Who and when gets the race? Two processing routes for the advantages and penalties of pronominal ambiguity resolution

Ruoqing Yao (ryao10@ucsc.edu)
Department of Linguistics, 1156 High St.
Santa Cruz, CA 95060 USA

Matthew Wagers (mwagers@ucsc.edu)
Department of Linguistics, 1156 High St.
Santa Cruz, CA 95060 USA

Abstract

The current study investigates how pronominal ambiguity is resolved in real-time, focusing on the role of referent bias and task context. In two self-paced reading experiments, we tested whether ambiguity leads to processing benefits or costs modulated by the presence of a biased referent and the task manipulations. Experiment 1 showed that the ambiguity advantage emerges only when a biased referent is not selected, supporting reanalysis-based accounts such as the unrestricted race model (Van Gompel, Pickering, & Traxler, 2000). Experiment 2, however, revealed a delayed ambiguity penalty, suggesting task-induced shifts in processing strategy that better fit a delayed interpretation account. These findings highlight that pronominal ambiguity resolution may involve two processing mechanisms shaped by the parser's evaluation space and the timing of selection.

Keywords: Ambiguity resolution; Pronoun resolution; Ambiguity advantage effect; Sentence processing; Psycholinguistics

Introduction

One key question in psycholinguistics is how people deal with multiple valid analyses simultaneously in memory. For example, in a sentence like *The son of the driver that had the mustache was pretty cool* (Traxler, Pickering, & Clifton, 1998), it could be *the son* or *the driver* who has the mustache. To resolve such an ambiguity, the parser needs to distinguish between multiple available parses, for which there are both single-path ("serial") and multi-path ("parallel") mechanisms. In multi-path *competition-based models*, all parses are actively maintained and compete with one another to satisfy constraints at multiple levels. In single-path *reanalysis-based models*, there is a single maintained parse that may sometimes require revision if wrong choices are made at points of ambiguity.

Many studies have shown that global ambiguity does not result in greater processing costs (Traxler et al., 1998; Grant, Sloggett, & Dillon, 2020; Swets, Desmet, Clifton, & Ferreira, 2008; Badecker & Straub, 1994; Van Gompel, Pickering, & Traxler, 2001; Van Gompel, Pickering, Pearson, & Liversedge, 2005). This contrasts with the prediction from *the competition-based account* that computing more parses simultaneously is necessarily more costly. In fact, many of these studies on structural ambiguity have shown that parsing ambiguous strings is easier compared to parsing their unambiguous counterparts, the so-called *the ambiguity advantage effect*. An idea that has received much attention in accounting for these results is *the unrestricted race model* (URM) first proposed by Van Gompel et al. (2000). URM states that all parses engage in a stochastic race, in which the parse that happens to formulate faster will be adopted. When both parses are valid, ambiguity does not trigger reanalysis, whereas reanalysis has a higher chance of being triggered in unambiguous parses.

One special type of ambiguity is the pronominal ambiguity – the use of pronouns with multiple possible antecedents. To resolve such a pronoun, the parser is required to commit to a single referent from all available ones. Pronominal ambiguity has shown an interesting mixture of ambiguity advantage (Grant et al., 2020) and disadvantage effects (Badecker & Straub, 1994) that could be sensitive to the experimental contexts (Swets et al., 2008; Stewart, Holler, & Kidd, 2007).

Grant et al. (2020) demonstrated an ambiguity advantage for ambiguous pronouns. In two eye-tracking studies, they found that the reading time in go-past reading time in the spill-over regions is faster when two available antecedents match the gender feature of the pronoun than when only one of the two human antecedents matches the gender of the pronoun. Crucially, in both of their experiments, they included a reference selection task, which found that participants balanced their responses between two available antecedents: the comprehension selections were split half-and-half between the two antecedents. This aligns with the ambiguity advantage found in their reading times; they argue that, like structural attachment, referential dependency resolution can also engage in an unrestricted race. Unambiguous pronouns receive penalties due to the reanalysis of erroneously co-indexing to the non-matching candidates.

Grant et al. (2020)'s data hints at an important assumption of the URM in predicting the effect of ambiguity advantage: all the analyses in the race only differ slightly in speed in the computation of the parser. In Van Gompel et al. (2001)'s studies on structural ambiguity, they manipulated the occurrence of initial bias and showed that the ambiguity advantage only occurs when there is no initial bias. If there is an initial bias, which means that the two possible parses have inherently different racing speed and chances to be selected, no reanalysis is predicted as long as the biased parse is valid. Grant et al. (2020) provided half of the evidence needed to fully support the URM with balanced comprehension. Our Experiment 1 aims to test the other half of the URM's predictions by conducting response-contingent reading time analysis. We pre-

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

dict that the reading-time cost of the pronoun would be conditioned by whether the comprehension response selects the biased referent. We find that this is indeed the case: when the subject is chosen for the comprehension response, there is no ambiguity advantage; when it is not, however, the ambiguity advantage is found.

A further question raised by our Experiment 1 concerns the fact that the URM does not account for ambiguity *disadvantage* effects that were previously discovered in Badecker and Straub (1994), Swets et al. (2008), and Stewart et al. (2007). Given our results in Experiment 1, one possible explanation for these conflicting results is that experimental contexts differ in the timing and speed of the evaluation. Specifically, some tasks may prompt early commitment to a single referent, favoring the rapid, race-like mechanism proposed by the URM, while others delay referential evaluation until more information becomes available. Broader and slower evaluation may increase competition and lead to ambiguity penalties.

The question asked in Experiment 2 is thus whether task effects can modulate the occurrence of reanalysis-based processing. Assuming the URM is the underlying mechanism of the ambiguity advantage effect, can the task effect nonetheless shut it down and result in a competition penalty? We conduct a partial replication of Badecker and Straub (1994) and Grant et al. (2020). Notably, with the same set of target materials, Badecker and Straub (1994) found a competition penalty for the ambiguous condition with self-paced reading, while Grant et al. (2020) found an advantage with eye-tracking. It could be the case that with their materials, eye-tracking allows for the URM while self-paced reading discourages any possibility of reanalysis from the URM. However, because Badecker and Straub (1994) included a probe recognition task and the statistical effects were only numerical, it is premature to state this hypothesis.

Thus, in our Experiment 2, we aim to replicate Grant et al. (2020) with self-paced reading, holding all other manipulations almost the same. We predict that changing the task from eye-tracking to self-paced reading will diminish the ambiguity advantage effect, which indicates a change in the underlying mechanisms from the URM to constraint-based processing. To preview the result, we did not replicate the ambiguity advantage found in Grant et al. (2020); instead, we found a delayed ambiguity disadvantage effect that supports the delayed-interpretation account.

Experiment 1

Participants

We recruited 49 college-age (mean 21) native speakers of English enrolled in the University of California, Santa Cruz. All participants were naive to the purpose of the experiment, provided informed consent, and received credit for a linguistics class. The experiment was conducted in person and lasted about 25 minutes.

Materials

The sentence structure in our materials are shown in (1). It contains three syntactically available antecedent positions that the target pronoun can grammatically bind with. Each sentence contains a subordinate clause, a matrix clause, and an embedded clause. The three antecedent positions are the subject of the subordinate clause, the subject of the matrix clause, and the object of the matrix clause. An embedded clause follows the matrix clause and includes an inanimate subject. The pronoun appears at the object position of the embedded clause. A temporal preposition phrase is always included after the pronoun in order to ensure that the pronoun is not the last word of the sentence.

(1) Sentence: When Michael/Lisa arrived, William/Crystal mentioned to Gregory/Amy that the government would hire him/her in two weeks.

Question: Who would be hired by the government in two weeks?
a. Michael/Lisa b. William/Crystal
c. Gregory/Amy d. I'm not sure

The experiment materials consisted of 54 sets of target items and 54 filler items. Item sets were constructed following the template in (1) in a 2x2x2 factorial design that manipulated the gender match of the pronoun with highly gender-typical names occurring at each of three positions (that is, subject of the subordinate class, subject of the matrix clause, and object of the matrix clause). The [Mismatch, Mismatch, Mismatch] condition was removed to make all sentences have at least one possible parse, which leaves a total of 7 conditions. Pronoun genders and name-to-position assignments were counterbalanced across all item sets. All item sets are available via Open Science Framework: https://osf.io/efxq4/.

Procedure

Participants were tested individually using an iMac computer through PCIbex (Zehr & Schwarz, 2018). After they consent to participate in the experiment, participants first engage in six practice trials that familiarize them with the procedure of self-paced reading and referent selection. Each sentence starts with a blank underscore mask at the beginning of the trial, and each word will appear when the participant clicks the space bar on the keyboard. After each click, the previous word will be masked again, and the next word will show up. This process is repeated until the end of the sentence. After reading each sentence, a four-choice comprehension question will appear on the screen. The question always asks about the interpretation of the pronoun in the context of the embedded clause. Three antecedents will appear in separate boxes under the question in random order, as well as an additional 'I'm not sure' box. The participant will then select one answer box, after which the next trial shows up with repeated processes. The entire experiment took an average of 25 minutes to finish.

Results

Comprehension Participants largely only answered the comprehension questions with gender-matching antecedents (Table 1), which suggested they were broadly paying attention to the experiment.

We indeed found the matrix subjects to be preferred in the comprehension selection. This is indicated by the fact that they are selected significantly more often when they are compatible with the gender feature of the pronoun. In addition, even when the matrix subject mismatches the gender of the pronoun, participants made more mistakes selecting the incompatible matrix subjects than those for the subordinate clause subject and the matrix object.

This consistent preference for the matrix subject reinforces the idea that structural prominence, such as subjecthood in the matrix clause, exerts a strong influence on referent selection. Here, we add to the literature of pronoun resolution preferences that the subjecthood of the main clause can override both the first-mention effect and the locality effect (Arnold, 2001). In our construction, the subjecthood in the matrix clause carries more interpretive weight in contexts involving multiple potential antecedents [1].

Reading time (RT) Given the comprehension bias, we conducted reading time analysis for the pronoun grouped by the referent selection responses of the matrix subject. We divide the analysis into two groups. Firstly, we analyzed the reading time when the subject is selected and compatible with the gender of the pronoun. The reading time at the pronoun did not show an ambiguity advantage (Figure 1). The reading times for having three matching antecedents (XXX), two matches (XXY and YXX), and no ambiguity (YXY) do not differ. We conducted a linear mixed-effect model and found that the main effect of the number of matches is not significant ($p>0.1$). In other words, whether the other two referents match or mismatch the gender of the pronoun was irrelevant to the reading time at the pronoun.

[1] We thank an anonymous reviewer for pointing out the value of our comprehension data.

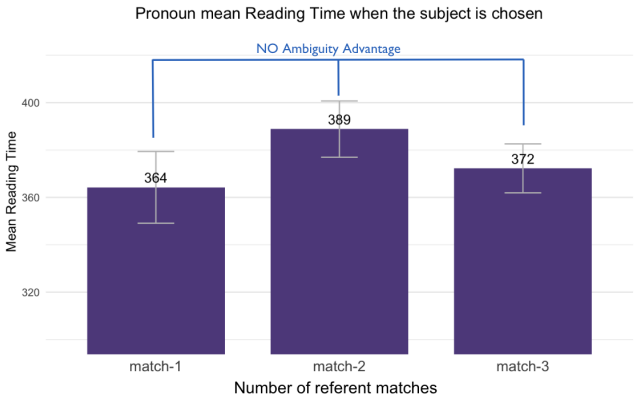

Figure 1: Mean RT at the critical region of the pronoun when the subject is selected for comprehension

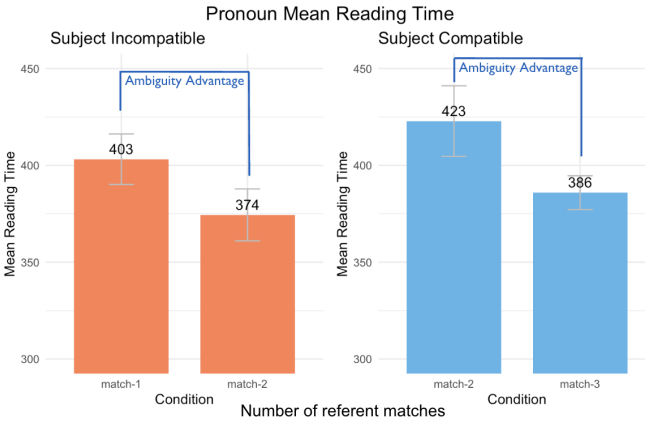

Figure 2: Mean RT at the critical region of the pronoun when the subject is NOT selected for comprehension, grouped by matrix subject compatibility.

However, in trials where the comprehender did not select the matrix subject for the comprehension question, an ambiguity advantage was found at the pronoun: faster RTs to multiple matches grouped by the availability of the subject (Figure 2). This effect is found both when the matrix subject matches and mismatches the gender feature of the pronoun.

Table 1: Comprehension percentage by conditions

	Subject Matches Pron.				Subject Mismatches Pron.		
	All Match	Two Match		Single Match	Two Match	Single Match	
X = Referent **matches** pronoun's gender **Y** = Referent **mismatches** pronoun's gender							
Response %	XXX	XXY	YXX	YXY	XYX	YYX	XYY
Sub. subject (Name 1)	21	29	2	4	27	2	**83**
Matrix subject (Name 2)	**39**	**56**	49	**86**	10	5	9
Object (Name 3)	24	3	**38**	4	**50**	**87**	3
I'm not sure	16	12	11	5	14	5	5

When it is incompatible (on the left, match-3 XXX condition is excluded because the matrix subject is compatible), having one more matching antecedent resulted in faster reading time. Moreover, even when the matrix subject is compatible with the gender of the pronoun (on the right, the match-1 YXY condition is excluded due to only 8% of the trials not selecting the matrix subject, see Table 1), the same effect is also found. Paired-wise tests were conducted for both groups and revealed a marginal significance for the subject unavailable group (p=0.06, 95% CI= (-76.7ms, 2.8ms)) and a significant effect for the subject available group (p=0.03, 95% CI=(-53.2ms, -2.6ms)).

Discussion

We found that pronominal ambiguity resolution is structured by the comprehension selection of the preferred referent. Our comprehension data confirmed the preference for the matrix subject in our manipulation. When we analyze the reading time data contingent upon the selection of the matrix subject, we found an ambiguity advantage effect only when the matrix subject is NOT chosen, no matter whether it matches the gender of the pronoun or not. When the matrix subject is selected as the antecedent, we do not see the same ambiguity advantage effect. Our result is summarized in Table 2.

Table 2: Summarized effects of Experiment 1

State of the biased referent		**Processing Effect**
Compatible?	**Selected?**	
–	–	Advantage
+	–	Advantage
+	+	No Advantage

This result is compatible with the unrestricted race model. According to the unrestricted race model, we should not expect the presence of the ambiguity advantage effect for all ambiguous constructions, but only for those that do not involve an initial bias. And this is exactly what we see in our current study.

It is particularly noteworthy that the ambiguity advantage does not depend merely on whether the matrix subject matches the pronoun's gender. We still observed an advantage effect even when the matrix subject was not selected as the antecedent. In fact, such trials constitute an average of 52% across the XXX, YXX, and XXY conditions. And in these trials, we see the same ambiguity advantage effect as when the preferred subject is incompatible. This suggests that what matters is not just the compatibility of the biased referent, but whether comprehenders actually select it as the referential target.

One way to explain the underlying mechanism of the URM is the Centering Theory (Grosz, Joshi, & Weinstein, 1995; Gordon & Chan, 1995), which posits that discourse processing is guided by a hierarchy of referential centers. According to this view, a "center" of attention is maintained in a heightened state of cognitive accessibility. When there is a determined center of the discourse, it will be immediately attached to the pronouns, resulting in no advantage of ambiguity. In contrast, when no referent emerges as the center, the parser engages in the URM.

Another way to explain this is through the activation strength. In this account, referents in discourse are encoded with varying levels of activation. A biased or structurally prominent referent, such as the matrix subject in our experiment, receives a substantially stronger activation boost compared to the others. It thus wins the race immediately, minimizing the chance of the other reference being considered, regardless of the compatibility of other referents.

Next, we will test whether the URM is sensitive to task manipulation. While our results provide additional support for the URM, the URM does not predict any ambiguity penalty. In Experiment 2, we test whether ambiguity resolution strategies are fixed across tasks. We wonder whether changing the task manipulation shifts the parser away from early commitment mechanisms like the URM by replicating Grant et al. (2020) with self-paced reading while holding all other manipulations almost the same.

Experiment 2

Participants

90 native English speakers were recruited via Prolific for compensation. All participants were naive to the nature of the study. Four participants were excluded due to unusual reading time patterns[2], leaving 86 participants in the analysis.

Materials

24 items and 48 fillers were adopted from Experiment 2 in Grant et al. (2020). An example item set is shown in Table 3. In each item, the critical pronoun is at the subject position of the embedded clause, while both referents are introduced in the matrix clause, one as the subject and one as the object. The referents in half of the items are gender-typical proper names (e.g., *Grace*), and the other half contains definite descriptions that had definitional gender (e.g., *the king*). There are three conditions in the study: an ambiguous condition in which the two referents both share the same gender of the pronoun, a subject-match condition in which only the subject referent but not the object referent matches the gender of the pronoun, and an object-match condition in which only the object referent matches the gender of the pronoun. The comprehension questions for the items were constructed to probe the interpretation of the pronoun, and those for fillers ask either who is the agent of an action or some general information relevant to the sentence. A sample target item set is shown in Table 3.

[2]We discovered four participants with consistent patterns of unusually long reading time (>500ms) for the first word of each sentence, followed by unusually short reading times (<100ms) for the rest of the sentence.

Table 3: Sample itemset in Grant et al. (2020), Experiment 2

Condition	Item
Ambiguity	The young prince$_i$ showed the revered king$_j$ that he$_{i/j}$ would be a fine leader of the Tharassian empire.
Subject match	The young prince$_i$ showed the revered queen$_j$ that he$_{i/*j}$ would be a fine leader of the Tharassian empire.
Object match	The revered queen$_i$ showed the young prince$_j$ that he$_{*i/j}$ would be a fine leader of the Tharassian empire.

Procedure

The procedure largely follows Experiment 2 in Grant et al. (2020), which is also similar to the procedure in our Experiment 1. The study is presented in PCIbex (Zehr & Schwarz, 2018). In each trial, participants first read a sentence one word at a time by pressing the space bar on the keyboard. After reading the sentence, a random half of all trials (i.e., 12 items and 24 fillers) will be followed by a speeded two-choice comprehension question. Participants were instructed to always place their index fingers on keys F and J and select a comprehension answer by pressing the corresponding key as soon as possible. Word-by-word reading times, comprehension reaction time, and comprehension selection were recorded.

Results

Comprehension

Table 4 summarizes the mean proportions and response times for each condition. The overall selection proportions are comparable to the results reported in Grant et al. (2020). In the unambiguous conditions, participants mostly selected gender-matching antecedents. In the ambiguity condition, the selection is balanced between the two antecedents, which confirms the comprehension results in Grant et al. (2020). In the two unambiguous conditions, the feature-matching antecedents are selected about the same amount, further indicating the lack of bias in the items. Note that, unlike Experiment 1, because of the use of *transfer-of-information* verbs, we did not observe a subjecthood bias in this study. This interaction of subjecthood prominence with verb semantics is found in (Kaiser, (2009), cf. Grant et al., (2020)).

The comprehension selection reaction times are also presented in Table 4. On average, participants took longer to respond in the ambiguous trials compared to the unambiguous ones. Notably, there was no significant difference in reaction times between the two unambiguous conditions. In both of these conditions, participants overwhelmingly selected the feature-matching antecedent, and it is costly to select the less-selected mismatching antecedents. Interestingly, in the ambiguous condition, participants took more time to select the object antecedent than the subject antecedent, despite both options being chosen at comparable proportions.

Table 4: Comprehension selection percentages and mean reaction times (ms) by condition

Condition	Subject	Object	Average RT
Ambiguity	51.3%	48.7%	
	2958	3304	3126
Subject match	81.5%	18.5%	
	2721	3678	2899
Object match	17.9%	82.1%	
	3318	2829	2917

We also attempted a comprehension-contingent analysis on reading time, but we found some caveats that discourage the same analysis. The comprehension bias of multiple items were created by unintended parts of the sentence, which means that the reading time at the pronoun and its spill-over regions may not reflect accurate comprehension-related effects. Thus, we refrain from reporting response-contingent analysis for Experiment 2.

Reading time

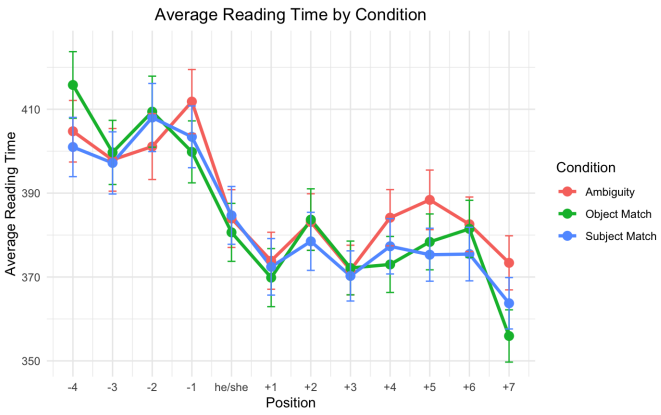

Figure 3: Mean RT in pre-critical, critical, and spill-over regions

We did not find the same ambiguity advantage effect shown in Grant et al. (2020). Instead, our results show a delayed ambiguity penalty. A linear mixed-effect model was run at positions +4 and +5 as indicated in Figure 3 with the main effect of the number of matches and a random effect of participant. We found a significant main effect of number of matches ($\beta = 9.527$, $SE = 3.796$, $p = 0.012$).

Discussion

To summarize, our Experiment 2 replicated the comprehension results reported in Grant et al. (2020) but not the real-time reading time effects. We found a delayed ambiguity

penalty despite finding no bias between the subject and object in the comprehension questions.

Notably, the competition penalty in our study emerged relatively late in the fourth and fifth words after the pronoun region. This delayed effect indicates that the resolution process in our study may arise not at the point of encountering the pronoun, but during the integration of downstream content that forces resolution. The items we used, taken from Grant et al. (2020), indeed have many items in which the content after the pronoun may have influenced the interpretation. For example, consider *The elderly gentleman warned the young man that he was unlikely to be hired by the advertising firm. A young man* is more likely to be hired than *an elderly gentleman*, but this information only comes well after the pronoun. Our result can thus be explained by delayed-interpretation accounts, in which referential commitment is postponed until comprehension demands require disambiguation.

Our reading time effect is in apparent conflict with Grant et al. (2020)'s finding. One obvious possibility for this disparity is the different task manipulations. We used self-paced reading, while they used eye-tracking to measure word-by-word reading times. We also had significantly fewer filler items, about one-third of the amount from Grant et al.'s original study. Perhaps this enabled participants to better guess the goal of the study and devote more effort to interpreting the pronoun. We thus conjecture that the processing of ambiguous pronouns will be very sensitive to task manipulations and procedures.

General Discussion

In Experiment 1, we observed that the ambiguity advantage effect is modulated by the comprehension outcomes, which supports the reanalysis-based unrestricted race processing model for resolving pronominal ambiguity. However, Experiment 2 did not replicate the ambiguity advantage reported by Grant et al. (2020) and misaligned with the predictions made by the URM despite closely following their materials. Instead, our Experiment 2 data fits better with a delayed-interpretation model.

Our Experiment 1 suggests a potential refinement to the URM mechanism. A key finding from Experiment 1 is that the ambiguity advantage effect is not merely modulated by whether a referent is structurally biased and feature-compatible. Rather, it depends on whether that referent is actually selected in interpretation. We observed a clear ambiguity advantage even when the matrix subject matched the gender of the pronoun as long as it was not selected during comprehension. This indicates that the URM is not just about the initial bias. If the parser tentatively evaluates multiple options through a URM race and does not strongly commit to the biased one, then the ambiguity can still confer a processing benefit. When it does, however, either the parser limits its evaluation space to only the biased referent or the other referents race with significantly slower speed, eliminating the ambiguity advantage effect.

We observed a delayed penalty for the ambiguity condition in Experiment 2, which can be explained by a constraint-based delayed-interpretation account. In this view, comprehenders only resolve the pronoun when there is enough downstream information. This posits that the processing mechanism of ambiguous pronouns is variable depending on the task. In some cases, the URM is triggered, while in others, it involves a constraint-based mechanism.

This result reinforces the task-dependent processing effects (Swets et al., 2008; Stewart et al., 2007). Although the comprehension was balanced in the ambiguous condition, suggesting no dominant referent, the task difference between our replication and Grant et al. (2020) may have pushed comprehenders away from the unrestricted race. Comparing the methods between our two experiments, both used self-paced reading and only differed in the form of the comprehension question. Our Experiment 1 had a multiple-choice question without a response time measure, while Experiment 2 had speeded binary judgment. We argue that ambiguous pronoun resolution may be more sensitive to task effects than previously assumed.

One possible connection between the URM mechanism (Van Gompel et al., 2001, 2005) with the depth of processing (Stewart et al., 2007; Swets et al., 2008) may be characterized by the evaluation space and the evaluation speed. Essentially, the URM allows for reshaping the evaluation space at any point and commits to a referent as fast as possible. The risk of reanalysis is allowed, such that the ambiguity advantage is sensitive to referent bias. However, under task pressure, the parser tends to avoid the risk of early commitment and instead delays the resolution until sufficient downstream information is available. All candidate referents are maintained in the evaluation space until late-stage selection, resulting in delayed penalties. Taken together, these findings point to a flexible ambiguity resolution mechanism that depends on how referential candidates are mentally weighted and the timing of selection.

It is worth noting that our Experiment 1 results are not contradictory with the strategic underspecification account (Swets et al., 2008). When the matrix subject was not selected, the parser may have refrained from committing to any specific referent, resulting in faster reading times under ambiguity. However, it may be harder for this account to explain the laddered ambiguity advantage observed in Experiment 1, comparing when there are three matching antecedents than when there are two. Both of them are ambiguous, but their processing costs are nonetheless different, which requires some refined underspecification to account for.

Through evidence from the real-time processing cost with referent bias and task effects, we offer a more comprehensive view of pronominal ambiguity resolution. Future work should manipulate the timing and depth of referential commitment more directly to further test the boundaries between facilitative and competitive processing under ambiguity.

Acknowledgment

We would like to thank the four anonymous reviewers, the audience at the 38th Annual Conference on Human Sentence Processing, the members of s/lab at UC Santa Cruz, and the participants of the graduate research seminar (LING290) at UC Santa Cruz for their valuable feedback on the project. This project also benefited from discussions with Amanda Rysling, Dustin Chacón, and Subhekshya Shrestha. M. Wagers gratefully acknowledges the support from the NSF, Award #2019804, to UC Santa Cruz.

References

Arnold, J. E. (2001, March). The Effect of Thematic Roles on Pronoun Use and Frequency of Reference Continuation. *Discourse Processes*, *31*(2), 137–162. Retrieved 2025-01-21, from https://doi.org/10.1207/S15326950DP3102$_0$2 (Publisher: Routledge _eprint: https://doi.org/10.1207/S15326950DP3102_02) doi: 10.1207/S15326950DP3102$_0$2

Badecker, W., & Straub, K. (1994). Evidence that binding principles participate in a constraint satisfaction process. *Poster session presented at the Seventh annual CUNY Sentence Processing Conference*.

Gordon, P. C., & Chan, D. (1995). Pronouns, passives, and discourse coherence. *Journal of Memory and Language*, *34*(2), 216–231. (Publisher: Elsevier)

Grant, M., Sloggett, S., & Dillon, B. (2020). Processing ambiguities in attachment and pronominal reference. *Glossa: a journal of general linguistics*, *5*(1). (Publisher: Open Library of Humanities)

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A Framework for Modeling the Local Coherence of Discourse. *Computational Linguistics*, *21*(2), 203–225. Retrieved 2025-02-03, from https://aclanthology.org/J95-2003/ (Place: Cambridge, MA Publisher: MIT Press)

Kaiser, E. (2009). Effects of Anaphoric Dependencies and Semantic Representations on Pronoun Interpretation. In S. Lalitha Devi, A. Branco, & R. Mitkov (Eds.), *Anaphora Processing and Applications* (pp. 121–129). Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-04975-0$_1$0

Stewart, A. J., Holler, J., & Kidd, E. (2007). Shallow processing of ambiguous pronouns: Evidence for delay. *Quarterly Journal of Experimental Psychology*, *60*(12), 1680–1696. (Publisher: SAGE Publications Sage UK: London, England)

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition*, *36*(1), 201–216. (Publisher: Springer)

Traxler, M. J., Pickering, M. J., & Clifton, C. (1998, November). Adjunct Attachment Is Not a Form of Lexical Ambiguity Resolution. *Journal of Memory and Language*, *39*(4), 558–592. doi: 10.1006/jmla.1998.2600

Van Gompel, R. P., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, *52*(2), 284–307. (Publisher: Elsevier)

Van Gompel, R. P., Pickering, M. J., & Traxler, M. J. (2000). Unrestricted race: A new model of syntactic ambiguity resolution. In *Reading as a perceptual process* (pp. 621–648). Elsevier.

Van Gompel, R. P., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language*, *45*(2), 225–258. (Publisher: Elsevier)

Zehr, J., & Schwarz, F. (2018, March). PennController for Internet Based Experiments (IBEX). Retrieved 2025-01-28, from https://osf.io/md832/ (Publisher: OSF) doi: 10.17605/OSF.IO/MD832