WHAT TAGALOG CAN TEACH US: THE INFLUENCE OF WORD ORDER IN REFLEXIVE PROCESSING*

Jed Sam Pizarro-Guevara
University of Massachusetts, Amherst
jpguevara@umass.edu

Brian Dillon
University of Massachusetts, Amherst
bwdillon@umass.edu

We investigated the interference profile of reflexive dependencies in Tagalog. We conducted a self-paced reading experiment using a number mismatch paradigm to examine the extent to which Tagalog comprehenders attend to the number features of the distractor (i.e., a non-c-commanding NP inside a relative clause modifying the antecedent). Our results indicated that even though the number feature of the distractor had very little impact on their final interpretations of the reflexive, comprehenders were attending to them in real-time. We discuss the implications of our results for models of cue-based retrieval.

1. Introduction

Many elements in natural language depend on other linguistic elements in their context for interpretation. Reflexive pronouns like *himself* are one example. Consider (1).

(1) Mario$_i$ said that Luigi$_j$ slapped himself$_{j/*i/*k}$

In English, *himself* in (1) can only refer to *Luigi*—not *Mario* or some other unmentioned entity. In generative linguistics, one well-known account of how we interpret reflexives is Principle A of the Binding Theory (Chomsky 1981). In (2) is a (simplified) reformulation of Principle A as conditions on the reflexive's antecedent:

(2) The antecedent must (i) be in the same clause as the reflexive, and (ii) c-command it

While Principle A captures the final interpretation of sentences like (1), its role in resolving a reflexive's reference in real-time remains an open question. This is the focus of the present paper.

Principle A could potentially guide how we comprehend reflexives in real-time. First, it identifies where comprehenders should search for the antecedent, and so can usefully constrain the search for an antecedent. Second, after the search space has been constrained, it also identifies what structural configuration has to hold for an element to be able to be the antecedent.

Researchers have probed the question of Principle A's role in guiding real-time processing by varying the features of potential antecedents that are not licensed by Principle A, and then looking at how such changes impact the processing of the reflexive. Consider (3).

*We thank Henrison Hsieh, Kristina Gallego, Rowena Garcia, and Aldrin Lee for helping us with the facts of the language; Michael Manahan for helping us with a pilot study of this experiment; and the audience at AFLA 28 for their comments and questions. All mistakes are our own. This work was supported by the National Science Foundation under NSF BCS 1941485 to BD.

(3) a. The man that the boy punched embarrassed himself
b. The man that the boys punched embarrassed himself

The potential antecedent that is not licensed by Principle A is singular in (3a); it is plural in (3b). To the extent that there is a difference between how *himself* is processed in (3a) and (3b) gives us a measure of the extent to which comprehenders attend to the features of potential antecedents that are not grammatically accessible. Any influence exerted by the features of these grammatically inaccessible NPs on the processing of a reflexive is broadly referred to as an INTERFERENCE EFFECT. More on this in section 2.1.

The present study investigates how reflexives are processed in Tagalog. Previous research has shown that Tagalog reflexives are subject to similar constraints as their English counterparts: the antecedent needs to be in the same clause as the reflexive and to c-command it in their base position (Richards 2013). Consider (4).

(4) Sinabi ni Mario na sinampal ni Luigi ang kaniyang sarili
say GEN PN LNK slap GEN PN NOM reflexive
"Mario$_i$ said that Luigi$_j$ slapped himself$_{j/*i/*k}$

Much like in English, the reflexive *kaniyang sarili* can only refer to *Luigi*—not to *Mario* or some other unmentioned entity. In this study, we are interested in whether—and to what extent—Tagalog comprehenders attend to the number feature of a potential antecedent that is not grammatically licensed. To our knowledge, the present study is one of the only two experimental investigations of reflexives in Tagalog—the other being Bondoc et al. 2019—and this is the first to investigate the interference profile in reflexive processing in the language.

The remainder of the paper is organized as follows. Section 2 provides an overview of cue-based retrieval as a framework for how comprehenders form linguistic dependencies in real-time. We then review some of the evidence used to argue for a model of retrieval that prioritizes structural cues over non-structural cues in reflexive processing. We then provide an alternative hypothesis that assigns no special status to structural cues in reflexive processing. Under this hypothesis, the Principle A-like effects observed in the literature are only apparent; they arise because of a potential confound in the experimental items used. Section 3 presents the self-paced reading experiment designed to factor out the potential confound by leveraging Tagalog's word order. Section 4 concludes by (i) discussing our findings and contextualizing what the empirical landscape in Tagalog can tell us about the influence of word order in reflexive processing and more generally, cue-based retrieval; and (ii) providing an overview of future experiments.

2. Cue-based retrieval

One of the central questions in psycholinguistics is how comprehenders form linguistic dependencies in real-time, that is, as a sentence unfolds over time. One view is that we rely on a cue-based retrieval mechanism to form linguistic dependencies in real-time (Van Dyke and Lewis 2003; Lewis et al. 2006). When we encounter a dependency that needs to be resolved, we use certain features of the element that instantiated the dependency to look for another element in working memory that could resolve it. Consider the example in (5), adapted from Dillon et al. 2013.

(5) The new executive$_{[\text{SG, +CMS}]}$ who oversaw the middle managers$_{[\text{PL, –CMS}]}$ apparently was$_{[\text{SG, CMS}]}$ dishonest about the company's profits.

Upon encountering the matrix verb *was* in (5), a comprehender knows that she has to resolve subject-verb agreement. She uses the features of the verb—[SG, CMS]—to search for the co-dependent element in working memory.[1] Because *the new executive*, which has the features [SG, +CMS], fully matches the features needed by the verb, it is retrieved from memory and is used to resolve the subject-verb dependency.

Before proceeding, here are some terms that will be used throughout. The term RETRIEVAL will be used to refer to the process of looking for an element in working memory that could resolve the dependency. In (5), this is the operation launched by the verb *was*. The term RETRIEVAL CUE will be used to refer to the features that launched the retrieval operation. In (5), this is the feature [SG, CMS] of the verb. The term TARGET will be used to refer to the element that resolves the dependency grammatically. In (5), this is the noun phrase *the new executive*, which has features that fully match the retrieval cues. Finally, the term DISTRACTOR(S) will be used to refer to the other elements in working memory that could potentially resolve the dependency but do not fully match the retrieval cues. In (5), this is the noun phrase *the middle managers*.

2.1. Interference effects

Retrieval is fast and associative, allowing comprehenders to rapidly form linguistic dependencies to support real-time comprehension. However, one of the disadvantages of this operation is that it is subject to what is called SIMILARITY-BASED INTERFERENCE. That is, retrieval errors can arise when features of the distractor match the retrieval cues. Research has identified two types of interference effects that can arise during dependency formation. Following Dillon 2011, we refer to the two types as INHIBITORY and FACILITATORY interference. Figure 1 summarizes the different types of interference effects and the contexts in which they can arise.

In the original Lewis and Vasishth model (Lewis and Vasishth 2005, henceforth, LV05), these interference effects are predicted to occur in different contexts (Engelmann et al. 2019). Inhibitory interference occurs when access to the target is disrupted because its features (partially) overlap with the features of other items in memory (i.e., the distractors). This overlap manifests as increased difficulty during dependency formation. Consider a sentence like (5), repeated here as (6a), and a minimally different sentence like (6b).

(6) Adapted from Dillon et al. 2013

a. The new executive$_{[\text{SG, +CMS}]}$ who oversaw the middle managers$_{[\text{PL, –CMS}]}$ apparently was$_{[\text{SG, CMS}]}$ dishonest about the company's profits.
b. The new executive$_{[\text{SG, +CMS}]}$ who oversaw the middle manager$_{[\text{SG, –CMS}]}$ apparently was$_{[\text{SG, CMS}]}$ dishonest about the company's profits.

In (6a), the features of the target (i.e., *the new executive*) completely match the retrieval

[1] The set of features comprehenders attend to and use during retrieval remain an area of active research in sentence processing (Parker et al. 2017). For this toy example, to simplify exposition, we assume that the verb wants something that has the following features: singular, abbreviated as SG, and a subject in the same clause, abbreviated as CMS for clausemate subject. Note that the feature CMS is only a placeholder for the constellation of features that constitute what it means to be a subject and another feature that differentiates the matrix subject from the subject inside the RC.

cues of the verb *was*; this is sometimes referred to as a grammatical (Wagers et al. 2009) or a TARGET MATCH configuration (Jäger et al. 2017; Engelmann et al. 2019). On the other hand, the features of the distractor (i.e., *the middle managers*) do not match any of the retrieval cues. Because there is no overlap between the features of the target and the distractor, this serves as the baseline for how subject-verb dependencies are resolved. In (6b), the features of the target again completely match the retrieval cues of the verb *was*. The features of the distractor now partially match the retrieval cues, namely they match in number. Van Dyke (2007), among others, found that this partial matching distractor causes a slowdown at the retrieval site. In LV05, this slowdown has been attributed to the spreading of activation assumed in the ACT-R architecture. The partial matches reduce the strength of association between the cue and the target. Thus, on average, it will take longer to retrieve the correct element (Engelmann et al. 2019).

By contrast, facilitatory interference occurs when there is no element in working memory that fully matches the retrieval cues and there are distractors that partially match them; this is variably called an ungrammatical or a TARGET MISMATCH configuration. This overlap manifests behaviorally as a speed-up during dependency formation. Consider a sentence like (7a), and a minimally different sentence like (7b).

(7) Adapted from Dillon et al. 2013

a. *The new executive$_{[SG, +CMS]}$ who oversaw the middle manager$_{[SG, -CMS]}$ apparently were$_{[PL, CMS]}$ dishonest about the company's profits.

b. *The new executive$_{[SG, +CMS]}$ who oversaw the middle managers$_{[PL, -CMS]}$ apparently were$_{[PL, CMS]}$ dishonest about the company's profits.

In (7a), the features of the intended target (i.e., *the new executive*) only partially match the retrieval cues of the verb *was*. The features of the distractor (i.e., *the middle manager*) do not match any of the retrieval cues. Researchers have used this as a baseline for how ungrammatical subject-verb dependencies are read. In (7b), the features of the intended target again only partially match the retrieval cues of the verb. Now the features of the distractor partially match the retrieval cues of the verb. Wagers et al. (2009), among many others, have found that this partial matching distractor causes a speedup at the retrieval site. Because there is no element that fully matches the retrieval cues, all partial matches are considered for retrieval. Thus, retrieval of any of the two items in (7b) will finish faster on average, compared to the retrieval of the one partial matching item in (7a), due to a higher probability of retrieving partial matches. This is a process referred to as STATISTICAL FACILITATION (Vasishth et al. 2019).[2]

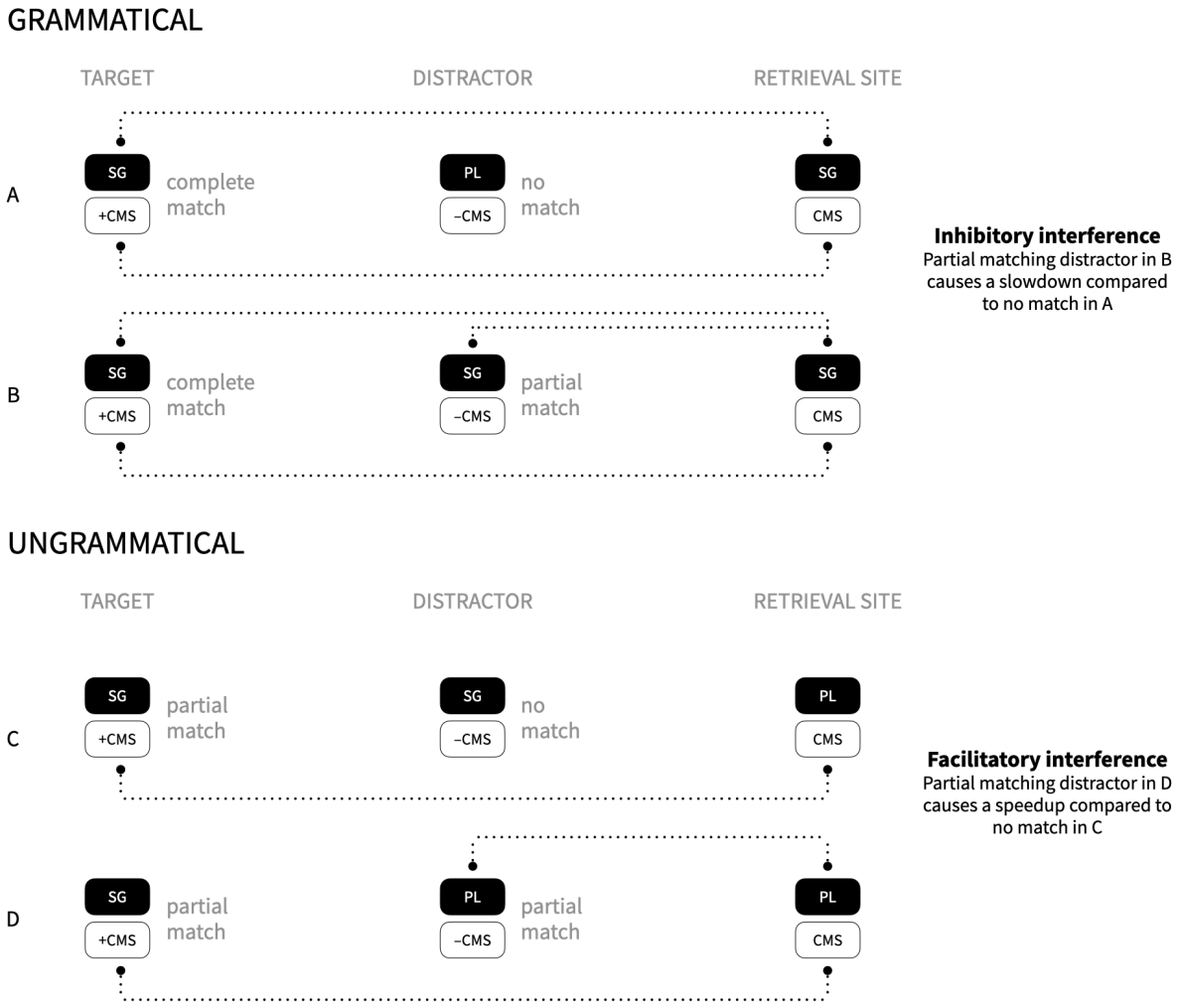
The model of retrieval discussed above assumes that all cues are weighted equally. Thus, we refer to it as a SIMPLE RETRIEVAL MODEL to contrast with the class of models that assume higher weighting of structural cues over non-structural ones. We refer to the latter models of retrieval as a SYNTACTICALLY GUIDED MODEL and they will be discussed in greater detail next.

2.2. Syntactically guided retrieval in reflexive processing

The example discussed above to illustrate the two types of interference effects involved subject-verb dependencies. However, cue-based retrieval is a general model of how dependencies are

[2] There are other possible linking hypotheses between the operations of a cue-based retrieval parser and patterns of facilitation/inhibition in real-time processing (Nicenboim and Vasishth 2018). For present purposes, we focus only on activation-based implementations of cue-based retrieval. We leave it to future work to explore direct-access models.

Figure 1: A schematization that summarizes the different types of interference effects and the contexts in which they arise. This figure is adapted from Jäger et al. 2020



resolved. Thus, it has also been used to account for how other linguistic dependencies are resolved, like reflexive dependencies (Dillon et al. 2013), verb-phrase ellipsis (Martin and McElree 2009), negative polarity items (Vasishth et al. 2008), among others. Here, we discuss how reflexive dependencies are resolved. Consider the sentence in (8), adapted from Dillon et al. 2013.

(8) The new executive$_{[\text{SG, +CCOM}]}$ who oversaw the middle managers$_{[\text{PL, –CCOM}]}$ apparently doubted himself$_{[\text{SG, CCOM}]}$ on most major decisions.

Even though the element that launches retrieval in example (8) differs from the element that launches retrieval in example (5), the idea remains the same. Upon encountering the reflexive *himself*, a comprehender knows that she has to resolve the reflexive dependency. She uses the features of the reflexive, which for ease of exposition, we assume to be [SG, CCOM], to search for

the co-dependent element in memory.[3] Because the features of *the new executive* fully matches the retrieval cues of the reflexive, it is retrieved from memory and used to resolve the dependency.

However, some researchers have argued that in the domain of reflexive processing, the retrieval operation at the reflexive is syntactically guided (Sturt 2003; Xiang et al. 2009; Dillon et al. 2013). Comprehenders deploy their knowledge of Principle A of the Binding Theory (Chomsky 1981) to limit the domain that they must search to find an antecedent. Proponents of a syntactically guided model of retrieval have conceived of how comprehenders apply their knowledge of Principle A in one of two ways. The first is that it functions as a categorical filter (e.g., Dillon et al. 2013). The second takes a more intermediate position by assuming differential cue weighting for structural and non-structural cues, where structural cues are weighted higher than non-structural ones (e.g., Kush 2013; Parker et al. 2017), creating a preference for using syntactic cues to resolve dependencies but stopping short of claiming syntax constitutes a categorical filter.

Proponents of a syntactically guided model of retrieval have used the difference between the interference profile of ungrammatical reflexive dependencies, as in example (9), and that of ungrammatical subject-verb dependencies, as seen previously in example (7).

(9) Adapted from Dillon et al. 2013

a. The new executive$_{[\text{SG, +CCOM}]}$ who oversaw the middle manager$_{[\text{SG, –CCOM}]}$ apparently doubted themselves$_{[\text{PL, CCOM}]}$ on most major decisions.

b. The new executive$_{[\text{SG, +CCOM}]}$ who oversaw the middle managers$_{[\text{PL, –CCOM}]}$ apparently doubted themselves$_{[\text{PL, CCOM}]}$ on most major decisions.

In (9a), the features of the intended target only partially match the retrieval cues of the reflexive; the features of the distractor do not match any of the retrieval cues. Thus, the example in (9a) has the same characteristics as the example in (7a) or Panel C in Figure 1. In (9b), the features of both the target and the distractor partially match the retrieval cues. Thus, the example in (9b) has the same characteristics as the example in (7b) or Panel D in Figure 1. Despite their similarities, many earlier researchers have found no evidence for facilitatory interference in reflexive dependencies (Sturt 2003; Xiang et al. 2009; Dillon et al. 2013). In fact, one meta-analysis found that reflexives showed inhibitory interference effects in target mismatch configurations, that is, when the features of the target fully mismatch the retrieval cues (Jäger et al. 2017). There are, however, some studies that have provided clear evidence for facilitatory interference in reflexive resolution. Parker et al. (2017) showed that facilitatory interference can be seen when the target is a very poor feature match to the retrieval cues. In a more recent high powered study, Jäger et al. (2020) have found some evidence for facilitatory interference in target-mismatch configurations for reflexives, while still observing some evidence for a contrast between subject-verb agreement and reflexive dependencies. At present, more research is necessary to clarify the empirical record, as it remains unclear what type or how much interference impacts reflexive processing in English.

[3] We acknowledge that SG and CCOM are inherently different features. SG is static and can exist independent of the other elements in the sentence. From the perspective of an incremental comprehender, upon encountering an NP, she knows whether to encode SG or PL. By contrast, CCOM is relational and can only exist relative to another element: "no item can be a c-commander per se, but can only be in a c-commanding relation with another syntactic constituent" (Jäger et al. 2017). From the perspective of an incremental comprehender, upon encountering an NP, she does not know whether the feature ±CCOM is a feature that even needs to be encoded since its presence needs to be evaluated against another element. This look-ahead problem is an implementational question that cue-based parsing models face when relational features are involved. It is outside the scope of our paper to pursue the question of how an incremental comprehender can dynamically mark constituents with relational features. See Kush 2013 for one implementation.

An important implication of the different interference profiles of subject-verb and reflexive dependencies is that the comprehender uses different memory operations depending on the linguistic context, or differentially weights various sources of information for different dependencies. In the next subsection, we provide an overview of an alternative explanation for some of the effects discussed above, which will be referred to as the RECENT ACTIVATION HYPOTHESIS. The recent activation hypothesis holds that some of the observed differences between subject-verb and reflexive dependencies can be attributed to differences in how activated different elements are across a parse. One attractive feature of this alternative is that there is no need to assign special status to structural cues. Thus, the comprehender uses the same memory operation irrespective of linguistic context.

2.3. Recent activation hypothesis: An alternative to a syntactically guided retrieval

One potential confound in the literature of reflexive processing is that the experimental items used in most of the previous studies involved reflexives occupying the position immediately after the verb (King et al. 2012; Dillon et al. 2013; Kush and Phillips 2014). This word order configuration could give the reflexive privileged access to the target (i.e., the antecedent) via recent activation. Consider (10). Circled numerals represent relevant timepoints where the target is retrieved.

(10) Adapted from Dillon et al. 2013
   a. The new executive who oversaw the middle managers apparently (1) doubted (2) himself on most major decisions.
   b. The new executive who oversaw the middle manager apparently (1) was dishonest about the company profits.

For reflexive-dependencies, as in (10a), the target is reactivated at time point (1), prior to the critical retrieval, which is at time point (2). Upon encountering the matrix predicate, comprehenders retrieve the target for thematic integration. Because they encounter the reflexive immediately after, the target could still be enjoying a high level of activation and this could confer an advantage over the distractor at the critical retrieval operation at time point (2). This advantage could either (i) ensure the correct retrieval of the target by giving the target an overwhelming activation advantage or (ii) make retrieval unnecessary upon encountering the reflexive. In either case, comprehenders would have privileged access to the target and its features with no need to posit syntactically guided retrieval cues. Crucially, as we see in (10b), the target is not retrieved immediately prior to the critical retrieval.

The recent activation hypothesis predicts interference effects to obtain when the reflexive is not immediately adjacent to the predicate. In an eye-tracking-while-reading experiment, King et al. (2012) provided evidence consistent with this prediction. They compared sentences like in example (11).[4] They leveraged the property of some predicates in English to undergo dative alternation. This allowed them to compare how reflexives were processed when they were immediately

[4] Note that the design used by King et al. (2012) differs from the examples we have seen thus far from Dillon et al. (2013). Instead of manipulating grammaticality by varying the number feature, they manipulated prototypical gender congruence between the target and the reflexive. For example, in (11), the target *bricklayer* is stereotypically masculine, per their norming study. When *himself* is encountered downstream, the reflexive and the target matched. When it was *herself*, the reflexive and the target mismatched. The distractors *Gregory* or *Helen* match or mismatch the gender features of the reflexive.

adjacent to the predicate, as in (11a), and when they were not, as in (11b). They found no evidence for interference when the reflexive was immediately post-verbal, as in (11a). However, they found evidence for interference when there was more distance between the verb and the reflexive, as in (11b).

(11) Adapted from King et al. 2012
   a. The bricklayer who employed Gregory/Helen shipped himself/herself sacks of mortar...
   b. The bricklayer who employed Gregory/Helen shipped sacks of mortar to himself/herself...

As mentioned previously, an attractive feature of the recent activation hypothesis is that there is no need to assign special status to structural cues and that there is no need to posit different memory operations for different linguistic contexts.

3. The present study

We explore the possibility that the recent activation hypothesis could contribute to difficulty observing interference effects in reflexive dependencies. In other words, the difference between the interference profiles of ungrammatical reflexive dependencies and that of ungrammatical subject-verb dependencies could be an artifact of the word order configuration of the experimental items used in these studies, as discussed in the previous section. In the present study, we leverage the word order properties of Tagalog to maximize the chances that the extent to which we observe interference effects in reflexive processing cannot be reduced simply to the recent activation of the target immediately before the critical retrieval operation (i.e., at the reflexive). Our goal is to investigate the interference profile in Tagalog after controlling for recent activation of the target as a potential confound. Consider a schematization of our experimental items in (12). Circled numerals represent time points where something relevant happens.

(12) Verb ① Target [$_{RC}$ Verb Distractor ...] ② Reflexive ...

Upon encountering the target at time point ①, it can immediately be thematically integrated with the matrix verb. There is a relative clause intervening between the target and the reflexive, creating temporal distance from when thematic integration occurred (i.e., at time point ①) and when retrieval for antecedent resolution is launched (i.e., at time point ②). We argue that this word order configuration allows us to factor out the contribution of the target being recently activated and independently investigate the retrieval process at the reflexive.

Empirically, the present study asks two questions: (i) To what extent do comprehenders attend to the number feature of the target? (ii) To what extent do they attend to the number feature of the distractor?

3.1. Participants

We recruited 80 participants, ranging from 18 to 63 years old (M = 31), through the Prolific Academic online platform (`https://www.prolific.co`). Of these, 64 of them self-reported as L1 speakers of Tagalog. The remaining 16 were sequential bilinguals. Participants provided informed consent and were compensated at a rate roughly equivalent to 12 USD/hr. On average, they took 25–35 minutes to complete the experiment. We excluded 10 participants before conducting

any statistical analyses based on their accuracy to comprehension questions and on their responses during debriefing. See section 3.4 for more information.

3.2. Materials

The experiment used a 2×2 design, crossing whether the sentence was grammatical or not (GRAM: G, UG) and whether the number feature of the distractor matched or mismatched the number feature of the reflexive (MATCH: M, MM). In the grammatical condition (G), the target (i.e., the antecedent) and the reflexive were both plural; in the ungrammatical condition (UG), the target was singular, while the reflexive was plural. In the match condition (M), the distractor and the reflexive were both plural; in the mismatch condition (MM), the distractor was singular, while the reflexive was plural. We created 24 experimental items, which were distributed evenly across four lists via Latin Square design and were randomized using 56 filler items. The filler items involved non-reflexive pronouns and relative clauses. They were different, but had comparable complexity as the experimental items. Table 1 provides a sample experimental item.

Each item was divided into presentation regions, which are indicated by the '/' in Table 1. The experimental items consisted of 13 presentation regions: the target always occupied the second region; the distractor, the fifth; and the reflexive, the seventh. The filler items consisted of 15–17 presentation regions.

Each item in experiment were followed by a comprehension question. In the grammatical condition, we probed their interpretation of the reflexive. These were critical comprehension questions and we report the results below. In the ungrammatical condition, we probed their comprehension of the other parts of the sentence that did not involve the reflexive-antecedent dependency. Sometimes, our question targeted the temporal or locative adverbial that occupy one of the presentation regions after the reflexive; sometimes, our question targeted the distractor by asking who the agent of the verb contained in the RC was. These were non-critical comprehension questions, and as will be described below, they were part of our exclusion criteria.

3.3. Method

We deployed the experiment online via IbexFarm, a web-based experimental presentation platform (Drummond 2016). Participants performed the experiment remotely on their own computers via a link distributed by Prolific Academic.

The items were presented in a self-paced phrase-by-phrase manner, as indicated in Table 1. Each phrase was center-aligned and presented non-cumulatively. Participants pressed the spacebar to move to the next region, and reading times were measured at each region.

During debriefing, we asked the participants the following: *Anong mga bagong hobbies at interes ang napick-up ninyo nang dahil sa pandemya?* 'What new hobbies and interests did you pick up due to the pandemic.' This free response question was included to ensure that those who participated were not computer bots and that they actually knew how to use the language.

3.4. Analysis

We excluded 10 participants before conducting any statistical analyses based on the following criteria. The first is if the participant had less than 75% accuracy in the non-critical comprehension questions, that is, questions that did not probe their interpretation of the reflexive. Six participants

were excluded for having accuracy in between 60%–73%. The second is if they had unnatural responses during debriefing. Two participants were excluded for having questionable proficiency as indicated by their responses during debriefing ('*Gusto ko naglaro sa internet ng League para my socialize because of the pandemya.*' and '*Bagong bobbies pagpipinta and maghurno*').

We fitted linear mixed effects regression models using the `lme4` package (Bates et al. 2015) in R (R Core Team 2018). We modeled the participants' log-transformed reading times (RT) as our dependent measure. The fixed effects included GRAM (G, UG) and MATCH (M, MM) and their interaction. GRAM and MATCH were sum-coded, such that G and M mapped to the negative coefficients. We also fitted a mixed effects logistic regression model using the package `lme4` in R. We modeled the participants' choices in response to the comprehension questions probing their interpretation of the reflexive (i.e., the questions after the grammatical conditions) as the dependent measure. When participants chose the distractor, we coded it as 0; when they chose the target, we coded it as 1. The fixed effect included MATCH, which was sum-coded such that M mapped to the negative coefficient. For all the models, we included the maximal random effects structure justified by the design that allowed the models to converge (Barr 2013). We used the package `lmerTest` (Kuznetsova et al. 2017) to calculate the *p*-values.

3.5. Results

We first present how the participants ultimately interpreted the reflexives, based on their responses to the comprehension questions that were presented after the grammatical sentences. After, we present how they incrementally interpreted the reflexives, based on their reading times at the critical region.

In short, we found the following: (i) the number features of the distractor had very little impact on how they ultimately interpreted the reflexive; and (ii) comprehenders were attending to them in real-time, such that sentences with feature-mismatching distractors took longer to read than sentences with feature-matching distractors.

3.5.1. Interpretation of the reflexive

Table 2 (left) provides the participants' rate of interpretation of the reflexive for grammatical sentences. Recall that our comprehension questions were designed so that the grammatical conditions probed their interpretation of the reflexive, while the ungrammatical conditions probed their comprehension of the other parts of the sentence. Thus, the values in Table 2 give us an index of how susceptible participants were to the number features of the distractor—put differently, an index of how consistent they were with Principle A of the Binding Theory—in how they ultimately interpreted the reflexive.

When the number feature of the distractor matched reflexive's (i.e., they were both plural), participants chose the target 70% of the time. When the number feature of the distractor mismatched the reflexive's (i.e., they were singular and plural, respectively), they chose the target 71% of the time. Table 2 (right) provides the summary of the mixed effects logistic regression. These results suggest that the number of the distractor had very little impact on how Tagalog comprehenders ultimately interpreted the reflexive (z = .03, p = .98). However, this does not give us any indication on whether they considered the distractor as the sentence was unfolding. We look at their reading times next to see the extent to which they attended to the distractor in real-time.

Table 1: Sample experimental item, manipulating GRAM (G, UG) and MATCH (M, MM). '/' indicates the different presentation regions.

GRAM	MATCH								
G	M	Pinupuri	/ ng mga dalaga	/ na	/ hinaharana	/ ng mga tambay	/ gabi-gabi	/ ang kanilang mga sarili	/ . . .
		praise	young women	LNK	serenade	loiterers	every night	themselves	. . .
		'The young women who the loiterers serenade every night praise themselves. . . '							
		Schematized: Verb / Target[PL] / Lnk / Verb / Distractor[PL] / Adverb / Reflexive[PL]							
G	MM	Pinupuri	/ ng mga dalaga	/ na	/ hinaharana	/ ng tambay	/ gabi-gabi	/ ang kanilang mga sarili	. . .
		praise	young women	LNK	serenade	loiterer	every night	themselves	. . .
		'The young women who the loiterer serenades every night praise themselves. . . '							
		Schematized: Verb / Target[PL] / Lnk / Verb / Distractor[SG] / Adverb / Reflexive[PL]							
UG	M	Pinupuri	/ ng dalaga	/ na	/ hinaharana	/ ng mga tambay	/ gabi-gabi	/ ang kanilang mga sarili	/ . . .
		praise	young women	LNK	serenade	loiterers	every night	themselves	. . .
		*The young woman who the loiterers serenade every night praise themselves. . .							
		Schematized: Verb / Target[SG] / Lnk / Verb / Distractor[PL] / Adverb / Reflexive[PL]							
UG	MM	Pinupuri	/ ng dalaga	/ na	/ hinaharana	/ ng tambay	/ gabi-gabi	/ ang kanilang mga sarili	/ . . .
		praise	young woman	LNK	serenade	loiterer	every night	themselves	. . .
		*The young woman who the loiterer serenades every night praise themselves. . . '							
		Schematized: Verb / Target[SG] / Lnk / Verb / Distractor[SG] / Adverb / Reflexive[PL]							
Comprehension question for G: Sino ang pinupuri? (Who is the one being praised?)									
Comprehension question for UG: Sino ang taga-harana? (Who is the one serenading?)									

Table 2: Participants' rates of interpretation of the reflexive for grammatical sentences only (left). Summary of the mixed effects logistic regression (right).

	Rate (%)	
MATCH	Target	Distractor
M	70	30
MM	71	29

	Est	*SE*	*z*	*p*
Intercept	1.2	.23	5.2	< .001
MATCH	.006	.23	.03	.98

N.B.: `glmer(Choice~Match + (1 + Match | Subject) + (1 + Match | Item))`

3.5.2. Reading times

In Figure 2, we plot their average region-by-region raw RTs. Table 3 provides a summary of the linear mixed effects models. We include the pre-critical region to provide a baseline for the effects that emerged in the critical region (i.e., at the reflexive).

At the pre-critical region, the fixed effects were not statistically significant. At the critical region, we found a main effect of GRAM: ungrammatical conditions took longer to read than grammatical conditions (t = 2.76, p = .01). We also found a main effect of MATCH: sentences with feature-mismatching distractors took longer to read than sentences with feature-matching distractors (t = 2.39, p = .02). The interaction was not statistically significant. These results suggest that even though the number features of the distractor had very little impact on how they ultimately interpreted the reflexive, comprehenders were attending to them in real-time.

Figure 2: Average raw reading times (*ms*) by GRAM and MATCH for the first 10 regions. In solid lines are the grammatical conditions; in dotted lines, the ungrammatical conditions. In gold are the match conditions; in blue, the mismatch conditions. Standard errors of the mean are also provided.

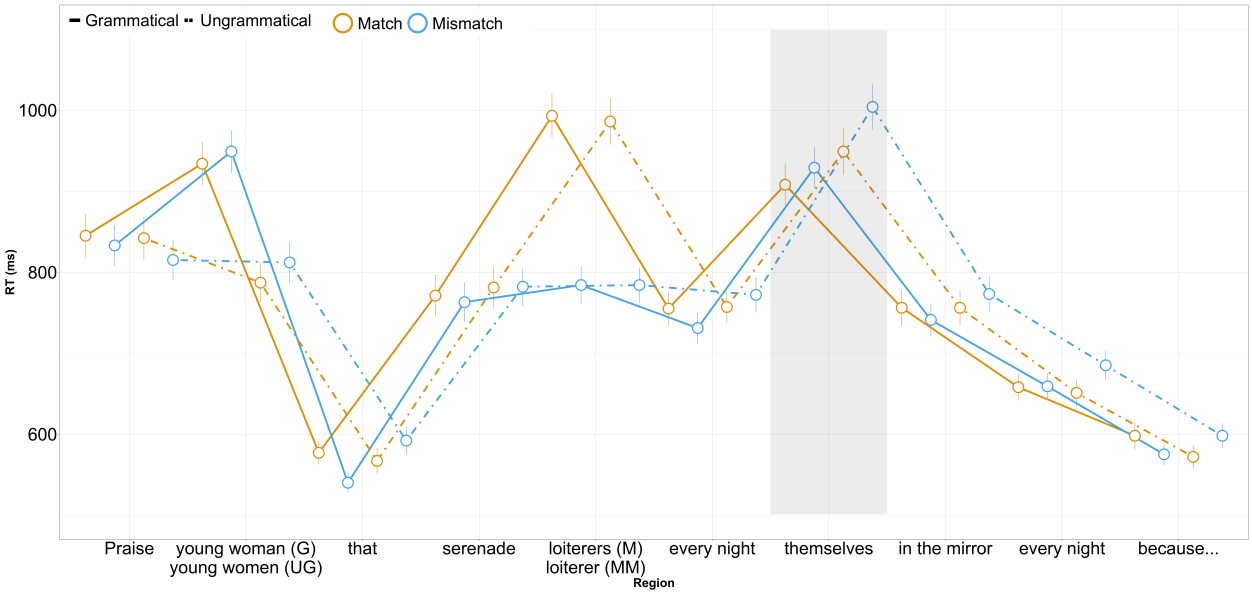


Table 3: Descriptive statistics at the pre-critical and critical regions (left). Summary of the linear mixed effects regression model at the pre-critical and critical regions (right).

GRAM	MATCH	*M*	*SE*
At the pre-critical region			
G	M	755	21
G	MM	731	19
UG	M	757	20
UG	MM	772	22
At the critical region			
G	M	908	25
G	MM	929	25
UG	M	949	29
UG	MM	1004	28

	Est	*SE*	*t*	*p*
At the pre-critical region				
Intercept	6.52	.037	174.87	< .001
GRAM	.028	.019	1.56	.12
MATCH	.007	.019	-.41	.68
Interaction	.037	.036	1.03	.31
At the critical region				
Intercept	6.73	.051	131.09	< .001
GRAM	.051	.019	2.76	.01
MATCH	.045	.019	2.39	.02
Interaction	.037	.037	.98	.32

N.B.: `lmer(log(RT)~Gram*Match + (1 | Subject) + (1 | Item))`

4. General discussion and conclusion

The goal of the present study was to investigate the interference profile of reflexive processing in Tagalog after controlling for recent activation of the target as a potential confound. We found the following. First, the number feature of the distractor had very little impact on how participants ultimately interpreted the reflexives. Second, the number feature of the distractor did affect how comprehenders read the sentences. After controlling for the reactivation of the target as a potential confound, reflexive processing in Tagalog is susceptible to retrieval interference: those with number-mismatching distractors took longer to read than those with number-matching distractors. In short, even though their final interpretation of the reflexive was consistent with Principle A, their incremental interpretation showed facilitatory interference in both grammatical and ungrammatical conditions. This across-the-board facilitation is a surprising result for both a simple model of retrieval and a syntactically guided model.

Before we discuss the implications of our results, we first address a potential concern regarding the localization of our effects given the methodology that we used. We then zoom out and discuss what our results have to say about models of cue-based retrieval. After, we outline some future directions.

4.1. Localization of interference effects

A discussion on how localized our effects are is in order. It is well-known in the psycholinguistics literature that self-paced reading studies typically have poor temporal resolution (Boyce et al. 2020). Effects typically appear not in the region where they are expected, but rather they "spill over" a region or two downstream. To make spillover effects more concrete, let's take Wagers et al. 2009 as an example. They investigated the interference profile in English subject-verb dependencies by comparing sentences like in (13).

(13) The key to the {cell/cells} unsurprisingly {was/were} rusty from many years of disuse.

To the extent to that there will be interference effects, the earliest time these effects should be observable is at the point where the matrix verb *was/were* is encountered (i.e., when retrieval of the target is launched). Their results indicated that their effects consistently spilled over and emerged one region after the retrieval site.

In the present study, we argue that our reading time results show localized effects, appearing at the region where retrieval is launched (i.e., at the reflexive). Here are some reasons why we think that this is the case for our study. First, at any given region, the mean reading times of our participants were slower than the mean reading times reported in existing self-paced reading studies. For context, consider experiment 4 in Wagers et al. 2009. The slowest mean reading time reported there is a little less than 400 ms. By contrast, in the present study, the fastest mean reading time is about 500 ms (at region 3). Based on our experience with running self-paced reading experiments, there seems to be a trade-off between how much time participants spend at any given region and how localized the effects are. We do not know of any studies, however, that explicitly test this purported trade-off. Second, our effects are localized throughout—even in our non-critical regions. For example, consider the schematization of our items in (14). '/' represent regions of presentation. We repeat Figure 2 as Figure 3 below for ease of reference.

(14)a. Verb / mga Noun / linker ... *Grammatical (*G*)*
b. Verb / Noun / linker ... *Ungrammatical (*UG*)*

In region 2, when participants encountered the target, our grammatical conditions had an extra morpheme *mga* to make the target match the plural reflexive *ang kanilang mga sarili* that they will eventually encounter downstream. As seen in Figure 3, the solid lines, which correspond to the grammatical conditions, diverge from the dotted lines, which correspond to the ungrammatical conditions. This divergence is a length effect and crucially, it appeared at the same region when the plural morpheme *mga* was introduced in the grammatical conditions.

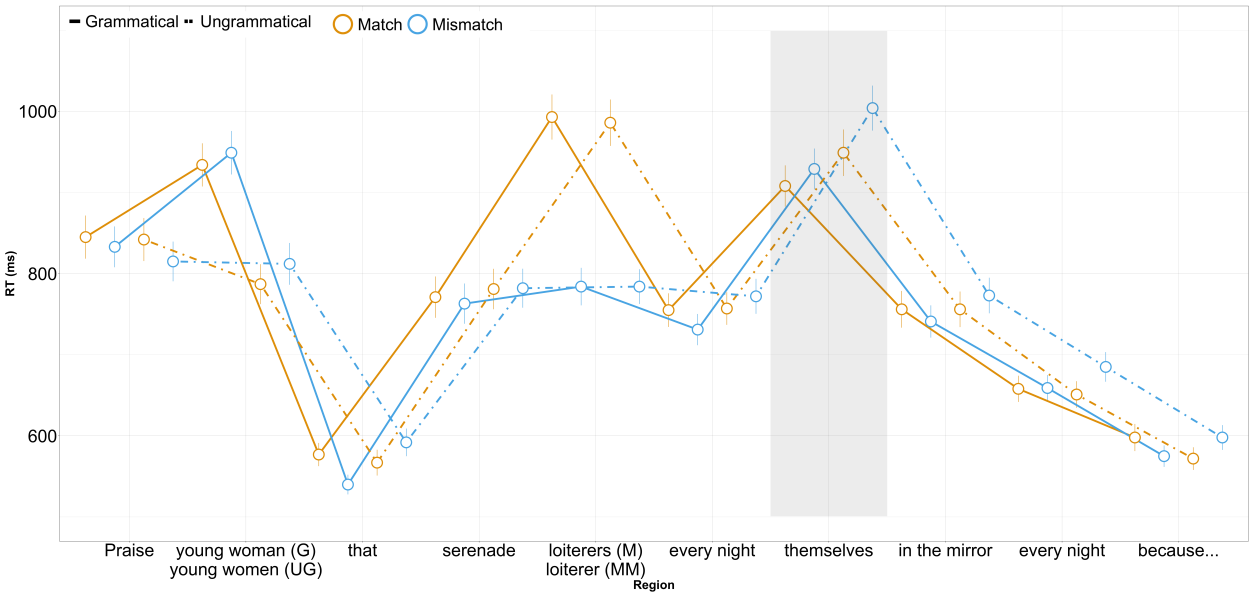
We also see another localized length effect in region 5. When participants encountered the distractor, our match conditions had an extra morpheme *mga* to make the distractor match the plural reflexive *ang kanilang mga sarili* that they will eventually encounter downstream. As seen in Figure 3, we see that the gold lines, which correspond to the match conditions, diverge from the blue lines, which correspond to the mismatch conditions. This appeared at the same region when the plural morpheme *mga* was introduced in the match conditions.

4.2. Implications for models of cue-based retrieval

As mentioned above, the facilitatory interference that we observed at the critical region—in both the grammatical and ungrammatical conditions—is a surprising result for any of the models of retrieval discussed in section 2. The results of the present study are inconsistent with the predictions of a simple model of retrieval and a syntactically guided model.

The direction of the interference effect in the present study cannot be accounted for by a simple model of retrieval, like LV05. A simple model of retrieval predicts the interference profile of reflexive dependencies to pattern like the interference profile of subject-verb dependencies, as schematized in Figure 1: inhibitory interference in the grammatical condition and facilitatory interference in the ungrammatical condition.

Figure 3: Average raw reading times (*ms*) by GRAM and MATCH for the first 10 regions. In solid lines are the grammatical conditions; in dotted lines, the ungrammatical conditions. In gold are the match conditions; in blue, the mismatch conditions. Standard errors of the mean are also provided.



The presence of the interference effect in the present study cannot be accounted for by a syntactically guided model of retrieval, like the one proposed by Dillon et al. (2013). In its strongest form, this model predicts no interference effects because elements that are not structurally licensed are not candidates for retrieval.

Our results can be accounted for by a model of retrieval that scales interference effects vis-à-vis the difference in activation between the target and distractor. For example, Jäger et al. (2015) proposed an extension of LV05 by including a prominence principle that takes into account the distractors' relative activation. One of the key predictions of this extension is that the behavioral footprint of similarity-based interference effects in target-match conditions (i.e., when the features of the target fully matches the retrieval cues) can be substantially altered for distractors that are highly activated. According to Jäger et al. 2015, there are a number of factors that can influence the distractor's base-level of activation: grammatical role, discourse topicality, thematic role, to name a few. For the other factors that could potentially affect base-level activation, see their section 4.1 (p. 17). For our study, we conjecture that when retrieval was launched upon encountering the reflexive, the distractor was more activated than the target because of Tagalog's word order.

Consider the following scenario. In (15), we reschematize our experimental items. Circled numerals represent time points where something relevant happens.

(15) Verb ① Target [$_{RC}$ Verb ② Distractor Adv] ③ Reflexive ...

At time point ①, the target and the matrix verb are thematically integrated. At time point ②, the distractor and the verb inside the verb are thematically integrated. Before the retrieval operation for antecedent resolution at time point ③, the distractor could be enjoying a higher base-level activation than the target because it was recently thematically integrated with the RC-verb. In our attempt to control for the recent activation of the target by leveraging Tagalog word order, we

had inadvertently created a configuration that might have made the distractor more accessible than the target via recent activation.

Why should distractor prominence create facilitatory interference in target match configurations? Although the target is usually assumed to robustly outcompete any distractors in grammatical contexts (Wagers et al. 2009), this may not be the case when the distractor is highly prominent. Engelmann et al. (2019) show that when a distractor is highly prominent relative to a target, it is possible to see facilitatory interference due to statistical facilitation in target-match configurations. It is also possible that the increased prominence of the distractor leads it to be retrieved at a relatively high rate even in target match sentences.

Alternatively, Tagalog comprehenders could be more prone to interference because of how similar the target and the distractor were in the context of our experiment. Both were full NPs that were genitive-marked (i.e, they had *ng* before the lexical noun) and were subjects/agents of their respective clauses.

4.3. Future directions

In this paper, we conjectured that when retrieval was launched upon encountering the reflexive, the distractor was more activated than the target because of Tagalog's word order. The word order configuration of our items might have made the distractor more accessible via recent activation. In order to investigate whether this hypothesis is tenable or not, we can run an experiment that leverages the availability of head-initial and head-final RCs in the language. Consider the schematizations in (16). Circled numerals represent time points where something relevant happens.

(16)a. Verb ① **Target** [$_{RC}$ Verb ② Distractor] ③ Reflexive ...
b. Verb [$_{RC}$ Verb ① Distractor] ② **Target** ③ Reflexive ...
c. Target [$_{RC}$ Verb ① Distractor] ② Verb ③ Reflexive ...

The configuration in (16a) is the current configuration. It involves a head-initial RC, where the target is the head noun (in bold) and the distractor is contained inside a relative clause modifying the head noun. By hypothesis, the distractor is more activated than the target because it was thematically integrated with the RC-verb at time point ②. This is before the retrieval operation for antecedent resolution at time point ③. If our hypothesis is tenable, we should replicate the current findings in a configuration like (16a).

The configuration in (16b) involves a head-final RC. In a sense, this configuration has parallels with the items found in many of the previous studies in reflexive processing in English. Compare (16b) and (16c). In both, the target and the matrix verb are thematically integrated at time point ②, immediately before the retrieval for antecedent resolution at time point ③. If our hypothesis is tenable, we expect no/weak interference effect when the RC is head-final.

To investigate whether Tagalog comprehenders were more prone to interference because the target and the distractor were too similar, we can add another factor to the current design and manipulate whether the voice of the RC-verb is agent voice (AV) or patient voice (PV), as schematized in (17).

(17)a. Verb.PV GEN-Target$_{[SG/PL]}$ [$_{RC}$ Verb.PV GEN-Distractor$_{[SG/PL]}$] Reflexive$_{[PL]}$
b. Verb.PV GEN-Target$_{[SG/PL]}$ [$_{RC}$ Verb.AV DAT-Distractor$_{[SG/PL]}$] Reflexive$_{[PL]}$

This additional manipulation has two welcome consequences. When the RC-verb has AV morphology, as in (17b), the distractor and the target no longer have the same case marking. The target remains genitive-marked; the distractor is now dative-marked. Additionally, they no longer have the same thematic relations in their respective clauses. The target remains the agent of the matrix verb; the distractor is now the patient of the RC-verb. If the interference effects in the present study were due to how similar the target and the distractor were in their case marking and thematic role, then we expect to see less interference when the verb has AV, as in (17b), than when the verb has PV, as in (17a).

References

Barr, Dale J. 2013. Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology* 4:328. URL https://doi.org/10.3389/fpsyg.2013.00328.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67:1–48. URL https://doi.org/10.18637/jss.v067.i01.

Bondoc, Ivan Paul, Kamil Deen, Elsie Marie Or, and Ma. Clarisse Hemedes. 2019. Reflexives in adult and child Tagalog. In *Proceedings of the 43rd Boston University Conference on Language Development*, ed. Megan M. Brown and Brady Dailey, 82–93. Somerville, MA: Cascadilla Press.

Boyce, Veronica, Richard Futrell, and Roger P. Levy. 2020. Maze made easy: Better and easier measurement of incremental processing difficulty. *Journal of Memory and Language* 111:104082. URL https://doi.org/10.1016/j.jml.2019.104082.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht, The Netherlands: Foris.

Dillon, Brian. 2011. Structured access in sentence comprehension. Doctoral Dissertation, University of Maryland, College Park, MD.

Dillon, Brian, Alan Mishler, Shayne Sloggett, and Colin Phillips. 2013. Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language* 69:85–103. URL https://doi.org/10.1016/j.jml.2013.04.003.

Drummond, Alex. 2016. Ibex farm: Internet Based EXperiments [Computer program]. URL http://spellout.net/ibexfarm/.

Engelmann, Felix, Lena A Jäger, and Shravan Vasishth. 2019. The effect of prominence and cue association on retrieval processes: A computational account. *Cognitive Science* 43:e12800. URL https://doi.org/10.1111/cogs.12800.

Jäger, Lena A, Felix Engelmann, and Shravan Vasishth. 2015. Retrieval interference in reflexive processing: experimental evidence from Mandarin, and computational modeling. *Frontiers in Psychology* 6:617. URL https://doi.org/10.3389/fpsyg.2015.00617.

Jäger, Lena A, Felix Engelmann, and Shravan Vasishth. 2017. Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language* 94:313–39. URL https://doi.org/10.1016/j.jml.2017.01.004.

Jäger, Lena A, Daniela Mertzen, Julie A. Van Dyke, and Shravan Vasishth. 2020. Interference patterns in subject-verb agreement and reflexives revisited: A large-sample study. *Journal of Memory and Language* 111:104063. URL https://doi.org/10.1016/j.jml.2019.104063.

King, Joseph, Caroline Andrews, and Matthew Wagers. 2012. Do reflexives always find grammatical antecedents for themselves? Poster presented at the 25th Annual CUNY Human Sentence Processing Conference.

Kush, Dave, and Colin Phillips. 2014. Local anaphor licensing in an SOV language: implications for retrieval strategies. *Frontiers in Psychology* 5:1252. URL https://doi.org/10.3389/fpsyg.2014.01252.

Kush, David. 2013. Respecting relations: Memory access and antecedent retrieval in incremental sentence processing. Doctoral Dissertation, University of Maryland, College Park, MD.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82:1–26.

Lewis, Richard, and Shravan Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29:375–419. URL https://doi.org/10.1207/s15516709cog0000_25.

Lewis, Richard, Shravan Vasishth, and Julie A. Van Dyke. 2006. Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences* 10:447–54. URL https://doi.org/10.1016/j.tics.2006.08.007.

Martin, Andrea E., and Brian McElree. 2009. Memory operations that support language comprehension: Evidence from verb-phrase ellipsis. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35:1231–9. URL https://doi.org/10.1037/a0016271.

Nicenboim, Bruno, and Shravan Vasishth. 2018. Models of retrieval in sentence comprehension: A computational evaluation using bayesian hierarchical modeling. *Journal of Memory and Language* 99:1–34. URL https://doi.org/10.1016/j.jml.2017.08.004.

Parker, Dan, Michael Shvartsman, and Julie A. Van Dyke. 2017. The cue-based retrieval theory of sentence comprehension: New findings and new challenges. In *Language Processing and Disorders*, ed. Linda Escobar, Vicenç Torrens, and Teresa Parodi, 121–44. Newcastle: Cambridge Scholars Publishing.

R Core Team. 2018. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Richards, Norvin. 2013. Tagalog anaphora. In *Diagnosing syntax*, 412–33. Oxford: Oxford University Press.

Sturt, Patrick. 2003. The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language* 48:542–562. URL https://doi.org/10.1016/S0749-596X(02)00536-3.

Van Dyke, Julie A. 2007. Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33:407–30. URL https://doi.org/10.1037/0278-7393.33.2.407.

Van Dyke, Julie A., and Richard Lewis. 2003. Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language* 49:285–316. URL https://doi.org/10.1016/S0749-596X(03)00081-0.

Vasishth, Shravan, Sven Brüssow, Richard L. Lewis, and Heiner Drenhaus. 2008. Processing

polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science* 32:685–712. URL https://doi.org/10.1080/03640210802066865.

Vasishth, Shravan, Bruno Nicenboim, Felix Engelmann, and Frank Burchert. 2019. Computational models of retrieval processes in sentence processing. *Trends in Cognitive Sciences* 23:968–982. URL https://doi.org/10.1016/j.tics.2019.09.003.

Wagers, Matthew, Ellen Lau, and Colin Phillips. 2009. Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language* 61:206–37. URL https://doi.org/10.1016/j.jml.2009.04.002.

Xiang, Ming, Brian Dillon, and Colin Phillips. 2009. Illusory licensing effects across dependency types: ERP evidence. *Brain & Language* 108:40–55. URL https://doi.org/10.1016/j.bandl.2008.10.002.