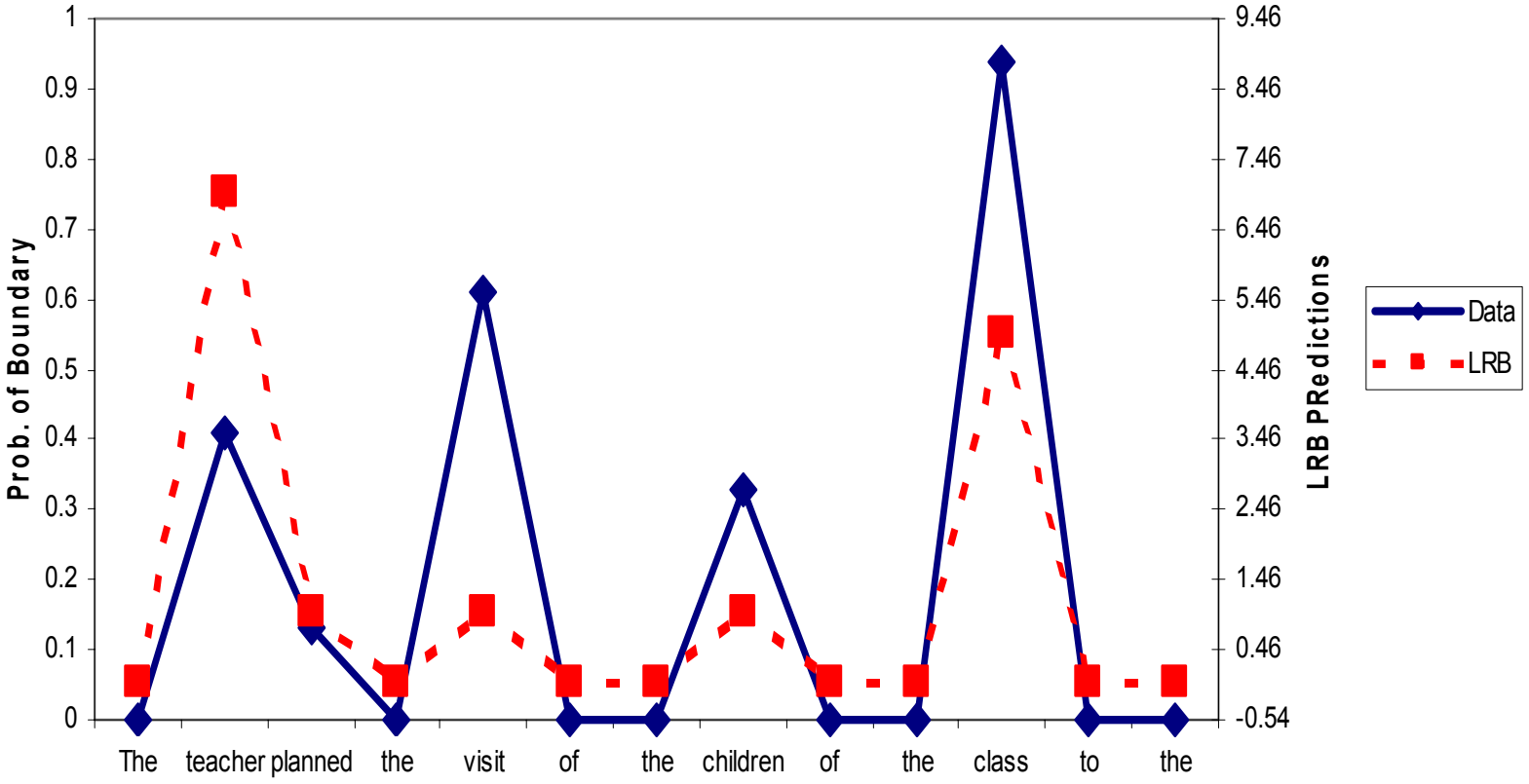
- 20 subjects
- 10 items
- 35 fillers
- listener - reader paradigm

Watson & Gibson (2004) Experiment 2

The teacher planned a visit of the children of the class to the zoo

Models	R^2	
LRB	.52*	$p < .01$
GG	.25	
X-bar	.22	

Watson & Gibson (2004) Experiment 2



LRB Conclusions

•Experiment 1 showed that on the sentences tested, the LRB can account for the same amount of variance as previous models, which have more parameters.

•Experiment 2 demonstrated that in some instances, the LRB does significantly better than previous models and that speakers prefer boundaries after long constituents.

Comprehension & Intonational Phrasing

The LRB describes a link between intonational phrasing and syntactic structure.

Given this link, listeners could potentially use intonational boundaries to make inferences about syntactic structure.

How are intonational boundaries used in parsing?

Using the LRB hypothesis in sentence comprehension (Watson & Gibson,in press)

The presence of an intonational boundary serves as a strong cue not to attach the upcoming word to the last potential attachment site before the boundary.

1. The LRB predicts that speakers tend to place an intonational boundary following a completed constituent, especially if the constituent is long.

   A boundary therefore often indicates that a syntactic constituent is complete and no further attachments should be made to it.

2. The LRB predicts that speakers tend to place an intonational boundary before a large upcoming constituent, if it is not an argument of the current head. A boundary would therefore suggest that upcoming constituent is not an argument of the current head.

Using the LRB hypothesis in sentence comprehension (Watson & Gibson, in press)

The Anti-Attachment Hypothesis (AAH):

Listeners prefer not to attach an incoming word to a lexical head that is immediately followed by an intonational boundary.

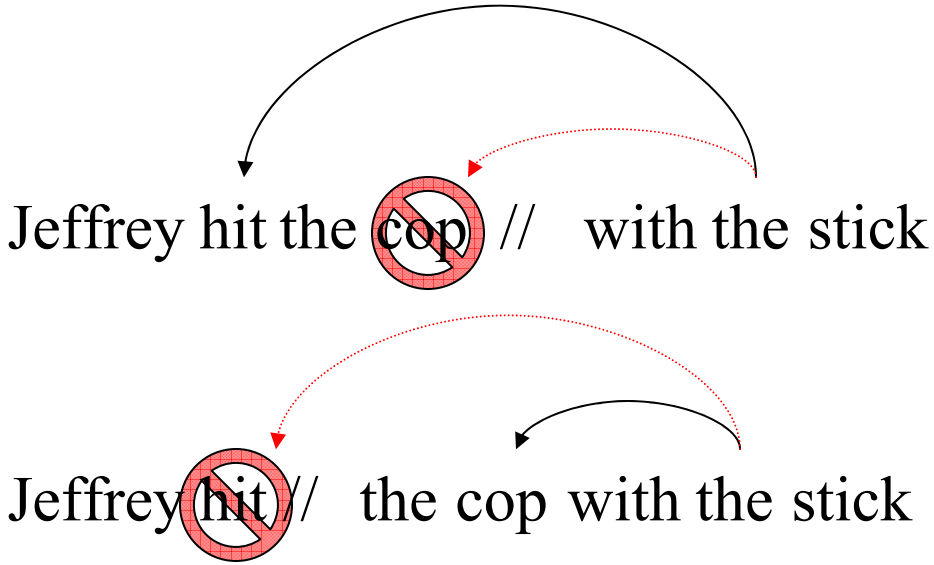
Implications:

1) The presence of an intonational boundary after a lexical head that is the site of subsequent attachment increases processing difficulty;

2) The presence of an intonational boundary after a lexical head that is not the site of subsequent attachment decreases processing difficulty.

Anti-Attachment

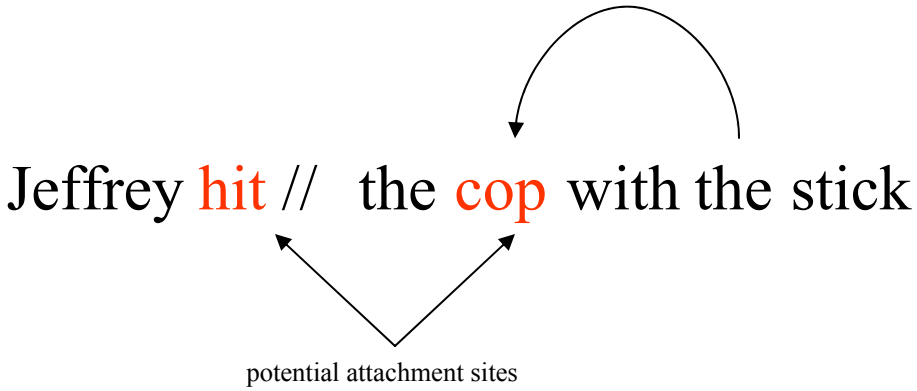
Anti-Attachment Hypothesis: When a listener detects an intonational boundary, there is a preference not to attach upcoming constituents to the head immediately before the boundary.



In unambiguous structures, comprehension will be facilitated when dependents are attached to heads that are not followed by an intonational boundary

An alternative view: Disruption theories

Disruption Theories (Slowiaczek, 1981; Schafer, 1997; Frazier & Clifton, 1998, Speer et al., 1996): Constructing dependencies between words within the same intonational phrase is easier than constructing dependencies between words in different intonational phrases.



Disruption Theory Predictions

In unambiguous sentences, comprehension will be disrupted by the insertion of an intonational phrase boundary.

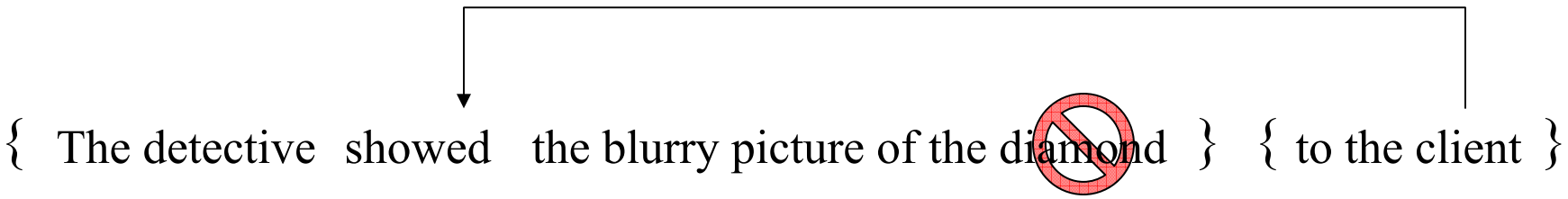
{ The detective showed the blurry picture of the diamond to the client }

{ The detective showed the blurry picture of the diamond } { to the client }

Watson & Gibson (in press) Experiment 1

Disruption theories predict this intonational structure will be easier:

The anti-attachment hypothesis predicts this intonational structure will be easier:



Watson & Gibson (in press) Experiment 1

In this experiment, subjects listened to sentences that varied in the position of their prosodic breaks, and judged how difficult they were to understand.

2 x 2 Design

The detective showed the blurry picture of the diamond // to the client // who was in an office that was on the fourteenth floor.

Watson & Gibson (in press) Experiment 1

The detective showed the blurry picture of the diamond // to the client // who was….

Each condition was produced by a speaker independently with the intended prosodic boundaries.

"of the diamond to the client who was in a office" was extracted from each condition and spliced into a control.

16 items

30 fillers

4 counterbalanced lists

Watson & Gibson (in press) Experiment 1

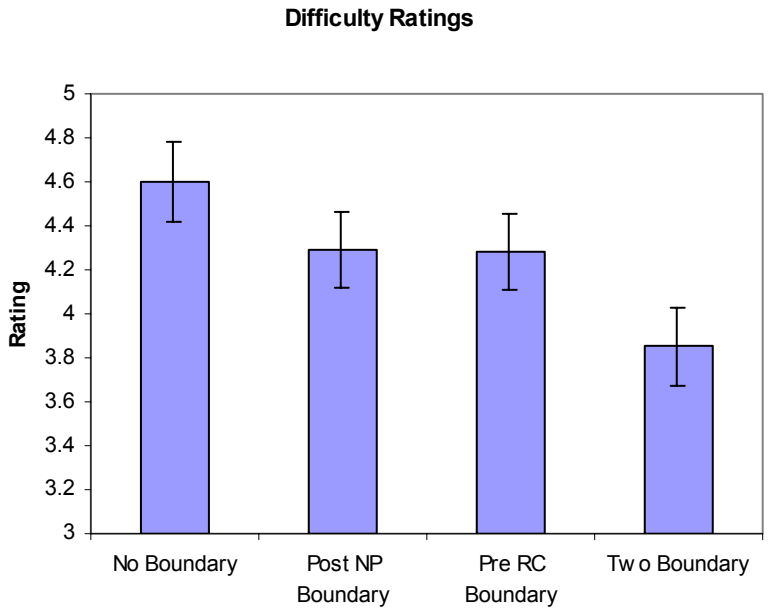
40 native English speakers from the MIT community

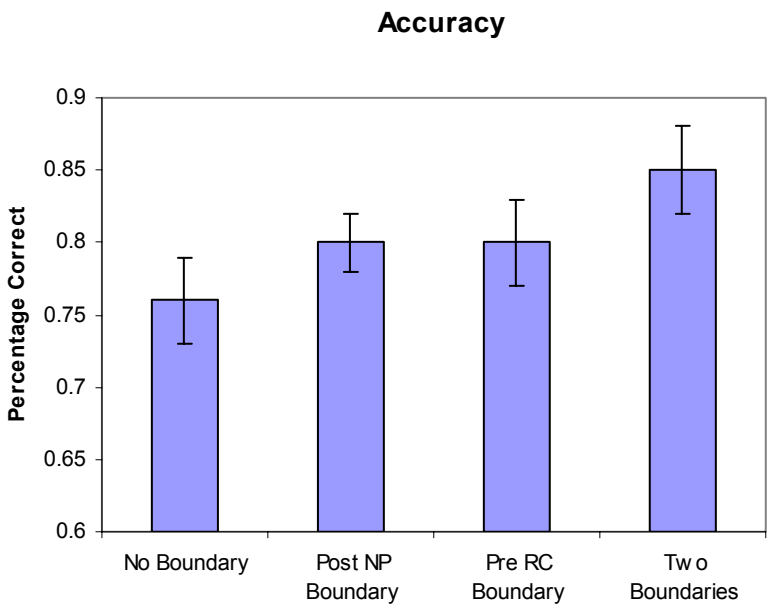
The experiment was presented as a questionnaire on a web page.

Subjects listened to each sentence on headphones and then rated the difficulty of the sentence on a scale of 1 to 7, 1 being easiest and 7 being the most difficult.

Question answering task to insure comprehension

Watson & Gibson (in press) Experiment 1





The detective showed the blurry picture of the diamond // to the client // who was….

Participants rated the sentences as "easier" and answered more questions correctly when there was an intonational boundary after the complex NP and/or a boundary before the relative clause.

Watson & Gibson (in press) Experiment 1
Conclusions

The presence of dependencies across intonational boundaries did not increase difficulty, contrary to the prediction of the disruption theories.

Remaining questions:

Was it simply the case that the presence of a boundary anywhere in the items in Experiment 1 made the sentence easier?

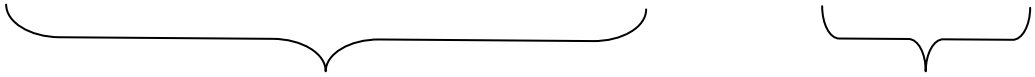
Is intonational boundary information used immediately on-line?

Watson & Gibson (in press) Experiment 2

Cross Modal Lexical Decision Task: The lexical decision time tells us the degree to which the subjects expect the word given the utterance's prosody.

The detective gave the picture . . . OF / TO
H* L-L%

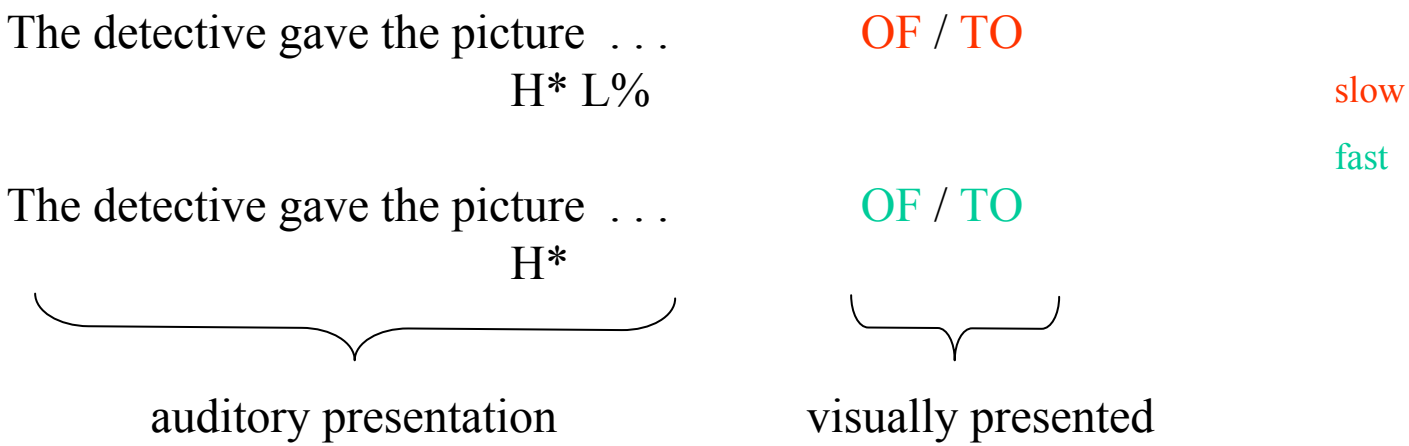
The detective gave the picture . . . OF / TO
H*



auditory presentation visually presented

Watson & Gibson (in press) Experiment 2

Disruption theories predict slower times when an intonational boundary is present than when an intonational boundary isn't present.



Watson & Gibson (in press) Experiment 2

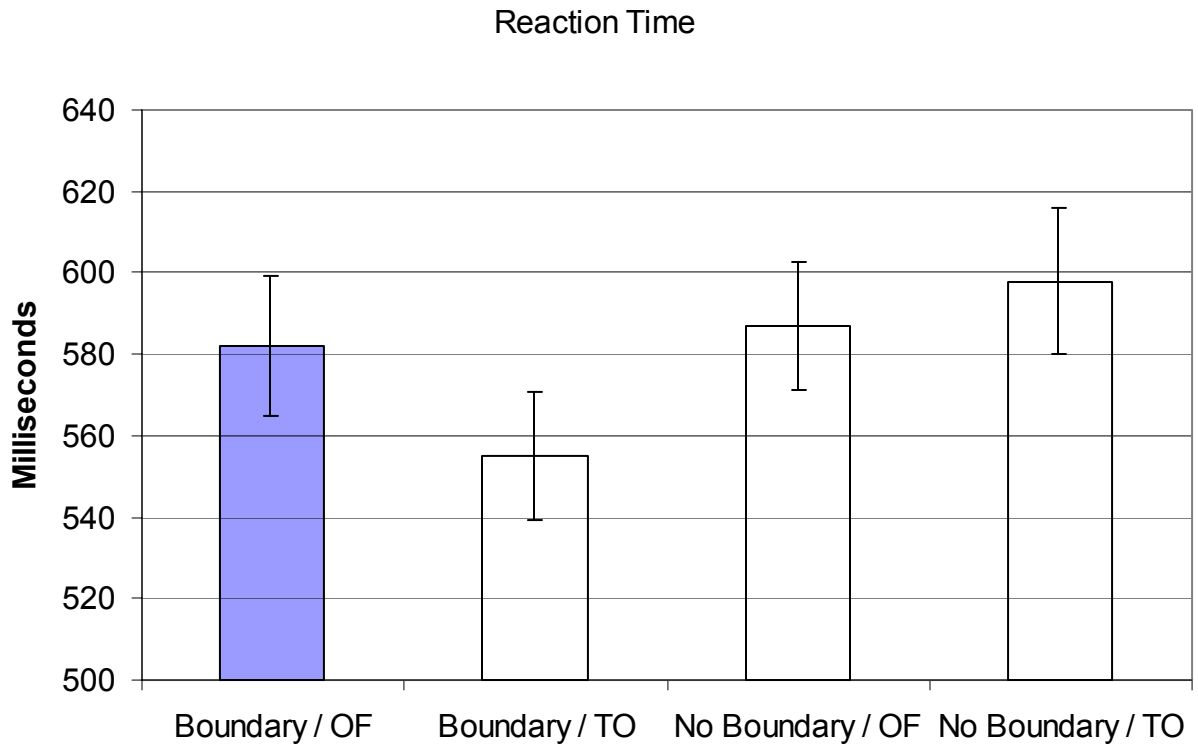
Anti-attachment theory predicts that a boundary after the direct object will decrease RT to “TO”. When there is no boundary, RT will be faster to “OF”.

The detective gave the picture . . . H* L% OF / TO slow

The detective gave the picture . . . H* OF / TO fast

auditory presentation visually presented

Watson & Gibson (in press) Experiment 2 Results



There is a significant interaction, F_1=7.78 (1,39), p <.01

Intonational boundary information appear to be used rapidly in on-line parsing.

Watson & Gibson (in press) Experiment 2 Results: Discussion

- Do the results fit the predictions of the AAH? Almost, but not quite.
- Odd data point: Slow RTs for the No-boundary local "of"-attachment: This should have been fast, according to the AAH
- Possible interpretation: The task is somewhat odd for the no-boundary condition, because there are prosodic cues of a boundary: the pause.
- RTs may be slow whenever some cues conflict with the observed attachment:
  - No cues conflict for non-local, boundary
  - Some cues conflict for the other three conditions, hence elevated RTs

Conclusions in Comprehension

1) Listeners prefer not to attach incoming constituents to lexical heads that are followed by an intonational boundary

2) This intonational boundary information is used rapidly on-line.

Conclusions

Speakers tend to place boundaries before and after large syntactic constituents.

Listeners use intonational boundaries as cues that signal where to make attachments.