Memory encoding/retrieval is sensitive to discourse status: through the lens of pronoun resolution

Sanghee J. Kim & Ming Xiang (The University of Chicago)

The 36th Annual Conference on Human Sentence Processing
The University of Pittsburgh
March 9-11, 2023

Discourse structure in working memory (WM)

- Much work on the interaction of WM and sentence comprehension in (morpho-) syntactic dependencies (Just & Carpenter (1992); MacDonald (1992); Caplan & Waters (1999); Gibson (2000); Lewis & Vasishth (2005))

How is information about discourse structure encoded and retrieved in WM?

Discourse structure in working memory (WM)

- Much work on the interaction of WM and sentence comprehension in (morpho-) syntactic dependencies (Just & Carpenter (1992); MacDonald (1992); Caplan & Waters (1999); Gibson (2000); Lewis & Vasishth (2005))

How is information about discourse structure encoded and retrieved in WM?

- “Discourse structure” focused on the division of [+ main/subordinate]
- Restrictive relative clauses (RRCs) vs. Appositive relative clauses (ARCs)
- Pronoun resolution

Discourse status of ARCs and RRCs

ARC
Appositive
Relative clause

The violinist, who the singers admired, came to the party.

Side comment
Subordinate discourse

Main discourse

RRC
Restrictive
Relative clause

The violinist who the singers admired came to the party.

RRC restricting the head NP
Part of **main discourse**

Discourse status of ARCs and RRCs

ARC	RRC
• Subordinate discourse structure • Side comments; not part of the main point of the utterance (Dehé & Kavalova (2007); Asher & Lascarides (2003); Koev (2013))	• Part of main discourse structure

Previous work

Information with distinct discourse status leads to different interference effects

Empirical evidence 1: Number agreement attraction effect

Ng & Husband (2017); McInnerney & Atkinson (2020); Kim & Xiang (2022)

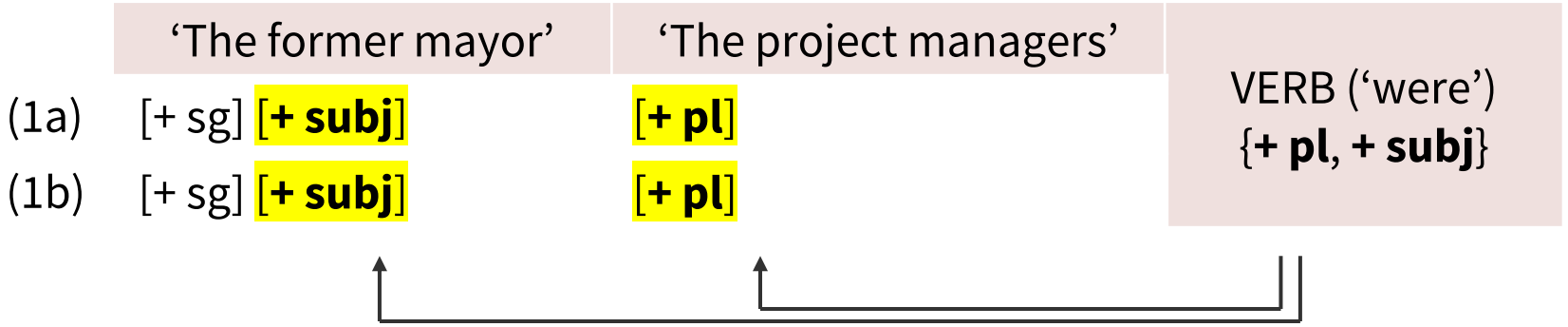
(1a) ***The former mayor** who hired **the project managers** [..] **were** upset [..].
(1b) ***The former mayor**, the one who hired **the project managers**, [..] **were** upset [..].

	‘The former mayor’	‘The project managers’	VERB (‘were’)
(1a)			
(1b)			

Empirical evidence 1: Number agreement attraction effect

Ng & Husband (2017); McInnerney & Atkinson (2020); Kim & Xiang (2022)

(1a) ***The former mayor** who hired **the project managers** [..] **were** upset [..].
(1b) ***The former mayor**, the one who hired **the project managers**, [..] **were** upset [..].



Prediction: Number agreement attraction effect in both (1a) and (1b)

Empirical evidence 1: Number agreement attraction effect

Ng & Husband (2017); McInnerney & Atkinson (2020); Kim & Xiang (2022)

(1a) ***The former mayor** who hired **the project managers** [..] **were** upset [..].
(1b) ***The former mayor**, the one who hired **the project managers**, [..] **were** upset [..].

	‘The former mayor’	‘The project managers’	VERB (‘were’) {**+ pl**, **+ subj**}
(1a)	[+ sg] [**+ subj**]	[**+ pl**]	
(1b)	[+ sg] [**+ subj**]	[**+ pl**]	

Absence of number agreement attraction effect in appositives (1b)

Why?

Empirical evidence 1: Number agreement attraction effect

Ng & Husband (2017); McInnerney & Atkinson (2020); Kim & Xiang (2022)

(1a) ***The former mayor** who hired **the project managers** [..] **were** upset [..].
(1b) ***The former mayor**, the one who hired **the project managers**, [..] **were** upset [..].

	'The former mayor'	'The project managers'	VERB ('were')
(1a)	[+ sg] [**+ subj**] [**+ main**]	[**+ pl**] [**+ main**]	{**+ pl**, **+ subj**, **+ main**}
(1b)	[+ sg] [**+ subj**] [**+ main**]	[**+ pl**] [+ subord.]	

Empirical evidence 1: Number agreement attraction effect

Ng & Husband (2017); McInnerney & Atkinson (2020); Kim & Xiang (2022)

(1a) ***The former mayor** who hired **the project managers** [..] **were** upset [..].
(1b) ***The former mayor**, the one who hired **the project managers**, [..] **were** upset [..].

	'The former mayor'	'The project managers'	VERB ('were')
(1a)	[+ sg] **[+ subj]** **[+ main]**	**[+ pl]** **[+ main]**	{**+ pl**, **+ subj**, **+ main**}
(1b)	[+ sg] **[+ subj]** **[+ main]**	**[+ pl]** [+ subord.]	

Different degree of partial feature match between the distractor and the retrieval cues

Empirical evidence 2: Syntactic complexity effect

- Appositive structures (ARCs) show reduced processing cost in syntactically complex structures, e.g., filler-gap dependency, wh-movement (Dillon et al. (2014); Dillon et al. (2017); Kroll & Wagers (2019))

Empirical evidence 2: Syntactic complexity effect

Dillon et al. (2017): Filler-gap dependency

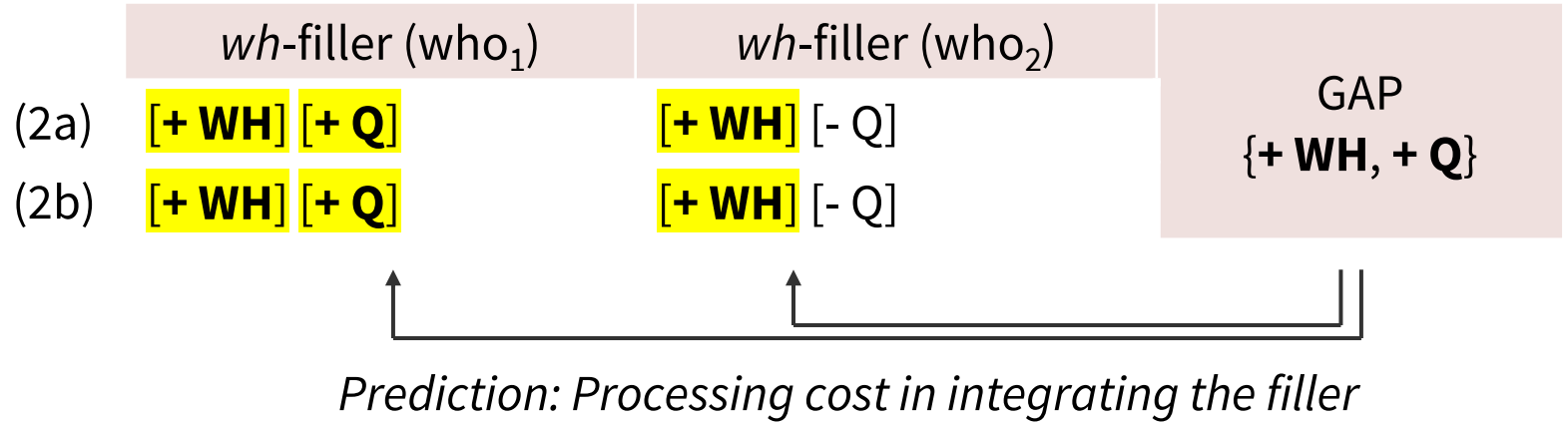
(2a) The butcher asked **who$_1$** the lady **who$_2$** bought Italian ham was cooking **dinner for __$_1$**.
(2b) The butcher asked **who$_1$** the lady, **who$_2$** bought Italian ham, was cooking **dinner for __$_1$**.

	wh-filler (who$_1$)	*wh*-filler (who$_2$)	GAP {**+ WH**, **+ Q**}
(2a)			
(2b)			

Empirical evidence 2: Syntactic complexity effect

Dillon et al. (2017): Filler-gap dependency

(2a) The butcher asked **who$_1$** the lady **who$_2$** bought Italian ham was cooking **dinner for __$_1$**.
(2b) The butcher asked **who$_1$** the lady, **who$_2$** bought Italian ham, was cooking **dinner for __$_1$**.



Prediction: Processing cost in integrating the filler

Empirical evidence 2: Syntactic complexity effect

Dillon et al. (2017): Filler-gap dependency

(2a) The butcher asked **who$_1$** the lady **who$_2$** bought Italian ham was cooking **dinner for __$_1$**.
(2b) The butcher asked **who$_1$** the lady, **who$_2$** bought Italian ham, was cooking **dinner for __$_1$**.

	wh-filler (who$_1$)	*wh*-filler (who$_2$)	GAP
(2a)	**[+ WH]** **[+ Q]**	**[+ WH]** [- Q]	{**+ WH**, **+ Q**}
(2b)	**[+ WH]** **[+ Q]**	**[+ WH]** [- Q]	

Reduced processing cost at the gap site with ARC (2b)

Why?

Empirical evidence 2: Syntactic complexity effect

Dillon et al. (2017): Filler-gap dependency

(2a) The butcher asked **who$_1$** the lady **who$_2$** bought Italian ham was cooking **dinner for __$_1$**.
(2b) The butcher asked **who$_1$** the lady, **who$_2$** bought Italian ham, was cooking **dinner for __$_1$**.

	wh-filler (who$_1$)	*wh*-filler (who$_2$)	GAP
(2a)	**[+ WH]** **[+ Q]**	**[+ WH]** [- Q]	{**+ WH**, **+ Q**}
(2b)	**[+ WH]** **[+ Q]**	**[+ WH]** [- Q]	

Empirical evidence 2: Syntactic complexity effect

Dillon et al. (2017): Filler-gap dependency

(2a) The butcher asked **who$_1$** the lady **who$_2$** bought Italian ham was cooking **dinner for __$_1$**.
(2b) The butcher asked **who$_1$** the lady, **who$_2$** bought Italian ham, was cooking **dinner for __$_1$**.

	wh-filler (who$_1$)	*wh*-filler (who$_2$)	GAP
(2a)	**[+ WH]** **[+ Q]** [+ main]	**[+ WH]** [- Q] [+ main]	{**+ WH**, **+ Q**, **+ main**}
(2b)	**[+ WH]** **[+ Q]** [+ main]	**[+ WH]** [- Q] [+ subord.]	

Different degree of partial feature match
between the intervening *wh*-word and the retrieval cues

Key takeaway of empirical findings

- At least for these two empirical cases, information contained in the subordinate discourse unit (ARC) does not introduce an interference effect whereas information contained in the main discourse unit (RRC) does

Key takeaway of empirical findings

- Suggests that discourse structure information is encoded as a feature and is used as a retrieval cue: [+ main], [+ subordinate]

‘The former mayor’	‘The project managers’	VERB (‘were’)
[+ sg] [**+ subj**] [**+ main**]	[**+ pl**] [- subj] [**+ main**]	{**+ pl**, **+ subj**, **+ main**}
[+ sg] [**+ subj**] [**+ main**]	[**+ pl**] [- subj] [+ subord.]	

Present study

- Previous work: Representation of discourse structure information in WM leads to a retrieval interference effect

- Current work: Encoding interference effect triggered by discourse structure information features

Present study

Key properties of our test case

1) Discourse structure information is not part of the retrieval cue

2) Discourse structure information features are encoded on retrieval targets

Test case

(3) The violinists, who admired the singers, invited **their** mentors to the party.

- 'The violinists' (main) & 'The singers' (subordinate) – ambiguous antecedent
- While the two NPs are part of different discourse structures, both NPs are available for pronoun resolution (AnderBois et al. 2015)
- Key contrast from previous cases (e.g., number agreement attraction):
  - Both NPs are grammatically possible antecedents for retrieval despite the distinction in discourse structure information

Test case

(4a) The violinists who admired the singers invited **their** mentors [..]

(4b) The violinists, who admired the singers, invited **their** mentors [..]

	‘The violinists’	‘The singers’	Pronoun (‘their’) {+ pl.}
(4a)	[+ pl, + main]	[+ pl, + main]	
(4b)	[+ pl, + main]	[+ pl, + subordinate]	

1) No retrieval cue directly associated with discourse structure information

Test case

(4a) The violinists who admired the singers invited **their** mentors [..]

(4b) The violinists, who admired the singers, invited **their** mentors [..]

	‘The violinists’	‘The singers’	Pronoun (‘their’) {+ pl.}
(4a)	[+ pl, + main]	[+ pl, + main]	
(4b)	[+ pl, + main]	[+ pl, + subordinate]	

2) Discourse structure information feature is encoded

Test case

(4a) The violinists who admired the singers invited **their** mentors [..]

(4b) The violinists, who admired the singers, invited **their** mentors [..]

	'The violinists'	'The singers'	Pronoun ('their') {+ pl.}
(4a)	[+ pl, + main]	[+ pl, + main]	
(4b)	[+ pl, + main]	[+ pl, + subordinate]	

Would feature overlap of discourse structure information during encoding lead to encoding interference effect?

Encoding interference effect

- Feature overlap during encoding leads to degraded representation of items in WM (Lewis & Vasishth (2005); Oberauer & Kliegl (2006); Vasishth et al. (2017); Rich & Wagers (2020); cf. Parker & Konrad (2020))

- Villata et al. (2018): Italian relative clause (also see Rich & Wagers (2020))

  - (a) The dancer-**FEM** that the waiter-**FEM** has surprised [..]
  - (b) The dancer-**FEM** that the waiter-**MASC** has surprised [..]
  - Gender feature was not a retrieval cue on the verb
  - Gender feature overlap led to processing cost on the verb: (a) > (b)

Test case

	‘The violinists’	‘The singers’	Pronoun (‘their’) {+ pl.}
(4a)	[+ pl, + main]	[+ pl, + main]	
(4b)	[+ pl, + main]	[+ pl, + subordinate]	

More feature overlap in (4a)

Test case

	‘The violinists’	‘The singers’	Pronoun (‘their’) {+ pl.}
(4a)	[+ pl, + main]	[+ pl, + main]	
(4b)	[+ pl, + main]	[+ pl, + subordinate]	

Prediction

Feature overlap of discourse structure information of possible antecedents leads to processing cost in pronoun resolution

Processing cost at pronoun (‘their’): (4a) > (4b)

Experimental design and predictions

Self-paced reading task (critical region ('their'))

Clause	Sentence
RRC	The / violinists / who / the / singers / admired / invited / **their** / **mentors** / to / the / party.
ARC	The / violinists, / who / the / singers / admired, / invited / **their** / **mentors** / to / the / party.

Experimental design and predictions

Self-paced reading task (critical region ('their'))

Clause	Sentence	Discourse status of antecedents
RRC	The / violinists / who / the / singers / admired / invited / **their** / **mentors** / to / the / party.	[+ main], [+ main]
ARC	The / violinists, / who / the / singers / admired, / invited / **their** / **mentors** / to / the / party.	[+ main], [+ subord.]

RT: RRC > ARC @critical region ('their')
Feature overlap – degraded representation

Experimental design and predictions

Self-paced reading task (critical region ('their'))

Clause	Sentence	Discourse status of antecedents
RRC	The / violinists / who / the / singers / admired / invited / **their** / **mentors** / to / the / party.	[+ main], [+ main]
ARC	The / violinists, / who / the / singers / admired, / invited / **their** / **mentors** / to / the / party.	[+ main], [+ subord.]
RRC		[+ main], [+ main]
ARC		[+ main], [+ main]

Experimental design and predictions

Self-paced reading task (critical region ('their'))

RC-Position	Clause	Sentence	Discourse status of antecedents
	RRC	The / violinists / who / the / singers / admired / invited / **their** / **mentors** / to / the / party.	[+ main], [+ main]
	ARC	The / violinists, / who / the / singers / admired, / invited / **their** / **mentors** / to / the / party.	[+ main], [+ subord.]
	RRC		[+ main], [+ main]
	ARC		[+ main], [+ main]

Experimental design and predictions

Self-paced reading task (critical region ('their'))

RC-Position	Clause	Sentence	Discourse status of antecedents
Medial	RRC	The / violinists / who / the / singers / admired / invited / **their** / **mentors** / to / the / party.	[+ main], [+ main]
Medial	ARC	The / violinists, / who / the / singers / admired, / invited / **their** / **mentors** / to / the / party.	[+ main], [+ subord.]
	RRC		[+ main], [+ main]
	ARC		[+ main], [+ main]

Experimental design and predictions

Self-paced reading task (critical region ('their'))

RC-Position	Clause	Sentence	Discourse status of antecedents
Medial	RRC	The / violinists / who / the / singers / admired / invited / **their** / **mentors** / to / the / party.	[+ main], [+ main]
Medial	ARC	The / violinists, / who / the / singers / admired, / invited / **their** / **mentors** / to / the / party.	[+ main], [+ subord.]
Final	RRC		[+ main], [+ main]
Final	ARC		[+ main], [+ main]

Experimental design and predictions

Self-paced reading task (critical region ('their'))

RC-Position	Clause	Sentence	Discourse status of antecedents
Medial	RRC	The / violinists / who / the / singers / admired / invited / **their** / **mentors** / to / the / party.	[+ main], [+ main]
Medial	ARC	The / violinists, / who / the / singers / admired, / invited / **their** / **mentors** / to / the / party.	[+ main], [+ subord.]
Final	RRC	The / singers / admired / the / violinists / who / invited / **their** / **mentors** / to / the / party.	[+ main], [+ main]
Final	ARC	The / singers / admired / the / violinists, / who / invited / **their** / **mentors** / to / the / party.	[+ main], [+ main]

Experimental design and predictions

Self-paced reading task (critical region (‘their’))

RT: RRC ≈ ARC @critical region (‘their’)
The same degree of feature overlap within these two conditions

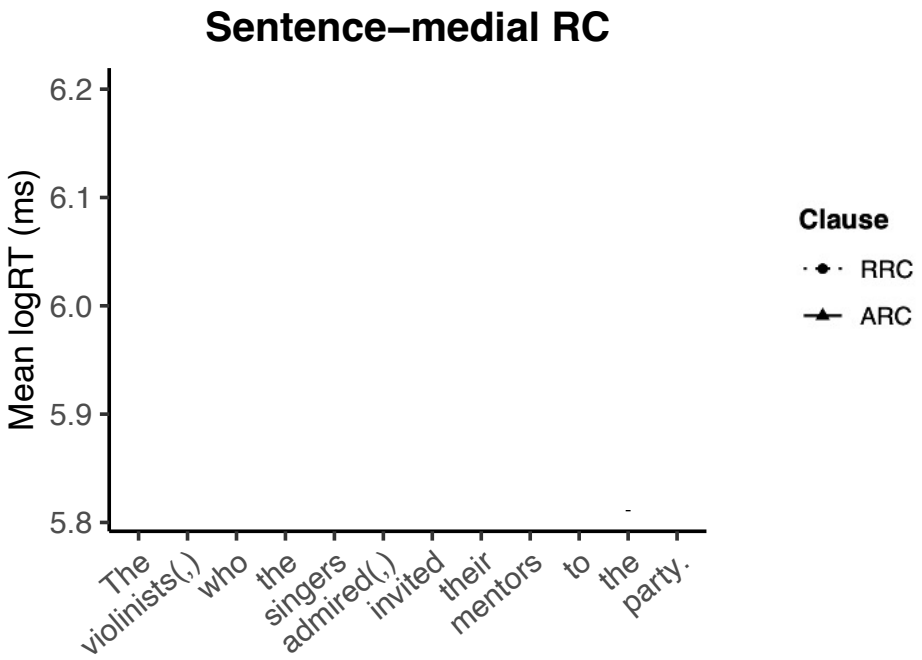
Final	RRC	The / singers / admired / the / violinists / who / invited / **their** / **mentors** / to / the / party.	[+ main], [+ main]
Final	ARC	The / singers / admired / the / violinists, / who / invited / **their** / **mentors** / to / the / party.	[+ main], [+ main]

Experimental design and predictions

- Native English speakers (n = 104); 32 target items & 20 fillers

- Forced antecedent choice task
  (e.g., “Who do you think ‘their mentors’ refer to?”)
  - (a) the singers
    (b) the violinists
    (c) someone else not mentioned in the sentence

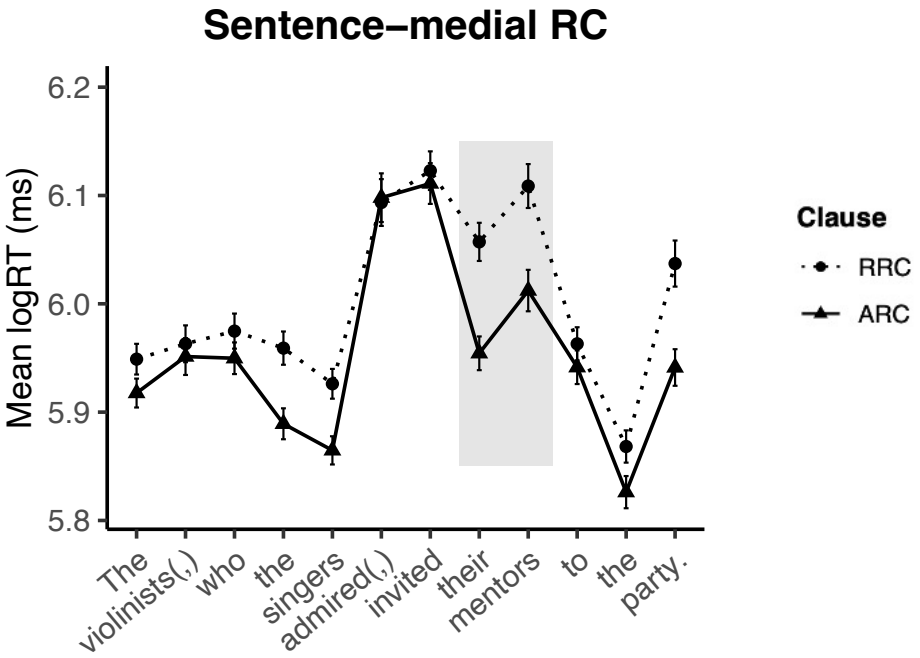
Reading times



The violinists who the singers admired invited **their** mentors to the party.

The violinists, who the singers admired, invited **their** mentors to the party.

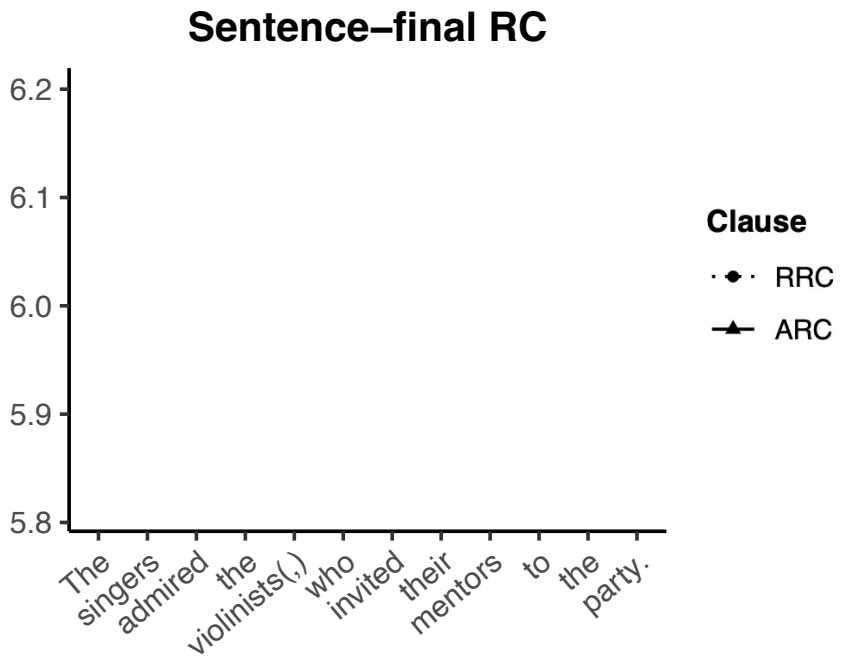
Reading times



The violinists who the singers admired invited **their** mentors to the party.

The violinists, who the singers admired, invited **their** mentors to the party.

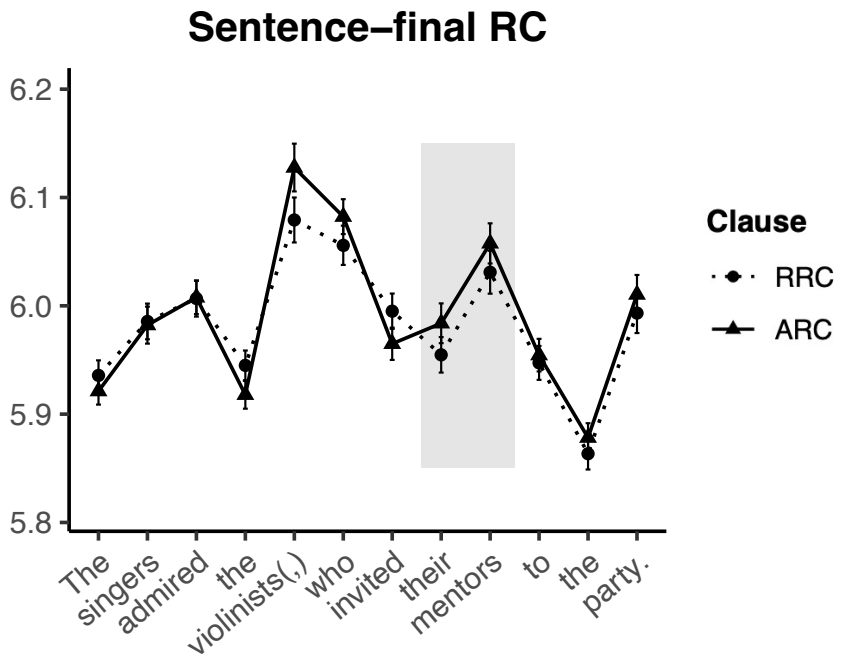
Reading times



The singers admired the violinists who invited **their** mentors to the party.

The singers admired the violinists, who invited **their** mentors to the party.

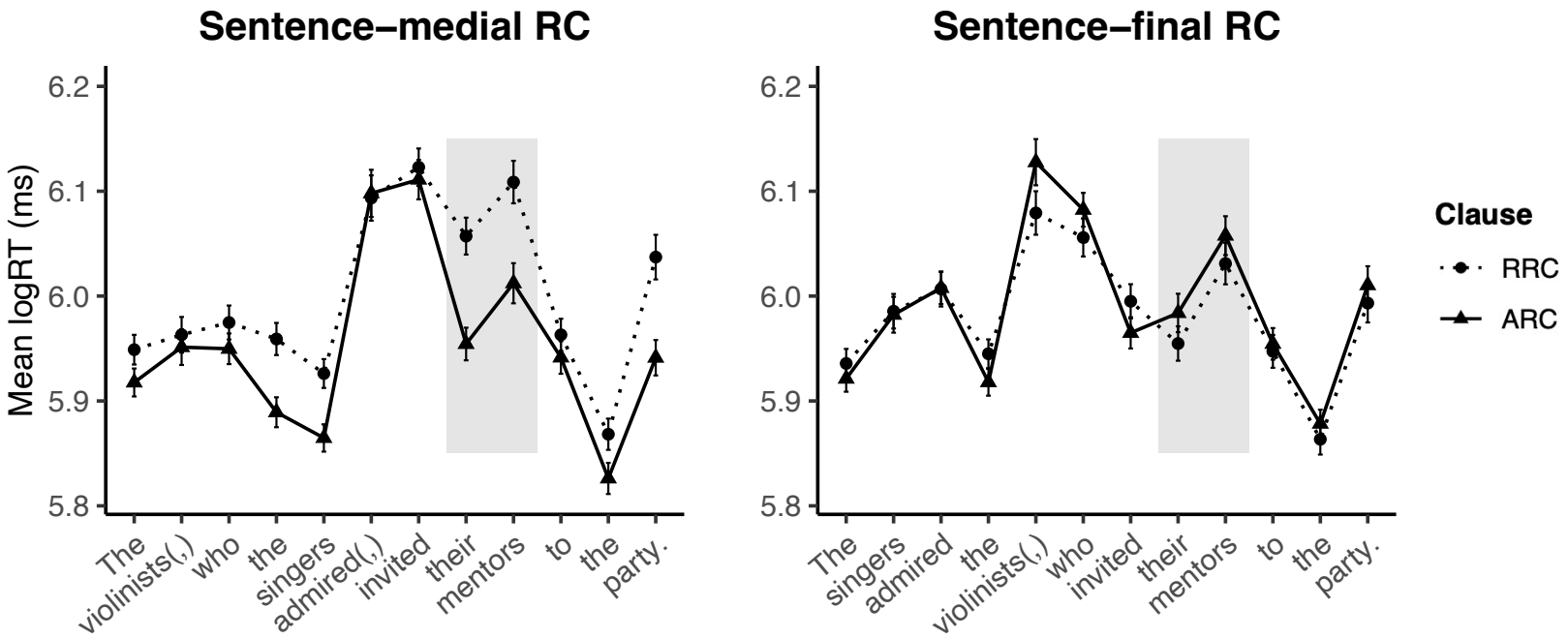
Reading times



The singers admired the violinists who invited **their** mentors to the party.

The singers admired the violinists, who invited **their** mentors to the party.

Reading times



RRC > ARC when RC is sentence-medial position; RRC ≈ ARC otherwise

- @Critical ('their'): 2-way interaction, *b*=0.129, *se=0.025*, *t=5.128*
- @Spillover ('mentors'): 2-way interaction, *b*=0.094, *se=0.029*, *t=3.233*

Reading times: summary

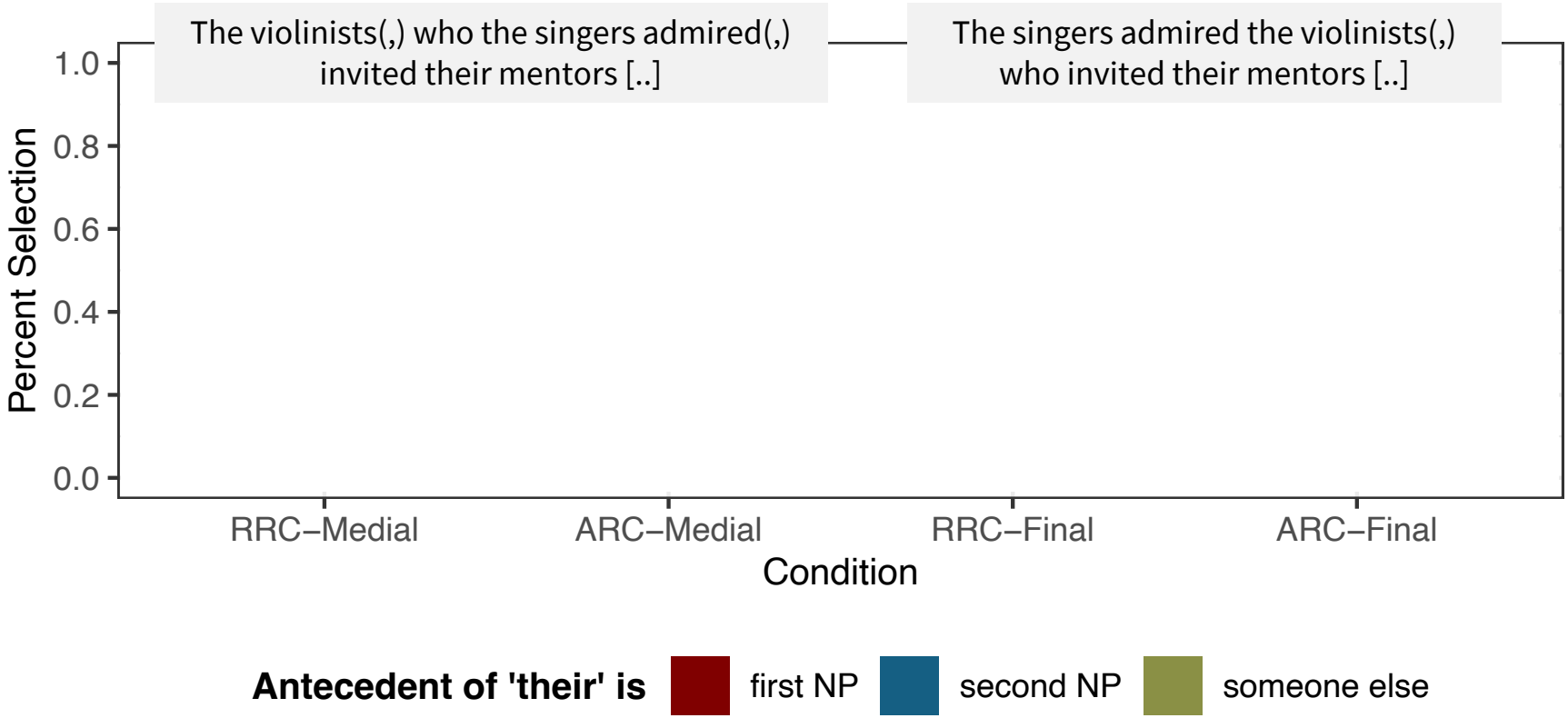
	Condition	Possible antecedent 1	Possible antecedent 2	
(a)	Medial-RRC	[+pl., + main]	[+pl., + main]	
(b)	Medial-ARC	[+pl., + main]	[+pl., + subord.]	
(c)	Final-RRC	[+pl., + main]	[+pl., + main]	*Control conditions*
(d)	Final-ARC	[+pl., + main]	[+pl., + main]	

Processing cost (RTs) as predicted

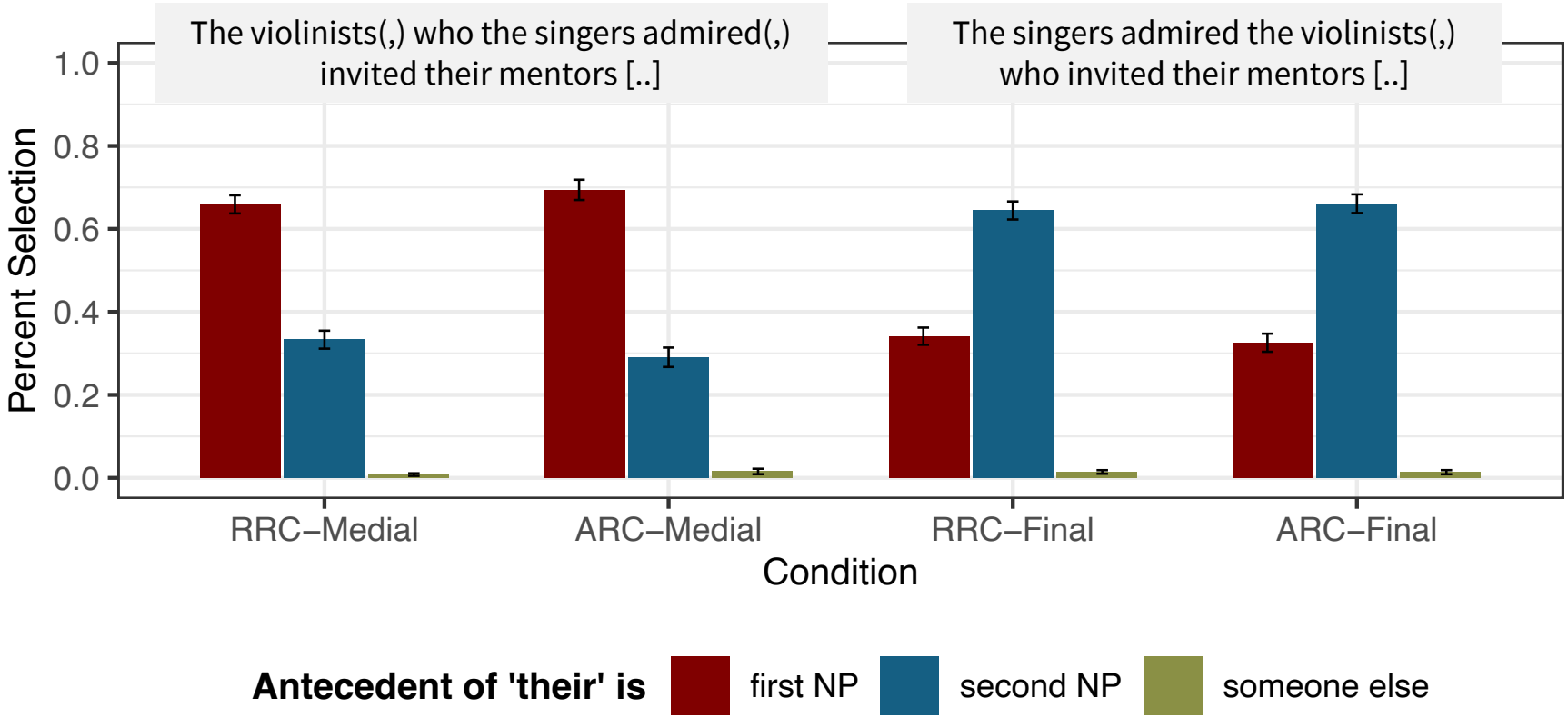
RT: (a) > (b) feature overlap
degraded representation of the antecedents

RT: (c) ≈ (d) same degree of feature overlap

Experiment 1: Antecedent choice task

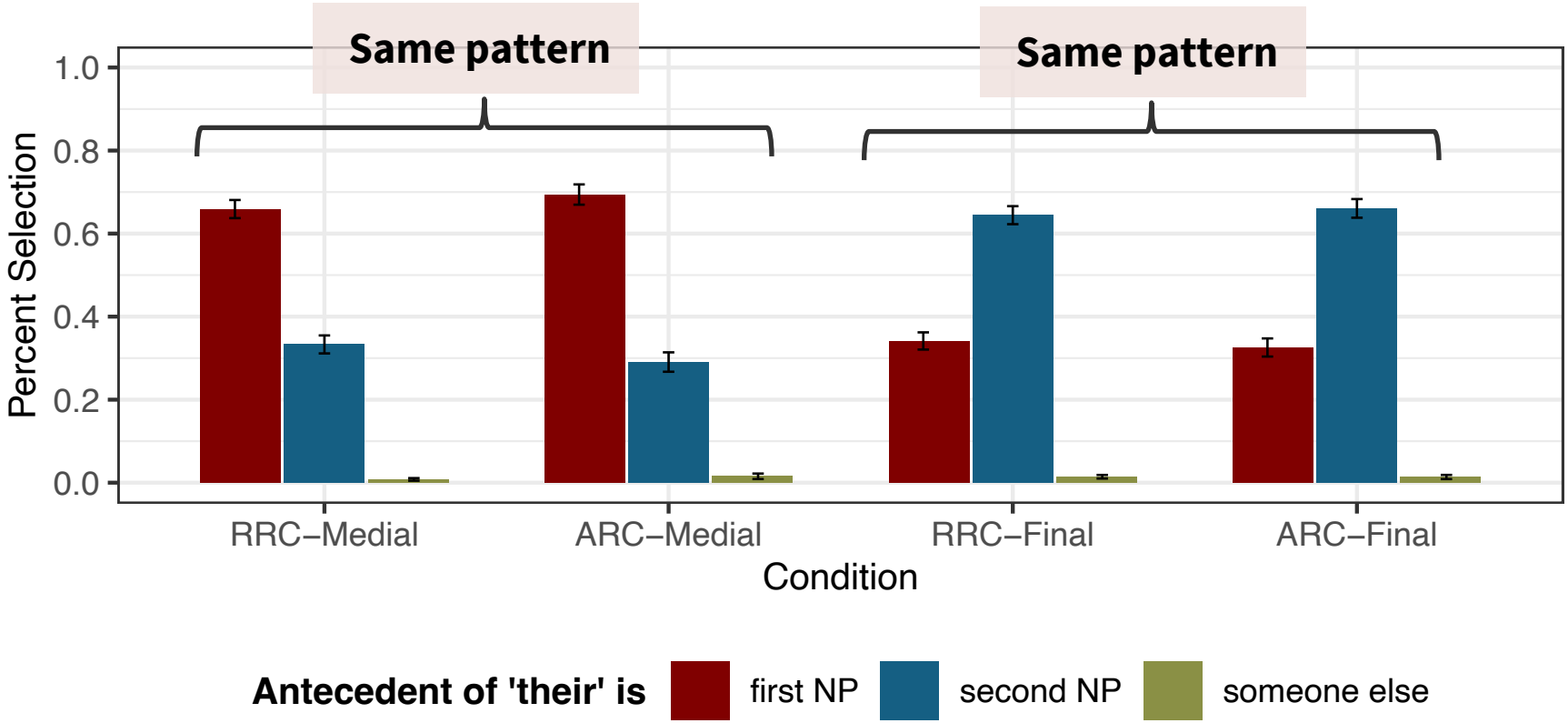


Experiment 1: Antecedent choice task

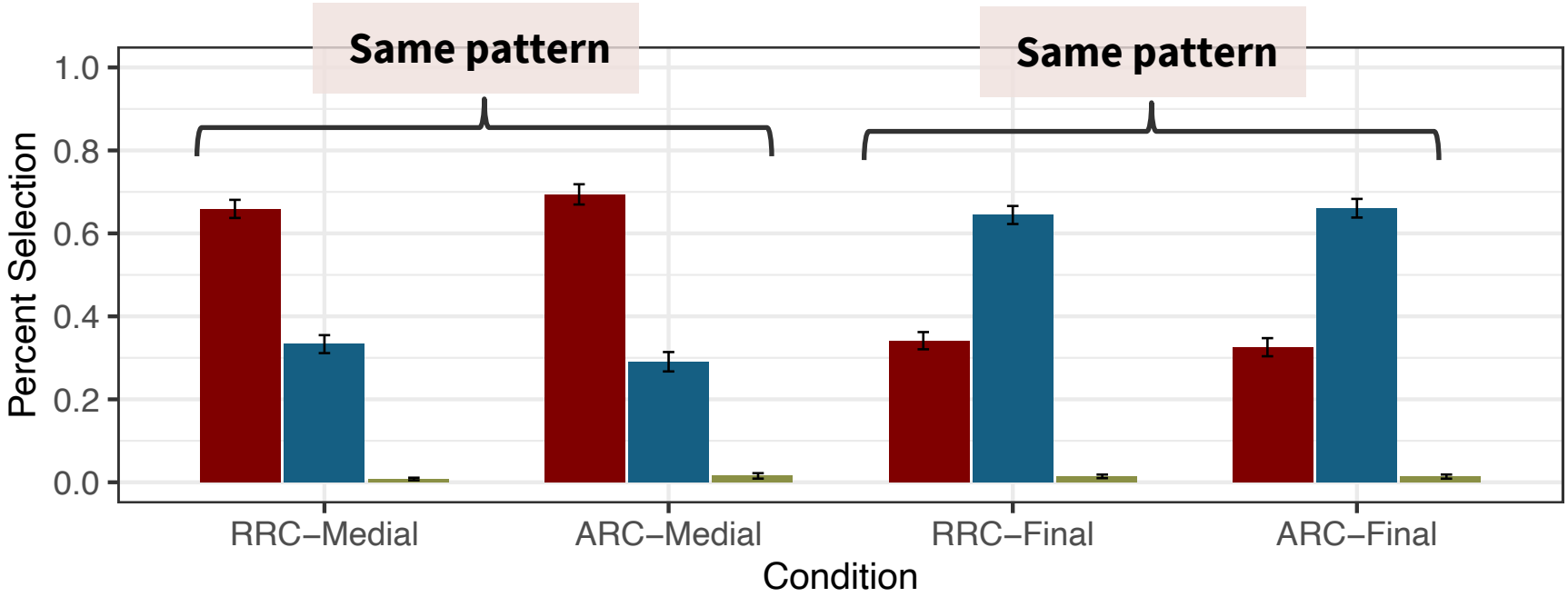


Interaction of Clause and Position (b=0.334, se=0.046, $p < 0.001$)

Experiment 1: Antecedent choice task



Experiment 1: Antecedent choice task



Antecedent choice preference data does not offer an alternative explanation for the RT data

Discussion and conclusion

- Previous work: Discourse structure information serves as *retrieval* cues and leads to a retrieval interference effect
- Current work: *Encoding* interference due to overlap of discourse structure information features
- Ongoing: visual world paradigm (eye-tracking) on the same phenomenon

Discussion and conclusion

- General understanding of what features serve as encoding and retrieval features in WM
    - Number, gender, case, structural constraint, animacy, semantic fit, etc.
    - Discourse structure information also serves as encoding and retrieval features

Acknowledgment

We are grateful for the generous funding support from the NSF DDRI-LING Grant (BCS-2214437).

We appreciate the UChicago Language Processing Lab members, Chris Kennedy, and Jennifer Arnold for discussions at different stages of this project.

We thank the anonymous reviewers for their feedback and comments.

Thank you for listening!

Selected bibliography

AnderBois, S., Brasoveanu, A., & Henderson, R. (2015). At-issue proposals and appositive impositions in discourse. *Journal of Semantics, 32*(1), 93-138.

Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Language and Linguistics Compass, 4*(4), 187-203.

Asher, N., & Lascarides, A. (2003). *Logics of conversation*. Cambridge University Press.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*(1), 77-94.

Creemers, A., & Meyer, A. S. (2022). The processing of ambiguous pronominal reference is sensitive to depth of processing. *Glossa Psycholinguistics, 1*(1).

Dehé, N., & Kavalova, Y. (2007). Parentheticals: An introduction. In N. Dehé & Y. Kavalova (Eds.), *Parentheticals* (pp. 1-22). John Benjamins Publishing Company.

Dillon, B., Clifton, C. J., & Frazier, L. (2014). Pushed aside: Parentheticals, memory and processing. *Language, Cognition and Neuroscience, 29*(4), 483–498.

Dillon, B., Clifton, C. J., Sloggett, S., & Frazier, L. (2017). Appositives and their aftermath: Interference depends on at-issue vs. not-at-issue status. *Journal of Memory and Language, 96*, 93–109.

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of memory and Language*, *51*(1), 97-114.

Selected bibliography

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. P. Marantz, Y. Miyashita, & O. Wayne (Eds.), *Image, language, brain* (pp. 95–126). Cambridge, MA: MIT Press.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122.

Kim, S. J., & Xiang, M. (2022). “Memory retrieval selectively targets different discourse units”. Talk presented at the *35th Annual Meeting of the Society for Human Sentence Processing*, UC Santa Cruz, CA, March 24-26.

Koev, T. (2013). *Apposition and the structure of discourse*. Rutgers University. Doctoral dissertation.

Kroll, M., & Wagers, M. (2019). Working memory resource allocation is not modulated by clausal discourse status. *Ms*. UC Santa Cruz.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375-419.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, *24*(1), 56-98.

Selected bibliography

McInnerney, A., & Atkinson, E. (2020). “Syntactically unintegrated parentheticals: Evidence from agreement attraction”. Talk presented at the *33rd CUNY Conference on Human Sentence Processing*, University of Massachusetts Amherst: Amherst, MA, March 19-March 21.

Ng, A., & Husband, M. (2017). “Interference effects across the at-issue/not-at-issue divide: Agreement and NPI licensing”. Poster presented at the *30th CUNY Conference on Human Sentence Processing*, MIT, Cambridge: MA, March 30-April 1.

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language, 55*(4), 601-626.

Parker, D., & Konrad, K. (2020). Teasing apart encoding and retrieval interference in sentence comprehension: Evidence from agreement attraction. In *Proceedings of the 42nd Annual Conference of the Cognitive Science Society,* 1446-1452.

Rich, S., & Wagers, M. (2020). “Semantic similarity and temporal contiguity in subject-verb dependency processing”. Talk presented at the *33rd CUNY Conference on Human Sentence Processing*, University of Massachusetts Amherst: Amherst, MA, March 19-March 21.

Vasishth, S., Jäger, L. A., & Nicenboim, B. (2017). Feature overwriting as a finite mixture process: Evidence from comprehension data. *arXiv preprint arXiv:1703.04081.*

Villata, S., Tabor, W., & Franck, J. (2018). Encoding and retrieval interference in sentence comprehension: Evidence from agreement. *Frontiers in Psychology, 9*, 2.