

Language comprehension

Lecture 4: Ambiguity resolution:
Resources; structural frequencies

9.591; 24.945

October 4, 2004

Ted Gibson

9.591 Course so far

- Lecture 1: Experimental methods; Informational constraints affecting sentence comprehension: Lexical frequency, plausibility, context and syntax; Modularity in sentence comprehension.
- Lecture 2: Resources and sentence complexity. The complexity of unambiguous sentences.
- Lecture 3: Working memory and sentence comprehension.
- **Lecture 4:** Ambiguity resolution: Resources; structural frequencies.

Sentence processing: Recap thus far

- Multiple factors are involved in processing unambiguous sentences and in ambiguity resolution: (a) Syntactic structure; (b) Lexical frequency; (c) Plausibility; (d) Discourse context; (e) Intonation; (f) all information sources are constrained by the amount of working memory resources that are available.
 - How exactly the factors are represented and processed are open questions: Minimal Attachment & Late Closure vs. Storage and Integration (DLT)?
Today: Evidence supporting the DLT in ambiguity resolution
 - **Frequency: What kind of frequency is the human sentence processor sensitive to?**
- What is the time course of information integration?
 - Modular (syntax-first)? Or non-modular?
- **Is the parser serial or constrained parallel?**

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 / Expectations: Predicting categories to complete the current structure.

Recap: Locality effects in unambiguous structures
Grodner & Gibson (in press)

Experiment 1:

Object-extracted relative clause:

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

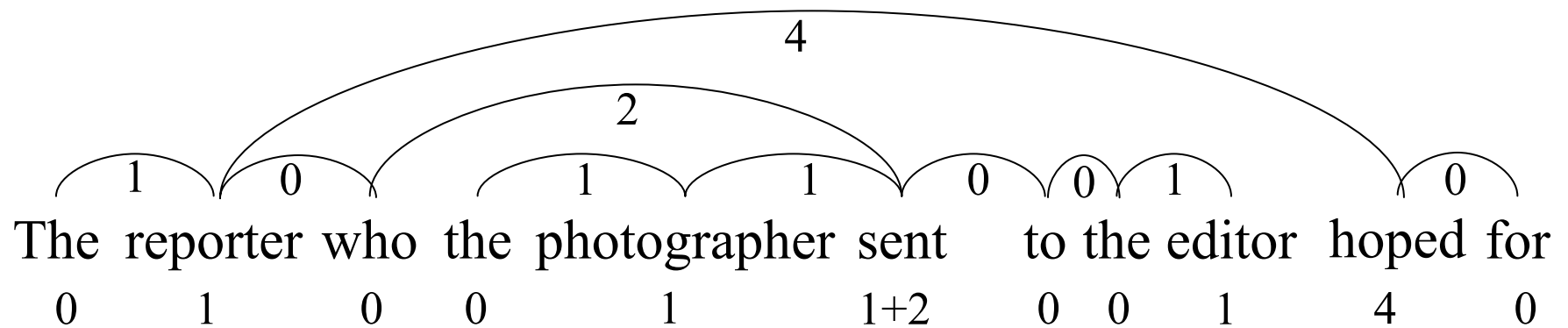
Subject-extracted relative clause:

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

Locality effects in unambiguous structures

Object-extracted relative clause:

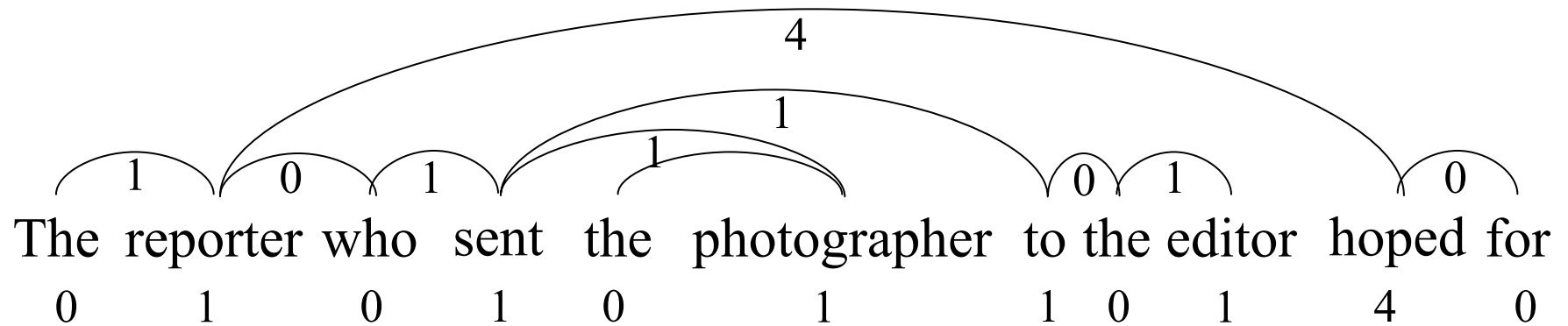
The reporter who the photographer sent to the editor hoped for a story.



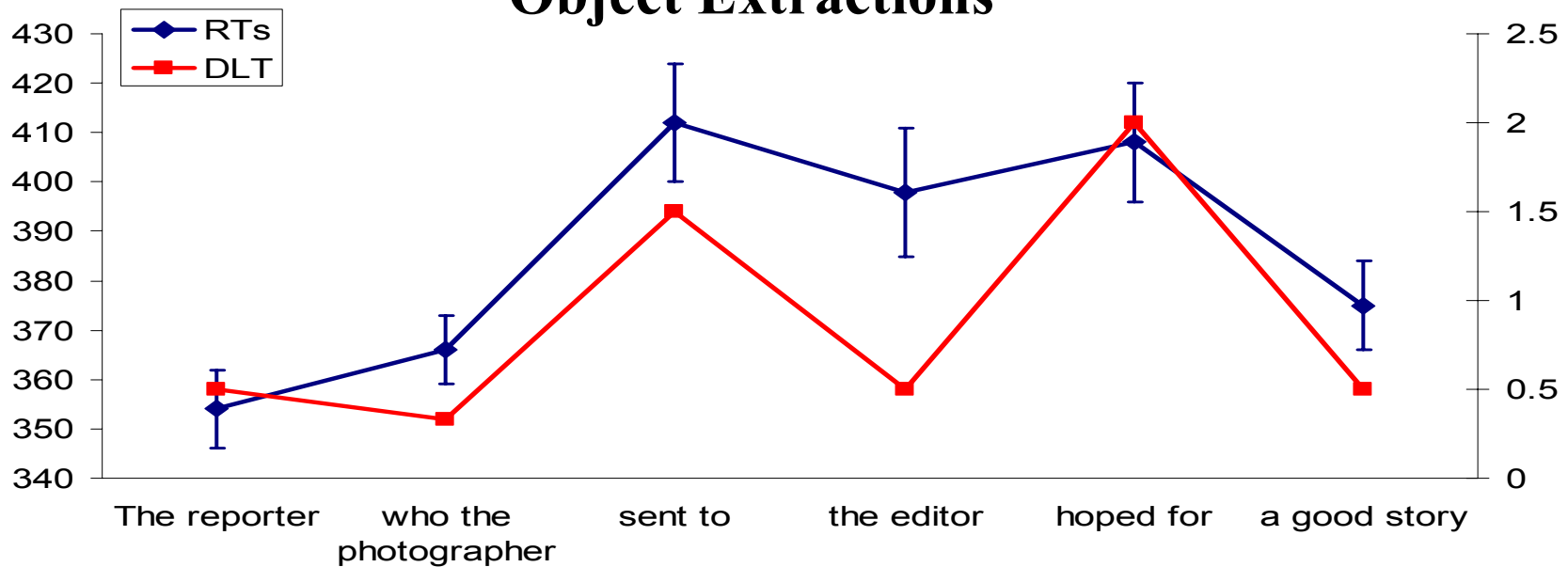
Locality effects in unambiguous structures

Subject-extracted relative clause:

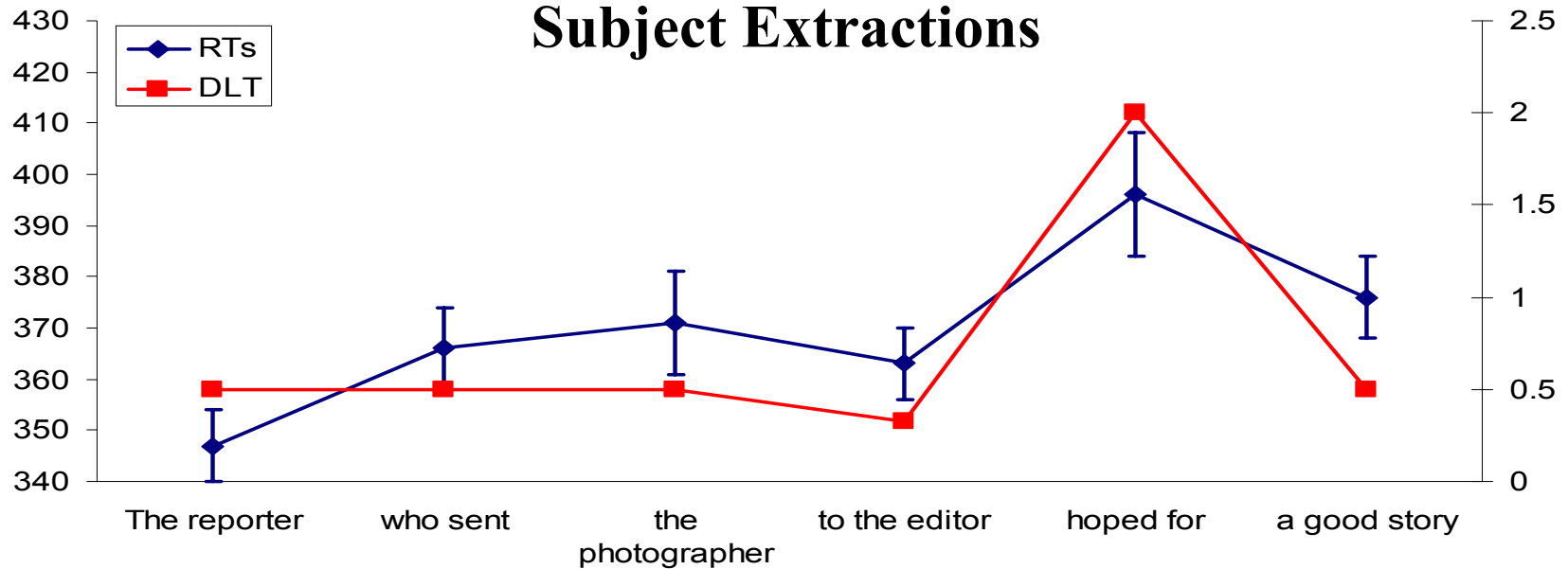
The reporter who sent the photographer to the editor hoped for a story.



Object Extractions



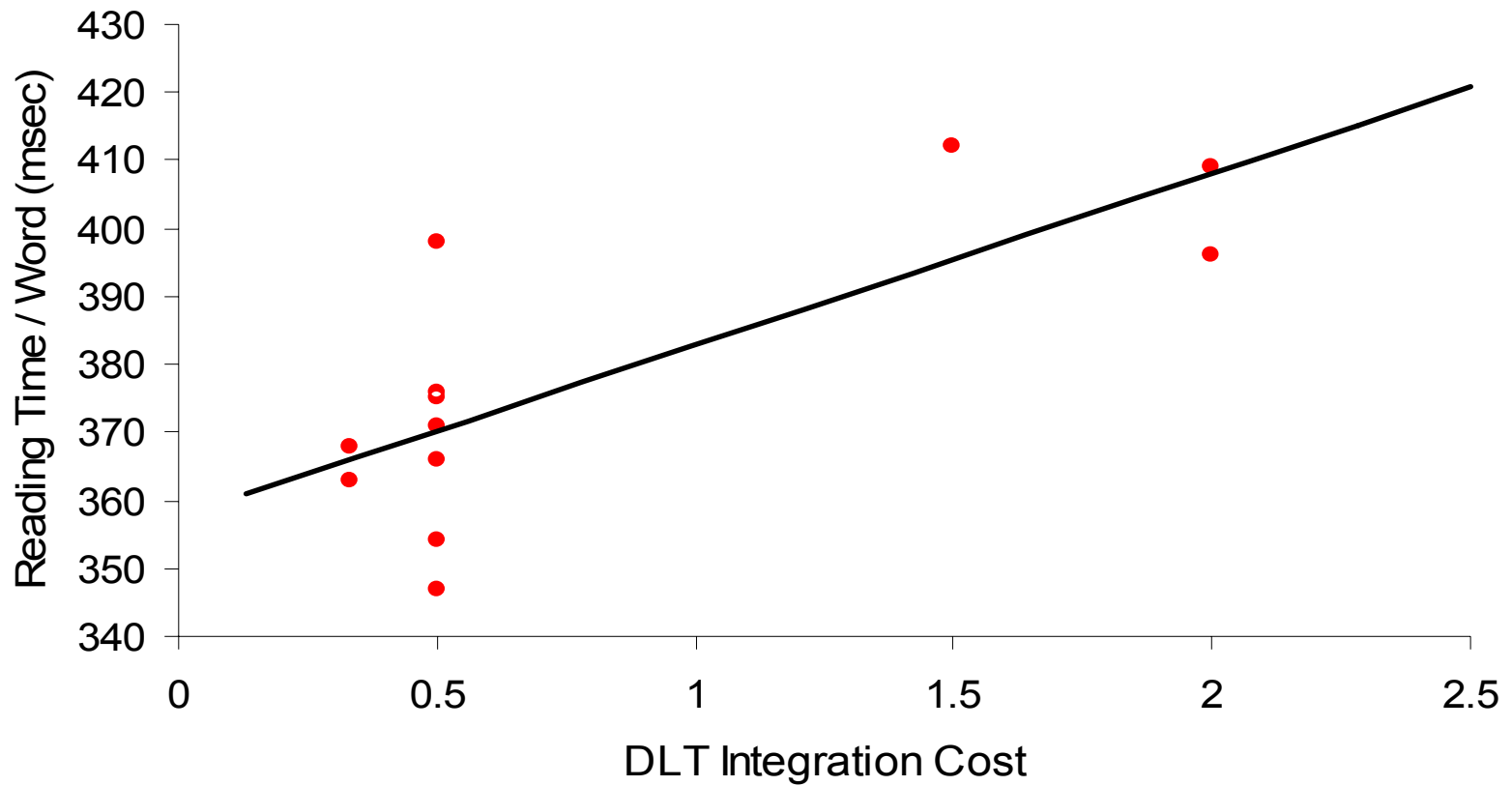
Subject Extractions



Experiment 1: DLT vs. RTs

(Grodner & Gibson, in press)

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



Grodner & Gibson (in press) results

- RTs are long at points of long dependencies
- Note that the data are not consistent with a “rare phrase structure rule” hypothesis (MacDonald & Christiansen, 2002; Hale, 2002, 2003)
- Such a hypothesis predicts RTs to increase when an uncommon PS rule is encountered: i.e., at the beginning of the object-extracted RC.
- No increase in RTs there: Only when the verb is encountered.
- Therefore, integration complexity cannot be reduced to phrase structure rule rarity.

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

- (1) The employee realized that the boss implied that **the company planned a layoff** and so he sought alternative employment.
- (2) The employee realized that the implication that **the company planned a layoff** was not just a rumor.
- (3) 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, in press)

Chen, Gibson & Wolf (in press): Experiment 1 results

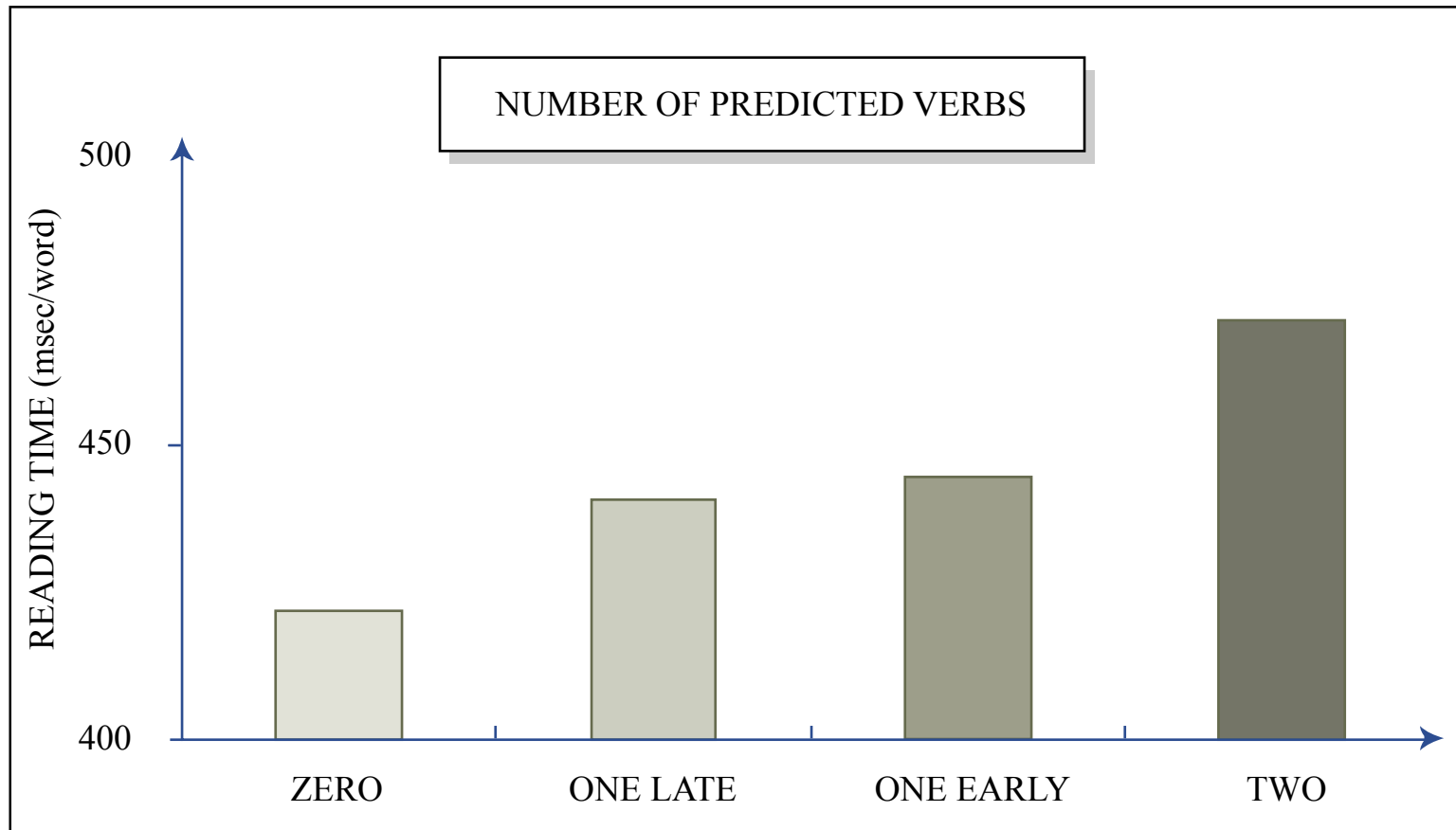


Figure by MIT OCW.

Chen, Gibson & Wolf (in press)

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

Predictions of other categories associated with storage cost?

Chen, Gibson & Wolf (in press)

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 (in press) Experiment 2 results

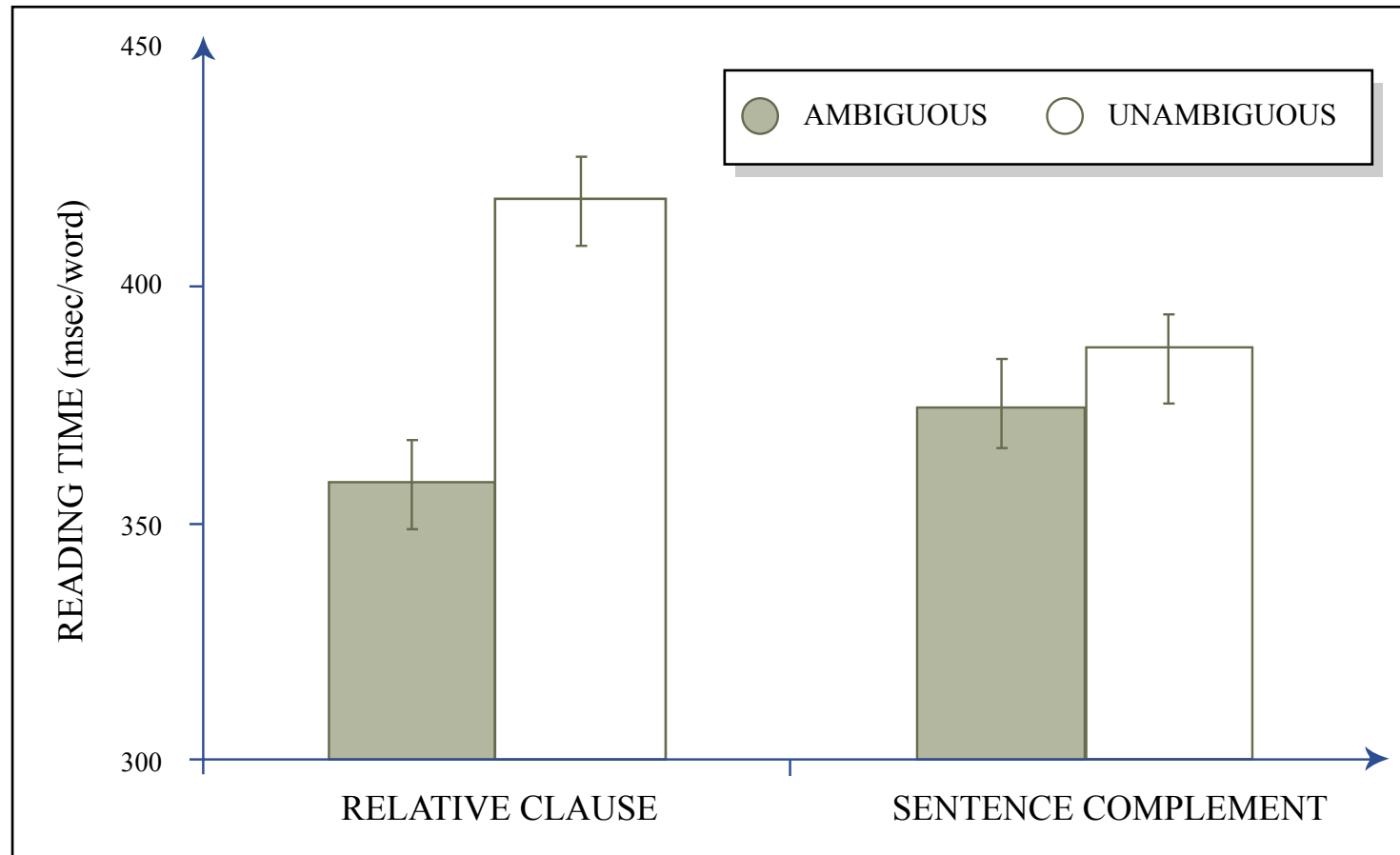


Figure by MIT OCW.

Chen, Gibson & Wolf (in press)

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.
- Experiment 3 replicated Expt 2, on simpler materials:
RC Structure:
The announcement [which **the baker from a small bakery in New York City** received ____] helped the business of the owner.
SC Structure:
The announcement [that **the baker from a small bakery in New York City** received the award] helped the business of the owner.

Chen, Gibson & Wolf (in press)

Potential explanations of the pattern of data:

- Incomplete clauses? No: Expts 2 & 3 results.
- Incomplete dependencies? OK for these data
 - Incomplete thematic role assignments
- Predicted syntactic heads? OK for these data
- Data from the processing of head-final languages strongly support the predicted-head view (German: Konieczny, 2000; Hindi: Vasishth, 2002; Japanese: Nakatani & Gibson, 2003)

Chen, Gibson & Wolf (in press)

Taken together with the processing results from the literature, the results of Gibson, Chen & Wolf therefore support a syntactic-head prediction theory of syntactic storage over and incomplete-dependency theory.

The results of Experiment 2 can only be accounted for with the inclusion empty categories mediating long-distance dependencies.

Therefore, these results provide **processing evidence in support of the existence of wh-traces in wh-dependencies.**

Ambiguity resolution

- Minimize integration distances
- Minimize storage costs

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

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 biased strongly for the RC:

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

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.

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

NN continuations:

(The) tool plumbers need to have big toolboxes because unforeseen problems often arise on the job.

(The) alley mice run rampant in dark streets because there isn't much traffic there to scare them away.

Application of DLT to the MV/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.

Application of DLT to 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

Application of DLT to 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

Application of DLT to 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.

Application of DLT to 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

Application of DLT to 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

Application of DLT to 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"

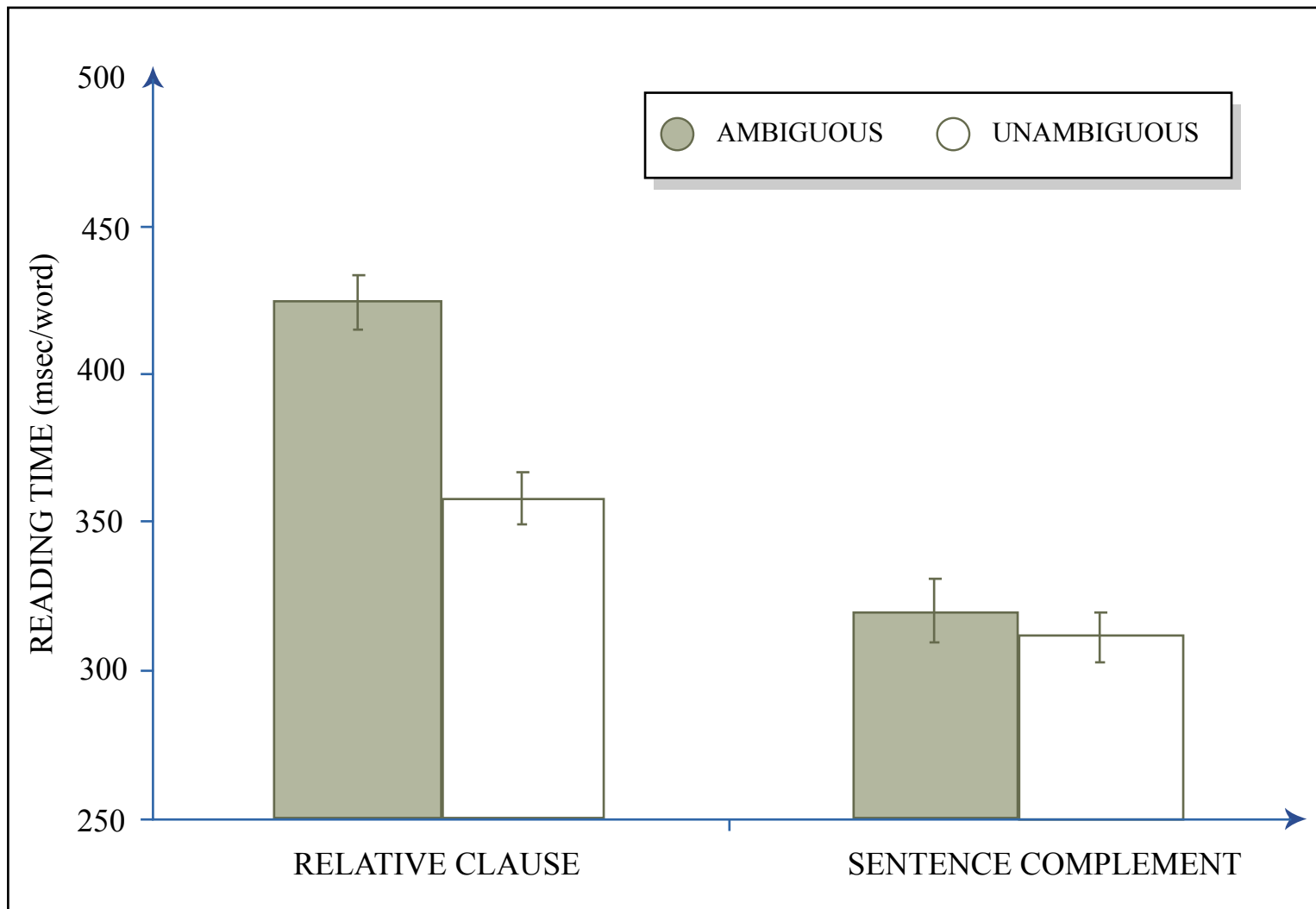


Figure by MIT OCW.

MV/RR (Grodner et al., 2002)

Graph removed for copyright reasons.

Conclusions so far

Storage and integration apply in accounting for the processing of both unambiguous and ambiguous structures, including:

- On-line reading times in unambiguous sentences
- The complexity of nested structures cross-linguistically
- Heaviness effects: putting heavy constituents at the end
- Numerous ambiguity effects cross-linguistically

Sentence processing: Recap thus far

- Multiple factors are involved in processing unambiguous sentences and in ambiguity resolution.
 - How exactly the factors are represented and processed are open questions: Minimal Attachment & Late Closure vs. Storage and Integration (DLT)?
Evidence supporting the DLT in ambiguity resolution
 - **Frequency: What kind of frequency is the human sentence processor sensitive to?**
- What is the time course of information integration?
 - Modular (syntax-first)? Or non-modular?
- **Is the parser serial or constrained parallel?**

The relationship between the frequency and the complexity of a syntactic structure

Ted Gibson

Department of Brain & Cognitive Sciences

Department of Linguistics & Philosophy

MIT

To do today:

1. The importance of lexical frequencies.
E.g., MacDonald et al. (1994); Trueswell (1996)
2. Cross-linguistic similarities and differences in modifier attachment preferences: The tuning hypothesis (Mitchell et al., 1996)
3. Next time: Contingent frequencies? (Tabor et al., 1997)

Cross-linguistic similarities and differences in modifier attachment preferences

Ambiguous two-NP-site relative clause attachments
(Cuetos & Mitchell, 1988)

El periodista entrevistó a [NP₁ la hija del
[NP₂ coronel]] [CP que tuvo el accidente]

The journalist interviewed [NP₁ the daughter of
[NP₂ the colonel]] [CP who had had the accident]

Cuetos & Mitchell (1988): Off-line evidence

- English: 60% NP₂ attachment (low attachment)
- Spanish: 60% NP₁ attachment (high attachment)

Cuetos & Mitchell (1988), Mitchell & Cuetos (1991):
On-line evidence

(1) Alguien disparó contra el criado de la actriz [que estaba en el balcón **con su marido**].

“Someone shot the servant (masc) of the actress (fem) who was on the balcony with her husband.”

(2) Alguien disparó contra la criada del actor [que estaba en el balcón **con su marido**].

“Someone shot the servant (fem) of the actor (masc) who was on the balcony with her husband.”

con su marido (with her husband) is processed faster in (2) than (1).

NP₁ attachments are easier than NP₂ attachments in Spanish.

Theories

- Late Closure (Frazier, 1979) or Locality (Gibson, 1998):
no good on its own
- Construal: Late Closure plus Gricean implicature (Frazier & Clifton, 1995)
- Tuning (Mitchell et al., 1996)
- Locality and Predicate Proximity (Gibson et al. 1995)
- Locality and Anaphoric binding (Hemforth et al., 2000)
- Implicit Prosody (Fodor, 1998)

Theory 1: (Cuetos & Mitchell, 1988)

Maybe Spanish speakers favor non-local attachment because modifiers are generally post-head in Spanish, including adjectives.

So speakers are used to attaching the RC to the initial NP in a sequence NP-modifier-RC when the modifier is an adjective.

Theory 1: (Cuetos & Mitchell, 1988)

Evidence against this hypothesis:

Dutch and German have pre-head adjectives, but there is still a non-local attachment preference in NP1-Prep-NP2-RC ambiguities in these languages (Brysbaert & Mitchell, 1995; Hemforth et al., 1998).

Theory 2: Construal + Locality (Frazier & Clifton, 1995)

In addition to locality, there is a second factor at play in English:

A Gricean Maxim: The Maxim of Manner (Clarity / Avoid Ambiguity):

“Speakers should avoid ambiguity by being clear and to the point.”

English speakers might favor local attachment (to “colonel”) because if the speaker wanted the RC to attach to the non-local site (“daughter”), he/she could have used the Anglo-Saxon genitive form (e.g., “the colonel's daughter”) which allows unambiguous attachment to “daughter”.

Theory 2: Construal (Frazier & Clifton, 1995)

Spanish doesn't have an Anglo-Saxon genitive form, so the Gricean principle doesn't apply there.

Evidence against this hypothesis: Dutch has an Anglo-Saxon genitive, and there is still a non-local attachment preference in Dutch (Brysbaert & Mitchell, 1995).

But: the Anglo-Saxon genitive is very rare in Dutch, so it's not clear that this argument has weight.

Theory 3: Tuning hypothesis (Mitchell et al., 1996)

People tabulate resolutions of ambiguities.

More frequently occurring resolutions are easier to process.

Analyses of two-site RC attachments:

Spanish: 60% attach to the high site;

English: 38% attach to the high site.

Theory 3: Tuning hypothesis (Mitchell et al., 1996)

Grain-size questions: What is being tuned? Does the category matter? Intervening material? Preceding context?

Begging the question? Why do all human languages tend to have a locality bias in multiple VP attachments?

No: a production theory might explain this observation. Tuning is for comprehension. (MacDonald & Seidenberg, 1999)

Theory 4: Two factors: Recency (locality) and Predicate Proximity (Gibson et al., 1996)

Relevant observation: Always a local preference for VP attachments:

Juan dijo que Bill se murió (# morirá) ayer.
"John said Bill died (# will die) yesterday."

Theory 4: Two factors: Recency (locality) and Predicate Proximity (Gibson et al., 1996)

Three-NP-site ambiguities:

NP₁ Prep NP₂ Prep NP₃ RC

las lámparas cerca de las pinturas de la casa que fue dañada en la inundación

las lámparas cerca de la pintura de las casas que fue dañada en la inundación

la lámpara cerca de las pinturas de las casas que fue dañada en la inundación

“the lamp(s) near the painting(s) of the house(s) that was damaged in the flood”

Gibson et al., (1996): Experiment 1 Cumulative grammaticalities

Graph removed for copyright reasons.

Gibson et al., (1996): Experiment 1 RTs

Graph removed for copyright reasons.

Gibson et al., (1996): Experiment 2 Cumulative grammaticalities

Graph removed for copyright reasons.

Gibson et al., (1996): Experiment 2 RTs

Graph removed for copyright reasons.

Theory 4: Two factors: Recency (locality) and Predicate Proximity (Gibson et al., 1996)

Results from on-line and off-line experiments in English and Spanish: Low preference, then high, then middle in both languages.

Theory: Two factors: recency (locality, see DLT) plus predicate proximity

Predicate Proximity specifies a fixed cost associated with attachments which are not as close as possible to a predicate phrase (typically a VP).

Theory 4: Two factors: Recency (locality) and Predicate Proximity (Gibson et al., 1996)

Furthermore, the relative strength of Predicate Proximity is proposed to vary across languages, and that it is weak in English but strong in Spanish.

It is claimed to be derivable from the distance of arguments to the predicate in a language: Larger distance (Spanish), leads to a stronger predicate activation (predicate proximity).

Note: "Predicate Proximity" is essentially a stipulation.

Gibson, Pearlmutter & Torrens (1999)

Two sites, high attach: El astrónomo predijo la órbita de los planetas que se observó desde el satélite

Two sites, low attach: El astrónomo predijo las órbitas del planeta que se observó desde el satélite

Three sites, high attach: El astrónomo predijo el cambio de las órbitas de los planetas que se observó desde el satélite

Three sites, middle attach: El astrónomo predijo los cambios de la órbita de los planetas que se observó desde el satélite

Three sites, low attach: El astrónomo predijo los cambios de las órbitas del planeta que se observó desde el satélite

“The astronomer predicted (the changes of) the orbits of the planets that was observed from the satellite.”

Gibson, Pearlmutter & Torrens (1999)

Graph removed for copyright reasons.

Gibson, Pearlmutter & Torrens (1999)

Results: Two sites: High preference

Three sites: Low preference, then high, then middle.

Therefore the observed differences are not due to lexical effects.

Attachments to 3 VPs (Pearlmutter & Gibson, 2001)

The grandmother claimed that the fireman said that the arsonists set the fire \ to get herself into the news last week.

The grandmother claimed that the fireman said that the arsonists set the fire \ to get himself into the news last week.

The grandmother claimed that the fireman said that the arsonists set the fire \ to get themselves into the news last week.

Attachment preferences: Low, then Middle, then High.

This supports Predicate Proximity.

Theory 5: Anaphoric binding (Hemforth et al., 2000)

Two factors: Locality, and anaphoric binding

Anaphoric binding: The parser initiates a search for the appropriate referent for a pronoun when the pronoun is first encountered.

This process influences RC attachment because a search is initiated for the relative pronoun heading an RC (e.g., "who", "which", "that", "que") just as for any other pronoun.

Claim: pronoun-binding: Look at most focused positions first (something like predicate proximity).

Theory 5: Anaphoric binding (Hemforth et al., 2000)

Evidence:

(a) Die Tochter der Lehrerin, die aus Deutschland kam, traf John.

[The daughter of the teacher who came from Germany met John.]

(b) Die Tochter der Lehrerin aus Deutschland traf John.

[The daughter of the teacher from Germany met John.]

NP1-Prep-NP2-RC ambiguity: Non-local attachment preference for RC.

NP1-Prep-NP2-PP2 ambiguity: Local attachment preference for the PP.

(Minimal pairs: The RC is just the PP with “that was” inserted.)

Theory 5: Anaphoric binding (Hemforth et al., 2000)

To be explained: The local RC attachment preference in English.

Claim: It has something to do with the fact that the lexical pronoun/complementizer can be dropped in some English RCs (object-extracted).

Evidence from Brazilian Portuguese (Edson Miyamoto)

Local RC attachment preference in NP1-Prep-NP2-RC ambiguities.

Miyamoto also looked at the attachment of reduced relatives (e.g., “seen by the boy”): strong local attachment.

Brazilian Portuguese does not allow the complementizer to be dropped in RCs, so the local attachment preference is surprising for the anaphor-binding hypothesis.

Evidence from Brazilian Portuguese (Edson Miyamoto)

Evaluating the Brazilian Portuguese data within the Predicate Proximity framework:

Unlike Spanish, Brazilian Portuguese has rigid SVO word order (like English). So the local attachment is expected.

Testing the tuning hypothesis: Evidence from English conjoined NP attachments (Gibson & Schutze, 1999)

Diagram removed for copyright reasons.
"Figure 1."

Gibson, Schutze & Salomon (1996)

RC attachments to three sites are relatively rare in natural corpora, so they examined conjoined NPs:

Off line complexity rating study:

The salesman ignored a customer with a child with a dirty face and ...

... a wet diaper. [low]

... one with a wet diaper. [middle]

... one with a baby with a wet diaper. [high]

Gibson, Schutze & Salomon (1996)

Table removed for copyright reasons.
"Table 8."

Corpus evidence

No matter how the corpus counts were analyzed, low attachments were most frequent, followed by middle attachments, with high attachments least frequent.

Unfiltered frequencies of all possible categories attaching to one of three preceding NP sites in the Brown corpus

Attaching category	Attachment site		
	NP1	NP2	NP3
ADJP	1	1	305
ADVP	1	5	37
NP	4	5	792
PP	155	261	733
S	17	16	14
SBAR	6	2	150
VP	1	0	144
CC (conjunction)	54	107	357
Totals	239	397	2532

Filtered frequencies of the categories PP, S, SBAR and CC attaching to one of three preceding NP sites in the Brown corpus

Attaching category	Attachment site	
	NP1 (High)	NP2 (Middle)
PP	29	47
S	5	1
SBAR	0	0
CC (conjunction)	22	36
Totals	56	84

Reading time evidence

Subject-NP Verb NP1 Prep1 NP2 Prep2 NP3 and (the) one ...

High attachment: ... Prep3 NP4 Prep4 NP5 ...

Middle attachment: ... Prep4 NP5 ...

The talkshow host told a joke about a man with an umbrella and one ...

High attachment: ... about a woman with a dog but hardly anybody laughed.

Middle attachment: ... with a dog but hardly anybody laughed.

Reading time evidence

High attachment plausibility:

The talkshow host told two jokes: one joke about a man with an umbrella; and a second joke about a woman with a dog.

Middle attachment plausibility:

The talkshow host told a joke about two men: one man with an umbrella; and a second man with a dog.

Comprehension-question response accuracy in Experiment 1

Graph removed for copyright reasons.

Residual reading times in Experiment 1

Graph removed for copyright reasons.

Experiment 2

Same materials as Experiment 1, but with no line breaks before the disambiguating region.

Comprehension-question response accuracy in Experiment 2

Graph removed for copyright reasons.

Residual reading times in Experiment 2

Graph removed for copyright reasons.

Gibson & Schutze (1999)

Why the dissociation between frequency and complexity?

Gibson & Schutze's hypothesis: Locality (e.g., DLT) is at play in both comprehension and production, but Predicate Proximity may play a role only in comprehension.

Desmet & Gibson (2003)

Gibson & Schutze didn't examine the right grain size of tuning.

Investigate the anaphoric binding hypothesis and its predictions for G&S's materials.

Desmet & Gibson (2003)

Table 1. Corpus frequencies of middle and high attachments in the three site conjunction ambiguity.

	Brown corpus		WSJ corpus		Total	
	Mid	High	Mid	High	Mid	High
All Instances	107	54	88	68	195	122
Pronouns	9	4	7	3	16	7
One/Ones	1	0	0	1	1	1

Desmet & Gibson (2003)

Table 2. Corpus frequencies of middle and high attachments in other conjunction ambiguities.

	Brown corpus		WSJ corpus		Total	
	Low	High	Low	High	Low	High
<i>Two site conjunction ambiguity</i>						
All Instances	1508	604	1036	592	2544	1196
One / Ones	2	22	5	8	7	30
<i>All conjunction ambiguities (> 1 attachment site)</i>						
One / Ones	5	29	8	13	13	42

Reading experiment (Desmet & Gibson, 2003)

Attachment site (middle, high) x disambiguation (determiner-noun, pronoun)

A column about a soccer team from the suburbs and ...

Determiner noun:

...an article about a baseball team from the city were published in the Sunday edition.

...a baseball team from the city was published in the Sunday edition.

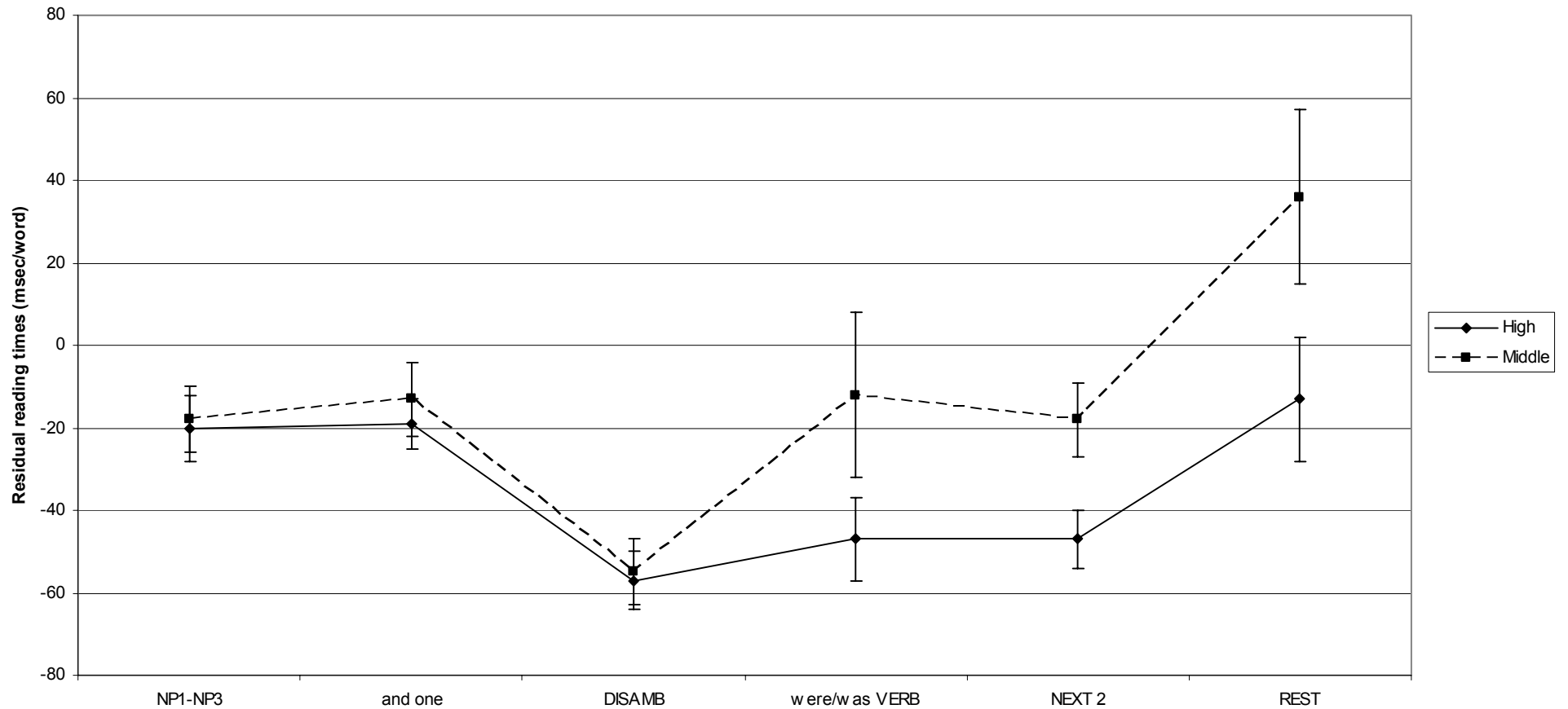
Pronoun:

...one about a baseball team from the city were published in the Sunday edition.

...one from the city was published in the Sunday edition.

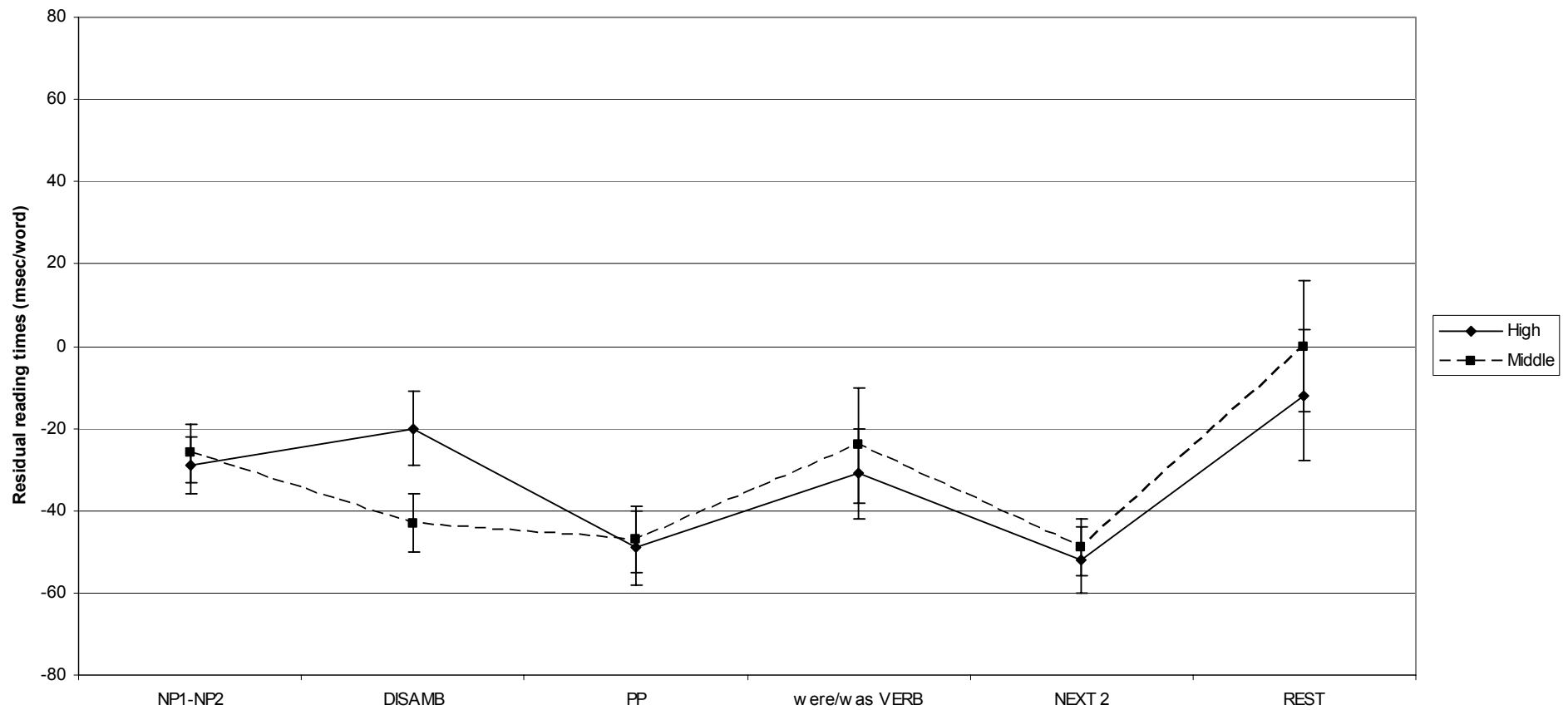
RTs in the "and one" conditions (Desmet & Gibson, 2003)

Figure 2. Reading Times of the "AND ONE" conditions.



RTs in the "and NP" conditions (Desmet & Gibson, 2003)

Figure 3. Reading Times of the "AND NP" conditions.



Desmet & Gibson (2003): Summary

In the sentences containing “and one”, the results of Gibson & Schutze (1999) were replicated: high attachment easier.

In the sentences disambiguated with “and NP”, the middle attachments were easier.

This result disconfirms G&S (1999).

This result is consistent with the tuning hypothesis.

Desmet & Gibson (2003): Summary

This result is consistent with the tuning hypothesis.

But caveat: The result is just another example of a correlation between frequency and complexity.

It does not demonstrate the direction of the causality between the two.

Sentence processing: Recap thus far

- Multiple factors are involved in processing unambiguous sentences and in ambiguity resolution.
 - How exactly the factors are represented and processed are open questions: Minimal Attachment & Late Closure vs. Storage and Integration (DLT)?
Evidence supporting the DLT in ambiguity resolution
 - **Frequency: What kind of frequency is the human sentence processor sensitive to? We don't know yet.**
Generalization: Corpus frequencies correlate with RTs, when the appropriate level of corpus analysis is used.
- What is the time course of information integration?
 - Modular (syntax-first)? Or non-modular?
- **Is the parser serial or constrained parallel?**