

Sentence comprehension IV

9.59; 24.905

March 3, 2005

Ted Gibson

To do today

- The dependency locality theory
 - Review integration;
 - New: storage
- Ambiguity effects (Gibson, 2000)
- An interesting test case of resource theories:
Chinese sentence comprehension (Hsiao & Gibson, 2003)

What sources of information do people use in processing sentences?

- Syntactic structure
- Word frequency
- Plausibility
- Discourse context
- Syntactic complexity
- Intonational information

Syntactic information use in sentence processing: The Dependency Locality Theory (DLT, Gibson, 1998, 2000)

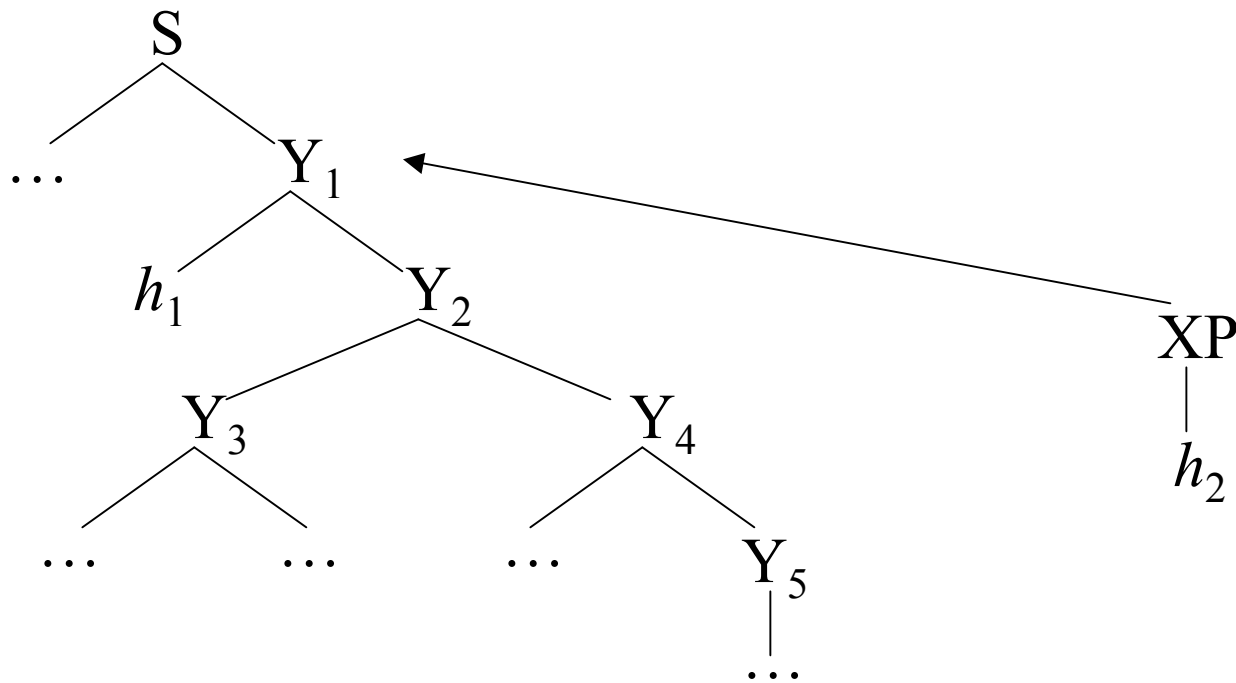
Resources are required for two aspects of language comprehension:

- (a) Integration: connecting the current word into the structure built thus far;
- (b) Storage: Predicting categories to complete the current structure.

Syntactic resource hypotheses

- Frazier (1978): Minimal Attachment & Late Closure:
 - Ambiguity resolution only
 - Assumed to be modular: Applying before other sources of information use
- Gibson (1998; 2000): Syntactic storage and integration
 - Apply in both ambiguous and unambiguous sentences
 - Assumed to be non-modular: Interact immediately with other sources of information (but this is not a crucial part of the theory).

Integration complexity depends on the *distance* or *locality* between the head and dependent being integrated.



Computational motivation: Integrating h_2 to h_1 involves re-accessing h_1 :

- h_1 's activation may have **decayed** for all the integrations that have taken place since it was last highly activated.
- **Interference** of intervening elements: similar to h_1 and/or h_2

Initial Integration Distance Hypothesis

Gibson (1998); Warren & Gibson (2002, *Cognition*)

The difficulty of integrating a new word h_2 to h_1 is proportional to the number of discourse objects and events (**nouns** and **verbs**, roughly) which were introduced since h_1 was last processed.

Locality effects in unambiguous structures

Gibson (1998, *Cognition*); Grodner & Gibson (in press, *Cognitive Science*)

Experiment 1:

Object-extracted relative clause:

The reporter who the photographer sent to the editor hoped for a good story.

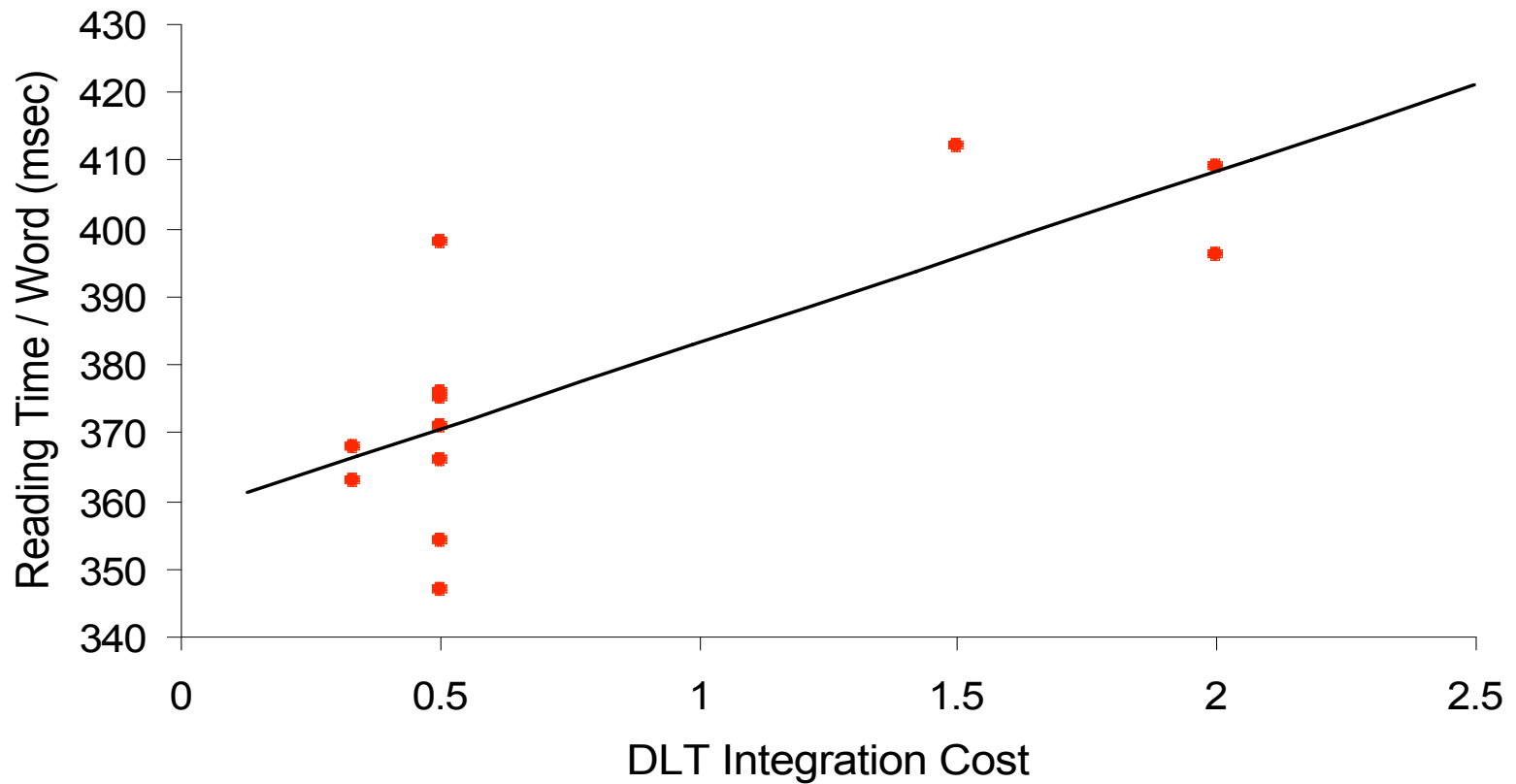
Subject-extracted relative clause:

The reporter who sent the photographer to the editor hoped for a good story.

Experiment 1: DLT vs. RTs

(Grodner & Gibson, in press, *Cognitive Science*)

Linear model: $r^2 = .582$, $p < .001$



Other effects of dependency distances

- Many construction comparisons cross-linguistically:
 - Heaviness effects: Put the heavy constituent later (Bever, 1970; Hawkins, 1994)
 - Nested vs. non-nested structures
 - Nested vs. cross-serial dependencies (Bach et al. 1986)

Nesting complexity effects

- (1) The reporter disliked the editor.
- (2) The reporter [who the senator attacked] disliked the editor.
- (3) The reporter [who the senator [who John met] attacked] disliked the editor.
- (4) John met the senator [who attacked the reporter [who disliked the editor]]].

Locality account of nesting complexity: Nested structures have longer distance dependencies than non-nested structures.

An old puzzle: relative clauses (RCs) and complement clauses (CCs)

(1) RC within CC: difficult, but processable

The fact [that the employee [who the manager hired] stole office supplies] worried the executive.

(2) CC within RC: unprocessable

The executive [who the fact [that the employee stole office supplies] worried] hired the manager.

Assumption: the cost of an integration peaks after a some intervening material.
The worst integration steps are therefore those that involve **multiple** integrations at the same point.

Maximal integration step in (1): 1 long integration, at "worried"; 2 short integrations at "hired".

Maximal integration step in (2): **2 long integrations, at "worried"**. (There is also 1 long integration at "hired".)

Nested vs. Cross-serial dependencies

Cross-serial dependencies are easier to process than nested dependencies (Bach, Brown & Marslen-Wilson, 1986)

Nested vs. Cross-serial dependencies


Nested (German)

... NP₁ NP₂ NP₃ VP₃ VP₂ VP₁

Nested vs. Cross-serial dependencies

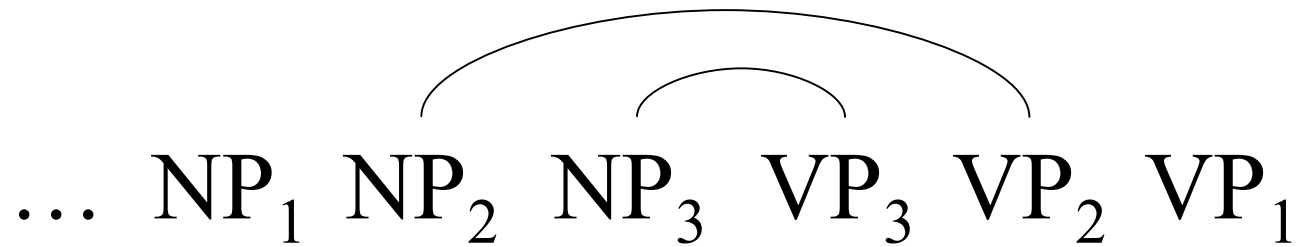
Nested (German)

... NP₁ NP₂ NP₃ VP₃ VP₂ VP₁



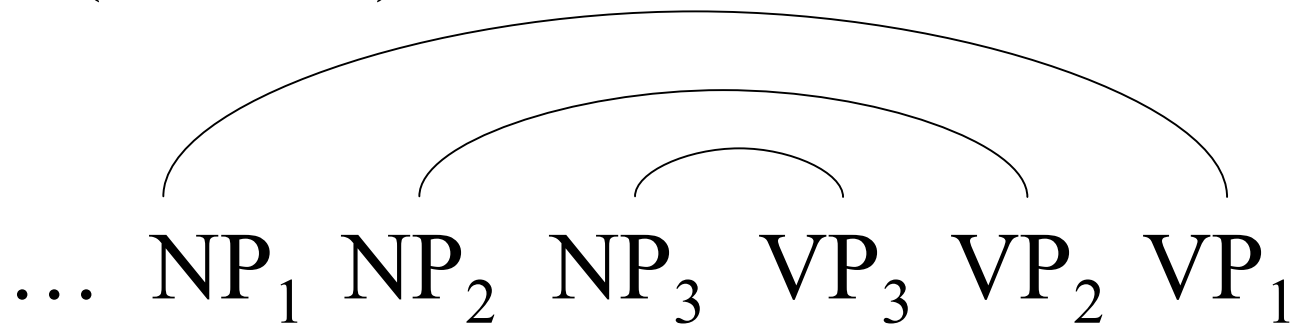
Nested vs. Cross-serial dependencies

Nested (German)



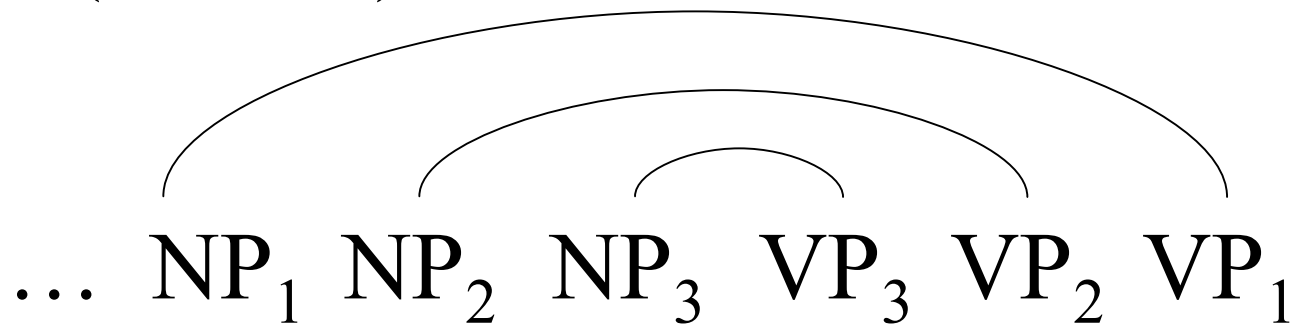
Nested vs. Cross-serial dependencies

Nested (German)



Nested vs. Cross-serial dependencies

Nested (German)



Johanna hat den Männern Hans die Pferde füttern lehren helfen.

Joanna has the men Hans the horses feed teach helped

“Joanna helped the men teach Hans to feed the horses.”

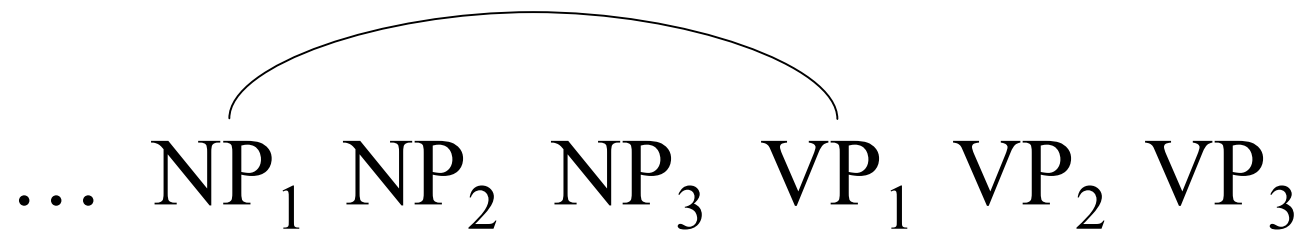
Nested vs. Cross-serial dependencies

Cross-serial (Dutch)

... NP₁ NP₂ NP₃ VP₁ VP₂ VP₃

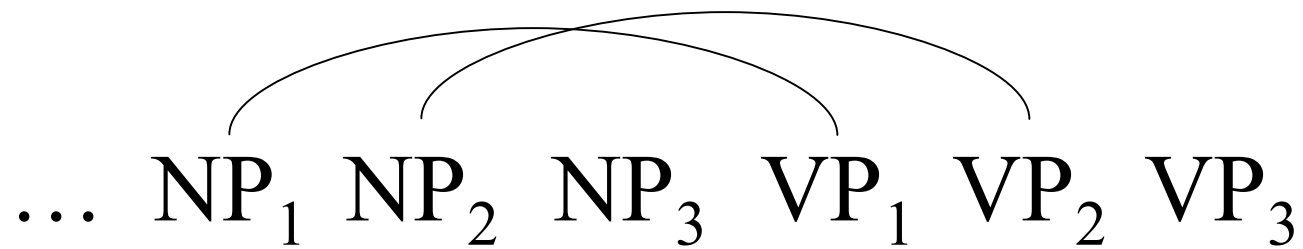
Nested vs. Cross-serial dependencies

Cross-serial (Dutch)



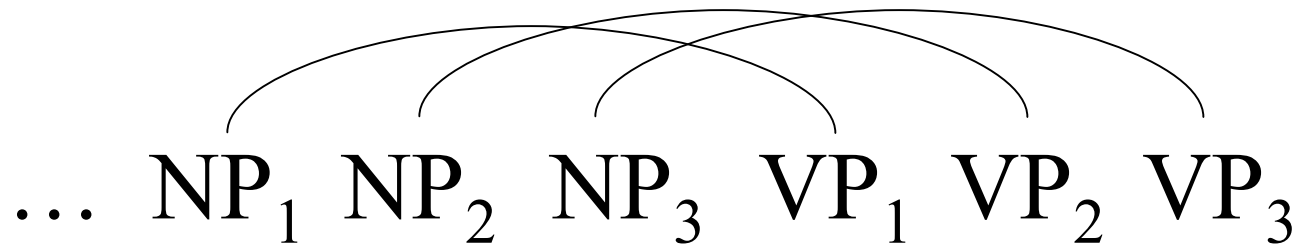
Nested vs. Cross-serial dependencies

Cross-serial (Dutch)



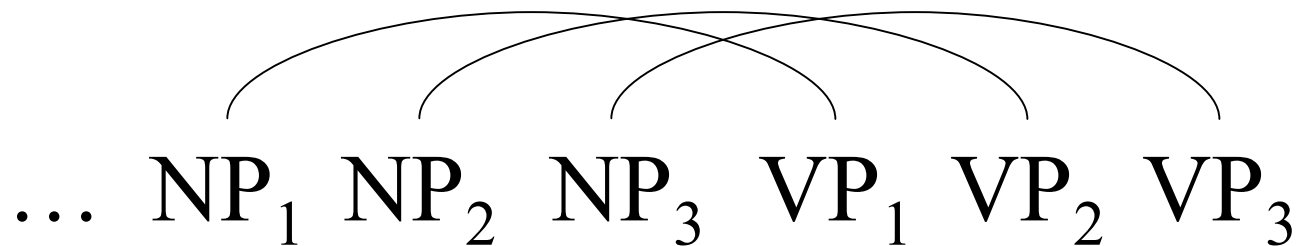
Nested vs. Cross-serial dependencies

Cross-serial (Dutch)



Nested vs. Cross-serial dependencies

Cross-serial (Dutch)



Jeanine heeft de mannen Hans die paarden helpen leren voeren.

Joanna has the men Hans the horses helped teach feed

“Joanna helped the men teach Hans to feed the horses.”

Open question: How to quantify distance?

- Words / syllables? (Gibson, 1998; Hawkins, 1994)
- Discourse structure? (Gibson, 1998; 2000; Warren & Gibson 2002)
 - New discourse referents (new nouns, verbs)
- Intervening interfering NPs? (Gordon et al. 2001, 2002; cf. Lewis)
 - Similar NPs (syntactic, semantic) cause more interference

Nested pronoun generalization

(Bever, 1974; Kac, 1981)

The low complexity of doubly-nested examples like:

- (1) The reporter who everyone that I met trusts said the president won't resign yet. (Bever, 1974)
- (2) A book that some Italian who I've never heard of wrote will be published soon by MIT Press. (Frank, 1992)

Cf. (3) # The reporter who the senator who John met trusts said that the president won't resign yet.

The relative ease of (1) and (2) can be explained by a discourse-based distance metric or an interference-based distance metric.

Null contexts:

The low complexity of nested pronouns

Experiment 6 Materials (Warren & Gibson, 2002)

First / second person pronoun

The reporter who the senator who you met attacked disliked the editor.

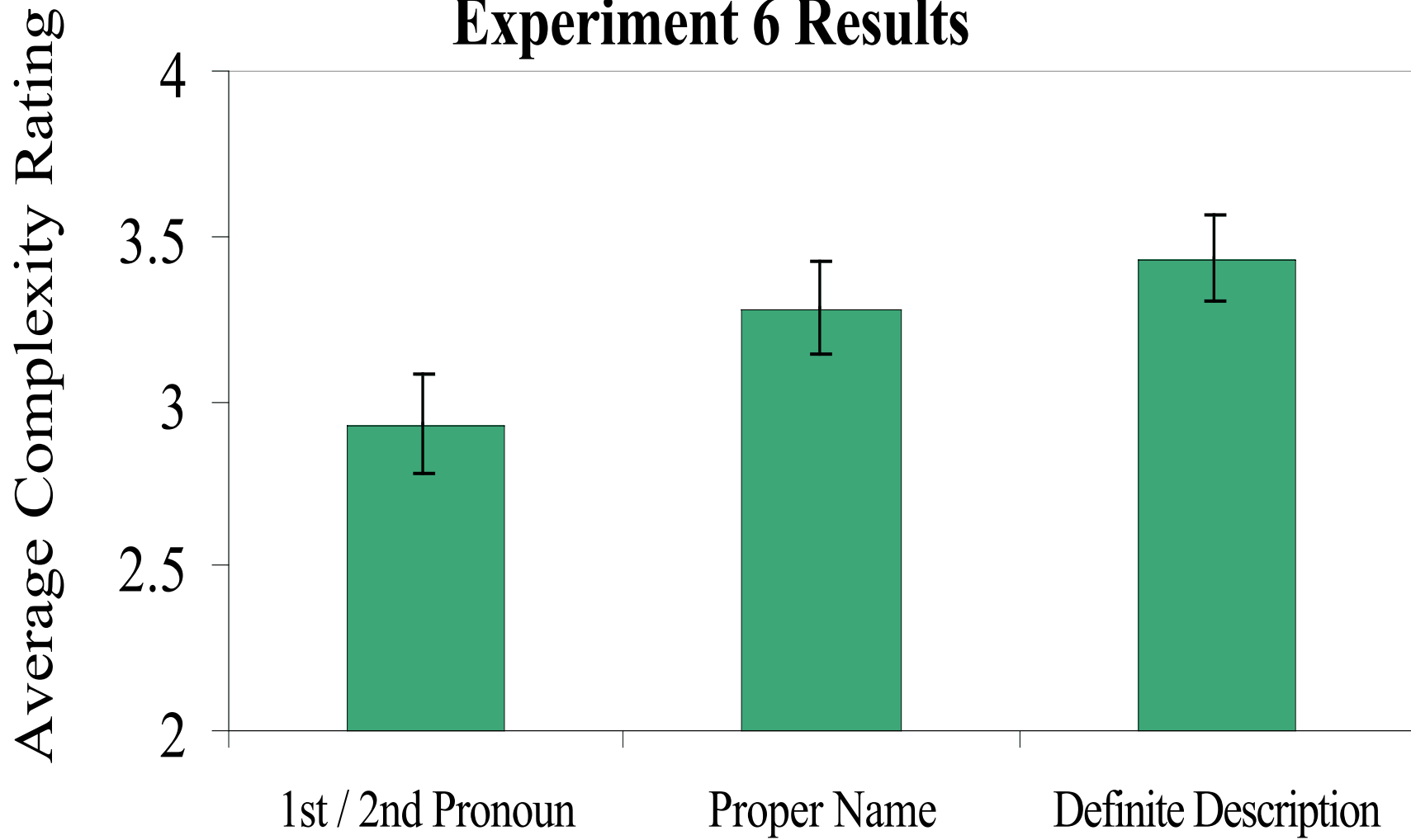
Proper name

The reporter who the senator who John met attacked disliked the editor.

Definite description

The reporter who the senator who the professor met attacked disliked the editor.

Experiment 6 Results



A discourse-based distance metric

(1) The reporter who the senator who **you** met trusts said that the president won't resign yet.

(2) # The reporter who the senator who **John** met trusts said that the president won't resign yet.

The relative ease of (1) vs. (2) can be explained in terms of a discourse-based distance metric: count the number of new NPs and verbs between two elements that need to be integrated.

Maximal integration cost in both sentences is at "trusts": 2 integrations: a) between "the senator" and "trusts"; and b) between "who" and "trusts".

A discourse-based distance metric

- (1) The reporter who the senator who **you** met trusts said that the president won't resign yet.
- (2) # The reporter who the senator who **John** met trusts said that the president won't resign yet.

In (1):

Integrating "trusts" to "senator": 2 new NPs/verbs ("trusts", "met") since "the senator".

Integrating "trusts" to "who": 3 new NPs/verbs ("trusts", "met", "senator") since "who".

Total cost: $2+3 = 5$ units.

In (2):

Integrating "trusts" to "senator": 3 new NPs/verbs ("trusts", "met", "**John**") since "the senator".

Integrating "trusts" to "who": 4 new NPs/verbs ("trusts", "met", "**John**", "senator") since "who".

Total cost: $3+4 = 7$ units.

Application to technical writing

- To help the reader, keep syntactic dependencies close together.

A second resource factor:
Syntactic storage

Syntactic predictions: processing cost for each head that is required to complete the current string as a grammatical sentence

- (1) The reporter claimed that **the baseball player would hold out for more money.**
- (2) The reporter's claim that **the baseball player would hold out for more money** turned out to be true.

Syntactic storage

- Syntactic predictions: processing cost for each head that is required to complete the current string as a grammatical sentence

E.g., after processing the word “the” sentence initially, the parser predicts 2 syntactic heads: a noun and a verb.

Syntactic storage

(1) 1 predicted verb:

The employee realized that the boss implied that **the company planned a layoff** and so he sought alternative employment.

(2) 2 predicted verbs:

a. The employee realized that the implication that **the company planned a layoff** was not just a rumor.

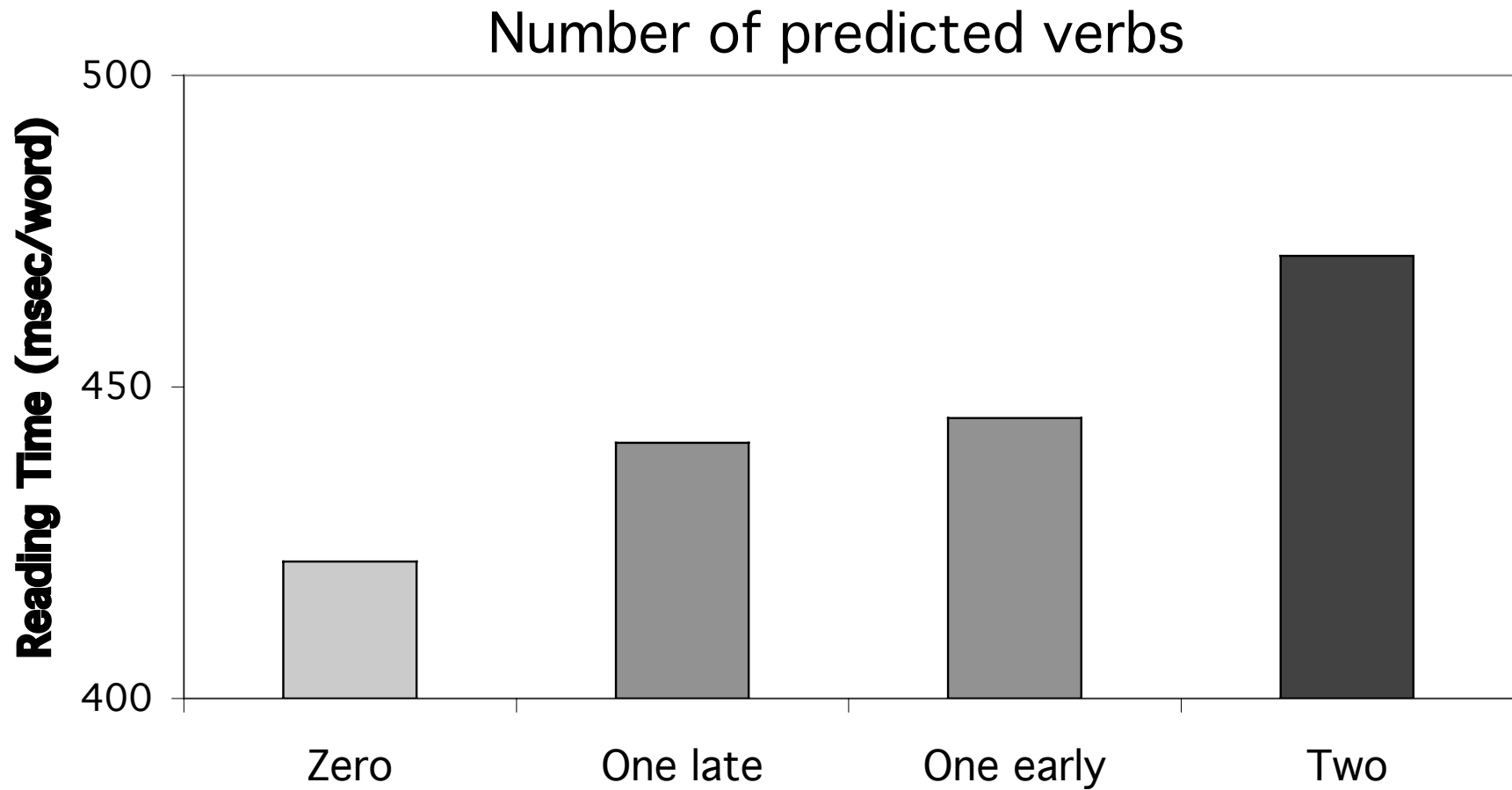
b. The realization that the boss implied that **the company planned a layoff** caused a panic.

(3) 3 predicted verbs:

The realization that the implication that **the company planned a layoff** was not just a rumor caused a panic.

RTs in the bold region are fastest in (1), intermediate in (2), and are slowest in (3). (Chen, Gibson & Wolf, 2005)

Gibson, Chen & Wolf (2002):
Experiment 1 results



Chen, Gibson & Wolf (2005)

The results from Experiment 1 indicate that there is storage cost associated with predicted verbs / incomplete verbal dependencies.

Predictions of other categories associated with storage cost?

Chen, Gibson & Wolf (2005)
Experiment 2

Do predictions of empty categories in wh-dependencies incur storage costs?

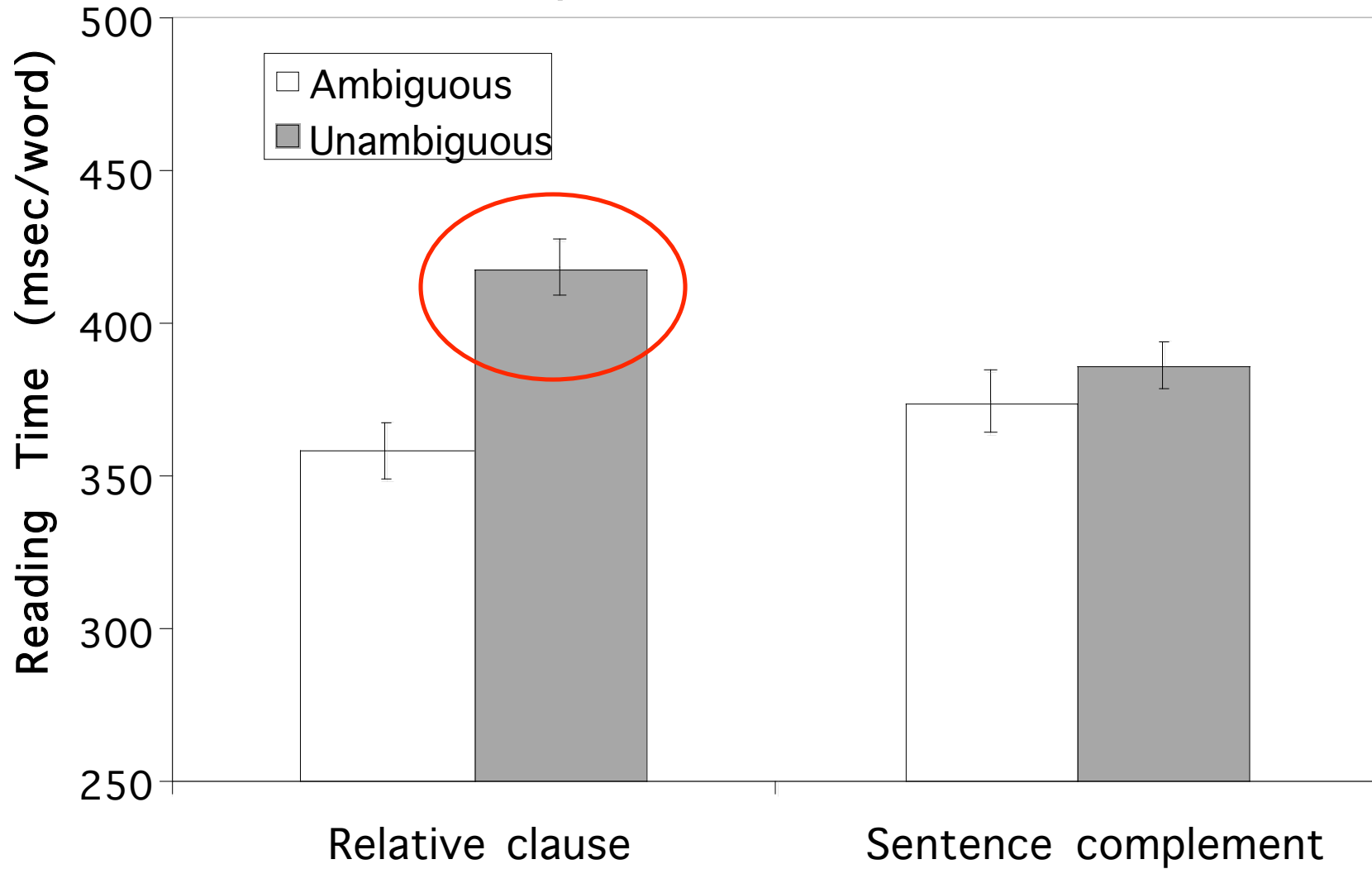
Sentential complement of a noun:

The claim (alleging) that **the cop who the mobster attacked** ignored the informant might have affected the jury.

Relative clause modifying a noun:

The claim which / that **the cop who the mobster attacked** ignored might have affected the jury

Chen, Gibson & Wolf (2005) Experiment 2 results



Chen, Gibson & Wolf (2005)
Experiment 2 results

- Unambiguous RCs were read slower than any of the other three conditions
- This result suggests that predicted empty categories in wh-dependencies incur storage costs.

Chen, Gibson & Wolf (2005)

Potential explanations of the pattern of data:

- Incomplete clauses? No: Expt 2 & 3 results.
- Incomplete dependencies? OK for these data
 - Incomplete thematic role assignments
- Predicted syntactic heads? OK for these data

Ambiguity resolution

- Minimize integration distances
- Minimize storage costs

- Small differences: easy ambiguity
- Big differences: hard ambiguity

Ambiguity resolution: Integration

(1) The bartender told the detective that the suspect left the country yesterday.

Yesterday is preferred as modifying *left* rather than *told*

Ambiguity resolution: Storage costs

Small (no) difference:

The desert trains young people to be tough.

The desert trains are tough on young people.

(Both readings involve local integrations of “trains”.)

Noun-noun reading of “the desert trains”:

one category needed to form a sentence : a verb

Noun-verb reading of “the desert trains”:

one category needed to form a sentence : a noun

Ambiguity resolution: Storage costs

- Big difference:
The cotton clothing is made of comes from Mississippi.

Noun-noun reading of "the cotton clothing":
one category needed to form a sentence : a verb

Relative clause reading of "the cotton clothing": three
categories needed: two verbs and a position in the relative
clause for "cotton".

Grodner, Gibson & Tunstall (2002): Noun-noun (NN) / Relative clause (RC) ambiguity

Item set 1: Plausibility and frequency factors were balanced between the RC and the NN:

The alley (which) mice run rampant in is damp and dimly lit but relatively clean.

Grodner, Gibson & Tunstall (2002): Noun-noun (NN) / Relative clause (RC) ambiguity

Item set 2: Plausibility and frequency factors were biased strongly for the RC:

The tool (which) plumbers need to have is a good monkey wrench for loosening rusty pipes.

NN/RC (Grodner et al. 2002)

Graph removed for copyright reasons.

NN/RC (Grodner et al. 2002)

Ambiguity effects for both the balanced items and for the NN-biased items

Conclusion: storage cost is an important factor in resolving ambiguity

Storage costs in the main-verb (MV) / reduced-relative (RR) ambiguity

The defendant examined ...

MV ...the evidence.

RR ... by the lawyer turned out to be unreliable

Integration costs: both the MV and RR are local integrations.

Storage costs in the MV/RR ambiguity

The defendant examined ...

Storage costs: no difference

MV structure: 1 head is required:
a noun (object of the verb "examined")

RR structure: 1 head is required:
the matrix verb

Storage costs in the MV/RR ambiguity

The defendant examined ...

Integration and storage costs: no differences

Therefore, lexical frequency and plausibility information play a major role in this ambiguity

Storage costs in the MV/RR ambiguity

MV/RR ambiguity embedded within an RC:

The witness who the defendant examined ...

MV: ...turned out to be unreliable.

RR: ...by the lawyer implicated turned out to be unreliable.

Storage costs in the MV/RR ambiguity

The witness who the defendant examined ...

Storage costs:

MV: 1 head is required:
the matrix verb.

RR: 3 heads are required:

(1) the embedded verb ("implicated"), (2) a gap-site for the wh-pronoun "who", (3) the matrix verb

Storage costs in the MV/RR ambiguity

The witness who the defendant examined ...

Storage costs: 3 heads vs. 1 head: MV preferred

Control ambiguity:

The witness said that the defendant examined ...

Storage costs: 1 head vs. 1 head: no storage cost preference

Storage costs in the MV/RR ambiguity

Grodner, Gibson & Tunstall (2002)

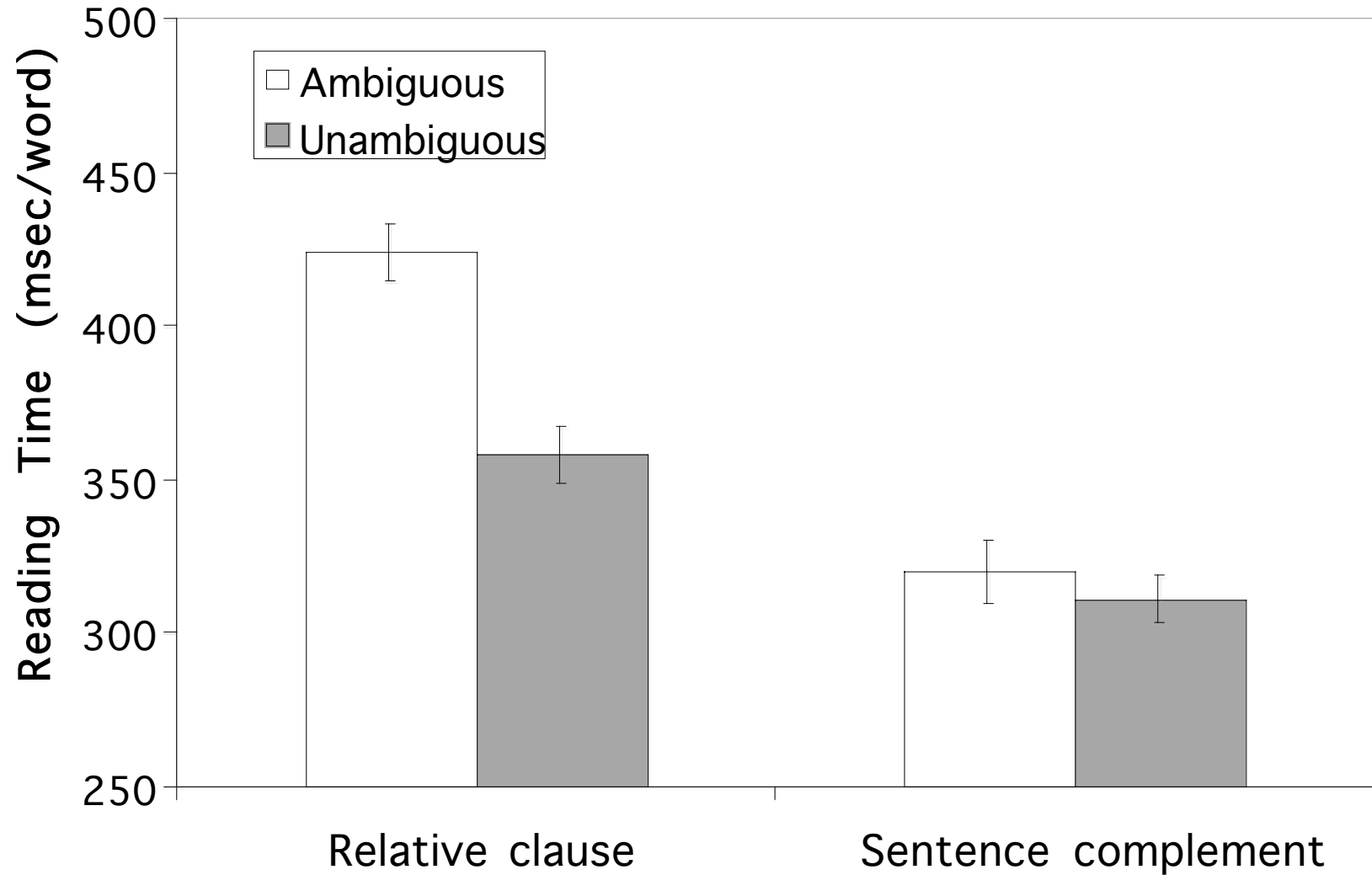
MV/RR embedded in relative clause:

The witness who the evidence (that was) examined by the lawyer implicated turned out to be unreliable.

MV/RR embedded in a sentence complement:

The witness said that the evidence (that was) examined by the lawyer was unreliable.

Reading times at the disambiguating region "by the lawyer"



MV/RR (Grodner et al., 2002)

Graph removed for copyright reasons.

Another test of resource theories: Chinese Relative Clauses

Hsiao & Gibson (2003, *Cognition*)

Observation:

Subject-extracted RCs are less complex than object-extracted RCs in English:

The reporter [who the senator attacked] admitted the error.

The reporter [who attacked the senator] admitted the error.

RC complexity processing theories

- Resources: more resources for object-extracted RCs than subject-extracted RCs, both integration and storage.
- Accessibility of syntactic positions: Subject position is more accessible than object position (Keenan & Comrie, 1977)

Test case: Chinese RCs

Chinese: Subject-Verb-Object word order, with RCs before the head noun:

Object-extraction:

[Subject-NP V e Comp] NP V NP

[the senator attacked who] the reporter admitted the error.

“The reporter who the senator attacked admitted the error.”

Subject-extraction:

[e V Object-NP Comp] NP V NP

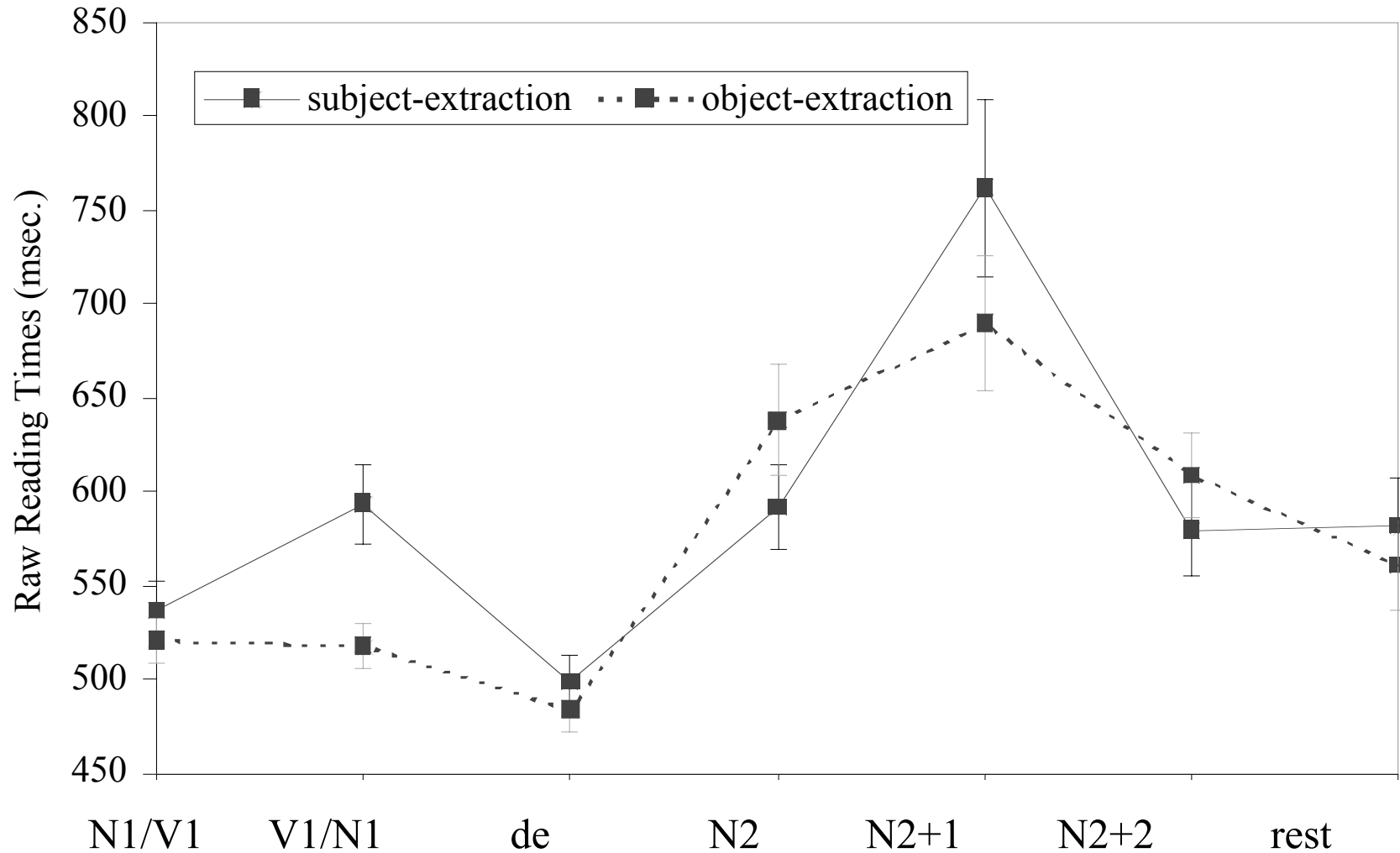
[attacked the senator who] the reporter admitted the error.

“The reporter who attacked the senator admitted the error.”

RC complexity processing theories: Predictions for Chinese RCs

- Resources: more resources for subject-extracted RCs than object-extracted RCs: both integration and storage.
- Accessibility of syntactic positions: Subject-extractions easier than object-extractions

Hsiao & Gibson (2003): Results for singly-embedded structures



Processing Chinese RCs: Conclusions

- Resource theories account for the pattern of data:
Both storage and integration
- Accessibility makes the wrong predictions