Do prosody and discourse status affect the allocation of working memory resources?  
Experimental evidence from appositive clauses

Margaret Kroll & Matt Wagers  
*University of California, Santa Cruz*  
margaretkroll@ucsc.edu
A well-established effect in the psycholinguistics literature is length effects: Increasing the length of a dependency within a sentence, such as by adding material between a subject and verb dependency, increases the processing difficulty of the sentence.

(Gibson 1998; Hale 2001; Warren & Gibson 2002; McElree et al. 2003; Van Dyke & Lewis 2003; Grodner & Gibson 2005; Lewis & Vasishth 2005; Levy 2008; Bartek et al. 2011; Dillon et al. 2014)
OVERVIEW

The man who was on the cruise tried to throw a waitress overboard.
OVERVIEW

The man who was on the cruise *Mary took to the Pacific Islands* tried to throw a waitress overboard.

Figure 1. Restrictive Relative Length Effects
OVERVIEW

The man who was on the cruise Mary took to the Pacific Islands tried to throw a waitress overboard.

Figure 1. Restrictive Relative Length Effects
In several recent experiments, Dillon et al. (2014) showed that length effects do not hold equal for all constructions. For example, let's look at a sentence that is similar to the one we just saw, but contains an appositive instead of a restrictive relative:

The man (who was on the cruise) tried to throw a waitress overboard.
In several recent experiments, Dillon et al. (2014) showed that length effects do not hold equal for all constructions. For example, let's look at a sentence that is similar to the one we just saw, but contains an appositive instead of a restrictive relative:

The man (who was on the cruise Mary took to the Pacific Islands) tried to throw a waitress overboard.
OVERVIEW

Crucial finding: Interaction of Length and Clause ($p < .001$)

Figure 2. Appositives vs. restrictive relatives length penalties

**Parentheticals have a weaker length penalty**

In parentheticals or restrictive relatives

Experiment 1 Summary. Error bars indicate standard error of the mean.
OVERVIEW

Why should appositives have a lesser length penalty than restrictive relatives?

Hypothesis proposed by Dillon et al.: Appositives express secondary discourse content and, as such, draw from separate working memory resources as their surrounding context. Restrictive relative clauses express primary content and, as such, draw from the same memory resources as matrix clause content.

Goal of today's talk: Test the discourse status hypothesis by explicitly manipulating the relationship between the appositive and its discourse context; then consider an alternative hypothesis involving implicit prosody.
PRIMARY VS. SECONDARY DISCOURSE CONTENT
Primary discourse content is content that expresses main or central information in an utterance, while secondary content expresses supplementary or ancillary information.

Unlike primary discourse content, secondary discourse content:

- is speaker-oriented (Potts 2005);
- can project beyond operators such as negation and propositional attitude verbs (Beaver 2001; Potts 2005/2012; Harris & Potts 2009; a.o.);
- does not contribute to resolving the current Question Under Discussion (QUD) of the discourse (Roberts 1996/2012; Ginzburg 1996, Simons et al. 2010, Tonhauser 2012).

Appositives pass our diagnostics for secondary discourse content:

They are speaker-oriented, can project, and are generally used to express information secondary to the current QUD:

Q: What bill does Trump think the Republicans passed?
A: Trump thinks that the Republicans, who apparently never realized that policy is complicated, passed a health care bill.
But, appositives *sometimes* pattern like primary discourse content as diagnosed by their ability to:

**Not Project/Not be Speaker-Oriented** (Amaral et al. 2007; Potts 2007; Schlenker 2003/2007; Sauerland 2007; Harris & Potts 2009):

Joan is crazy. She’s hallucinating that some geniuses in Silicon Valley have invented a new brain chip that’s been installed in her left temporal lobe and permits her to speak any of a number of languages she’s never studied. Joan believes that her chip, *which was installed last month*, has a twelve year guarantee.

**Contribute Main Discourse Information** (contribute to resolving a QUD) (Tonhauser 2012; Koev 2013):

Q: Who did you see at the party and *what did they bring*?
A: I saw Beatrice, *who brought a spongecake*. 
DISCOURSE STATUS AS A DRIVER OF LENGTH EFFECTS
Recall Dillon et al.'s hypothesis: Length effects in sentences containing appositive clauses are attenuated because appositives contribute secondary discourse content, and therefore draw from separate working memory resources as the matrix clause.

Current Proposal: Control for whether the appositive clause contributes primary or secondary discourse content.

We can control the discourse status of appositives in an experimental setting by varying whether the appositive content addresses an explicit experimental QUD.
Some Terminology

- I'll refer to appositives that contribute to resolving a QUD as **responsive appositives**.

  Q: Who did you see at the potluck and **what dish did they bring**?
  A: I saw Hilary, **who brought an artichoke dip**.

- I'll refer to appositives that do not contribute to resolving a QUD as **supplemental appositives**.

  Q: Who did you see at the potluck?
  A: I saw Hilary, **who brought her puppy**.
Experiment 2 tests whether the length effects observed by Dillon et al. and in Experiment 1 can be shown to be due to the discourse status of the appositive clause.

Materials are similar to Experiment 1, except the target sentences are embedded in a multi-exchange discourse organized as a simulated text message exchange between two interlocutors.

The target sentences were presented as answers to an explicit question in the discourse: responsive appositives were presented as answers to a coordinated question, and supplemental appositives were presented in answers to a single question.
EXPERIMENT 2

- Supplemental Appositive Condition
  
  Q: What is the bear wearing?  
  
  A: The bear (who is standing on the ball) is wearing a hat.

- Responsive Appositive Condition
  
  Q: Where is the bear standing and what is it wearing?  
  
  A: The bear (who is standing on the ball) is wearing a hat.
EXPERIMENT 2

DESIGN & MATERIALS

- 2 x 2 x 2 design, with the factors Length (long or short), Clause (parenthetical or restrictive), and Discourse Status (responsive or supplemental).
- Experiment 2 uses medial ARC materials.

Table 1. Experiment 2 Length x Clause Conditions

<table>
<thead>
<tr>
<th>Clause</th>
<th>Length</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parenthetical</td>
<td>Short</td>
<td>The bear (who is standing on the ball) is wearing a hat.</td>
</tr>
<tr>
<td></td>
<td>Long</td>
<td>The bear (who is standing on the ball the trainer rolled across the room) is wearing a hat.</td>
</tr>
<tr>
<td>Restrictive</td>
<td>Short</td>
<td>The bear that is standing on the ball is wearing a hat.</td>
</tr>
<tr>
<td></td>
<td>Long</td>
<td>The bear that is standing on the ball the trainer rolled across the room is wearing a hat.</td>
</tr>
</tbody>
</table>
Experiment 2: Responsive Parenthetical Screen

Questioner:

Hi Viewer, what animals do you see in the video?

Viewer:

I see a vulture, a tortoise, and a sparrow.

What is the sparrow doing and what is it wearing?

The sparrow (who is eating seeds the gardener scattered on the ground) is wearing a tiny hat.
EXPERIMENT 2

PROCEDURE

- Experiment was administered on Mechanical Turk using Ibex.

- Participants were asked to rate test sentences for their naturalness on a Likert scale of 1 (very unnatural) to 7 (very natural).

SUBJECTS

- 40 workers on Amazon's Mechanical Turk.

- IP addresses were restricted to the United States and all workers self-identified as native English speakers.
EXPERIMENT 2

RESULTS

Main effects of Length, Clause, and Discourse Status, and an interaction of Length and Clause (all ps < .01). No interaction of Length, Clause, and Discourse Status.

Figure 3. Experiment 2 Mean Ratings

Experiment 2 Summary. Error bars indicate standard error of the mean computed over items. n = 40
Crucial finding: Length effects show no significant difference between responsive appositive clauses and supplementary appositive clauses. This is contrary to what we would expect if the observed length effect differences between appositives and restrictive relatives is due to the discourse status of the clauses.

What is driving the observed length effects?
IMPLICIT PROSODY AS A DRIVER OF LENGTH EFFECTS
What is driving the observed length effects?

Implicit Prosody

- When we read silently, we impose a prosodic contour onto the material we're reading.

- There is evidence that implicit prosody can affect language processing (Fodor 1998/2002, Bader 1998, Hirotani et al. 2006; a.m.o.).

- For example, implicit prosody is argued to affect attachment ambiguities:
  1. The divorced bishop's daughter
  2. The recently divorced bishop's daughter
  3. The recently divorced bishop's daughter-in-law
What is driving the observed length effects?

Implicit Prosody

- When we read silently, we impose a prosodic contour onto the material we're reading.
- There is evidence that implicit prosody can affect language processing (Fodor 1998/2002, Bader 1998, Hirotani et al. 2006; a.m.o.).
- For example, implicit prosody is argued to affect attachment ambiguities:
  1. [The divorced bishop]'s daughter
  2. The recently divorced [bishop's daughter]
  3. [The recently divorced bishop]'s [daughter-in-law]
What is driving the observed length effects?

- Fodor calls this effect the **Same Size Sister Constraint** (part of the broader **Implicit Prosody Hypothesis** (Fodor 1998/2002; Quinn et al. 2000):

  Sentence comprehension involves a syntactic and a prosodic parser: the prosodic parser packages material into phrases roughly balanced in length, and the syntactic parser interprets only those sentence parts that are in a prosodic phrase.

- Additional work suggests that, when reading, readers dwell on material before punctuation to do any necessary integrative processing before moving on to the next word or phrase (Just & Carpenter 1980; Warren et al. 2009; Kuperman et al. 2010; a.o.).

- These observations have led to the hypothesis that intonational phrases correspond to units held in working memory (Hirotani 2004; Hirotani et al. 2006).
PROSODIC MANIPULATION 1:
LINE BREAKS
EXPERIMENT 3

- Experiment 3 uses a method inspired by Swets et al. (2007), in which sentences are presented in noncumulative chunks in order to impose a particular implicit prosody onto the items.

- Swets et al. presented ambiguous relative clauses in Dutch and English to participants in three timed chunks:

  [The servant of the princess] [who scratched herself in public] [is mildly ashamed]

- The authors found that presenting the sentence in chunks instead of all together yielded more N1, or high, attachments in both English and Dutch.
EXPERIMENT 3

DESIGN & MATERIALS

- 2 x 2 design, with the factors *Length* (long or short) and *Clause* (parenthetical or restrictive).

- Experiment 3 uses the same medial ARC materials as Experiment 2.

Table 2. Experiment 3 Length x Clause Conditions

<table>
<thead>
<tr>
<th>Clause</th>
<th>Length</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parenthetical</td>
<td>Short</td>
<td>The bear (who is standing on the ball) is wearing a hat.</td>
</tr>
<tr>
<td></td>
<td>Long</td>
<td>The bear (who is standing on the ball the trainer rolled across the room) is wearing a hat.</td>
</tr>
<tr>
<td>Restrictive</td>
<td>Short</td>
<td>The bear that is standing on the ball is wearing a hat.</td>
</tr>
<tr>
<td></td>
<td>Long</td>
<td>The bear that is standing on the ball the trainer rolled across the room is wearing a hat.</td>
</tr>
</tbody>
</table>
EXPERIMENT 3

DESIGN & MATERIALS

- All test items were presented in three windows.

- Each sentence chunk was presented for a fixed period of time in the center of the screen. Timing was determined based on word and character quantities.*

- Fillers were presented in the same manner as the test items. Following Swets et al. (2007), breaks were made only at major syntactic boundaries. The amount of material in each break varied, so that some windows included a short string of text (one to two words) and others included a long string of text (up to 12 words).

*Thanks to Ethan Poole and Brian Dillon for this code.
For example, here's what a short parenthetical item would look like:
The bear
(who is standing on the ball)
is wearing a hat.
How natural was the preceding sentence?

(Very unnatural) 1 2 3 4 5 6 7 (Very natural)

Use number keys to answer
EXPERIMENT 3

PROCEDURE

- Experiment was administered on TurkPrime using Ibex.
- Participants were asked to rate test sentences for their naturalness on a Likert scale of 1 (very unnatural) to 7 (very natural).

SUBJECTS

- 40 workers on TurkPrime.
- IP addresses were restricted to the United States and all workers self-identified as native English speakers.
EXPERIMENT 3
RESULTS

Main effects of Length and Clause ($ps < .001$) and an interaction of Length and Clause ($p < .05$).

Figure 4. Experiment 3 Mean Ratings

Line breaks did not change the length penalty
In medial parentheticals or restrictive relatives

Experiment 3 Summary. Error bars indicate standard error of the mean.
EXPERIMENT 3

The length penalty is attenuated compared to the conditions in Experiment 2.

Length penalties: (mean rating of short condition) - (mean rating of long condition)

Figure 5. Experiments 2 & 3 Length Penalty Comparisons
What can we conclude?

- Manipulating the presentation of the items attenuated the length penalties for appositives and restrictive relatives; however, the interaction is still statistically significant.

- Three possibilities:
  - This particular line break manipulation does not control implicit prosody;
  - prosody is not the driving factor of the line breaks;
  - not controlling for expertness of readers/individual differences (Jun & Bishop 2015).

- We tried one more prosodic manipulation...
PROSODIC MANIPULATION 2: ADVERBS
EXPERIMENT 4

Experiment 4 uses adverbs to force a particular prosodic structure onto items.*

For example:

The bear that is standing on the ball the trainer rolled across the room is, predictably, wearing a hat.

* Thank you to Junko Ito for suggesting this idea.
### EXPERIMENT 4

#### DESIGN & MATERIALS

- **2 x 2 x 2** design, with the factors *Length* (long or short) and *Clause* (parenthetical or restrictive) and *Adverb* (adverb or no adverb).
- Experiment 4 uses the same medial ARC materials as Experiments 2&3.

#### Table 3. Experiment 4 Length x Clause x Adverb Conditions

<table>
<thead>
<tr>
<th>Clause</th>
<th>Length</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parenthetical</td>
<td>Short</td>
<td>The bear (who is standing on the ball) is, predictably, wearing a hat.</td>
</tr>
<tr>
<td></td>
<td>Long</td>
<td>The bear (who is standing on the ball the trainer rolled across the room) is, predictably, wearing a hat.</td>
</tr>
<tr>
<td>Restrictive</td>
<td>Short</td>
<td>The bear that is standing on the ball is, predictably, wearing a hat.</td>
</tr>
<tr>
<td></td>
<td>Long</td>
<td>The bear that is standing on the ball the trainer rolled across the room is, predictably, wearing a hat.</td>
</tr>
</tbody>
</table>
EXPERIMENT 4

PROCEDURE

- Experiment was administered on TurkPrime using Ibex.
- Participants were asked to rate test sentences for their naturalness on a Likert scale of 1 (very unnatural) to 7 (very natural).
- The dialogue presentation from Experiments 2&3 was used.

SUBJECTS

- 40 workers on TurkPrime.
- IP addresses were restricted to the United States and all workers self-identified as native English speakers.
EXPERIMENT 4

RESULTS

Main effects of Length, Clause, and Adverb; interactions of Length and Clause and Length and Adverb (ps < .001).

Figure 6. Experiment 4 Mean Ratings

Error bars indicate standard error of the mean computed over items.
EXPERIMENT 4
RESULTS

The length effect decrease is driven by a lowering of the ratings of the long and short parentheticals and the short restrictives in the Adverb condition.

Figure 6. Experiment 4 Mean Ratings

Medial adverb effects
In parentheticals and restrictive relatives

Error bars indicate standard error of the mean computed over items.
EXPERIMENT 4

RESULTS

Why did we not see an adverb cost on the long restrictives?

- One possibility is that the long restrictives are at a ratings floor.
- However, filler ratings suggest that is not the case.

Table 4. Experiment 4 Filler Means

<table>
<thead>
<tr>
<th>Filler Type</th>
<th>Mean Rating</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical</td>
<td>5.34 (s = 1.46)</td>
<td>1400</td>
</tr>
<tr>
<td>In Between</td>
<td>3.2 (s = 1.84)</td>
<td>520</td>
</tr>
<tr>
<td>Ungrammatical</td>
<td>2.42 (s = 1.66)</td>
<td>640</td>
</tr>
</tbody>
</table>
EXPERIMENT 4

The length penalty in the Adverb Condition is attenuated compared to the conditions in Experiment 2.

Length penalties: (mean rating of short condition) - (mean rating of long condition)

Figure 7. Experiment 2, 3, and 4 Length Penalty Comparisons
What can we conclude?

- Manipulating the prosody of the items by adding in adverbs *did* appear to have some effect on ratings.

- However, this effect decreased the length penalty by lowering the rating of the long and short parentheticals and the short restrictives, as opposed to increasing the rating of the long restrictives.

- One possibility is that this effect is due to the implicit prosody imposed by the adverbs:
  - In the short conditions and long parenthetical condition, the adverbs chopped the sentences into smaller than ideal prosodic phrases;
  - In the long restrictive condition, the adverb created a parsing benefit that cancelled out the cost of creating a small prosodic phrase.
CONCLUSION AND FUTURE DIRECTIONS
What I've shown in this talk

- Additional evidence for a consistent length effect in parentheticals and restrictive relative clauses in which lengthening a restrictive relative clause decreases the acceptability of the containing sentence more than lengthening a comparable parenthetical (following Dillon et al. 2014).

- Evidence that changing the discourse status of the parentheticals does not change the length effect, counter to the predictions of Dillon et al.

- Some suggestive, but inconclusive, evidence that changing the prosody of the sentences containing restrictive relatives attenuates the length effect.
Thank you!  

Acknowledgments  

Thank you to Pranav Anand, Adrian Brasoveanu, Brian Dillon, Donka Farkas, Amanda Rysling, and audiences at UCSC and the CUNY Graduate Center for helpful comments and feedback.

Thank you also to Jed Pizarro-Guevara for help with my experimental setup, and to Caitlin Kronk for the illustrations.