2022. *Proc Ling Soc Amer* 7(1). 5210. https://doi.org/10.3765/plsa.v7i1.5210.

Processing profile for quantifiers in verb phrase ellipsis: Evidence for grammatical economy

Wesley Orth & Masaya Yoshida*

Abstract. Quantifier Raising leaves no overt marking to indicate movement has occurred, making the task of identifying when raising has occurred extremely difficult for the parser. Beyond this challenge, evidence from interpretation and judgement studies suggests that raising causes difficulty in processing. These two aspects taken together have led some to suggest that human sentence processor employs a strategy in which the construction of raised structures is avoided, commonly called processing scope economy. This contrasts to the traditional notion of grammatical scope economy, where Quantifier Raising is restricted in the grammar. In this paper we discuss the properties of these two theories. We conclude that the two approaches make different predictions about when raising should occur online, with processing scope economy predicting that the parser avoids raising whenever possible and grammatical scope economy predicting that the parser raises regularly and sometimes produces ungrammatical structures in the process. We then present an experiment which examines complex scope structures in verb phrase ellipsis to observe when penalties related to Quantifier Raising are observed online. We find that penalties appear in configurations where Quantifier Raising would be ungrammatical under grammatical scope economy, suggesting the parser attempts Quantifier Raising in these configurations. This evidence indicates that the parser's behavior matches the predictions of grammatical scope economy rather than those of processing scope economy.

Keywords. psycholinguistics; scope; quantifier raising; ellipsis; maze task

1. Introduction. An outstanding question in sentence processing concerns when the parser builds structures reflecting scope taking operations such as Quantifier Raising (QR) (May 1977). There are three main proposals for answering this question. The first proposal is that the parser avoids performing QR whenever possible, as performing QR is thought to be a costly operation (Anderson 2004, Tanaka 2015, Wurmbrand 2018). This is based on the observation that readers generally prefer surface scope and scopally complex sentences often have relatively low acceptability. This avoidance strategy has been operationalized as a parsing constraint, processing scope economy (PSE), which contends that the parser avoids raising whenever possible (Anderson 2004). Under the second proposal, the availability of complex scope structures in early processing is taken as a sign that the parser performs QR regularly to accommodate underlying grammatical principles like grammatical scope economy (GSE) (Brendel 2019, Fox 1995, Fox 2000, Syrett 2015a, Syrett 2015b). While others still suggest that the parser waits until near the end of the sentence to compute possible scope structures (Bott & Schlotterbeck, 2015). In this paper we focus on the two proposals related to GSE and PSE and report the results of an experiment testing the predictions of these alternatives in sentences with verb phrase ellipsis (VPE).

* This material is based upon work supported by the National Science Foundation under Grant No. 2116989. Thanks to the members of the Syntax Semantics and Sentence Processing Lab at Northwestern University for their feedback on both the manuscript and early stages of this work. Authors: Wesley Orth, Northwestern University (worth@u.northwestern.edu) & Masaya Yoshida, Northwestern University (m-yoshida@northwestern.edu).

In the following section we offer a brief overview of QR and its properties. In section 3, we introduce GSE in greater detail, discuss the issues it raises for parsing and offer a model for how the parser could deal with GSE (Cecchetto 2004, Fox 1995, Fox 2000, Syrett 2015b). We conclude that determining whether QR results in a scope informative structure or not is only possible if the post-QR structure is built. Thus, the parser must frequently perform QR to check if the post-QR structure is scopally informative. Because the parser must perform QR to check for scope informativity the parser will occasionally produce structures that are not scope informative which must be reanalyzed. In section 4 we discuss PSE in greater detail. This approach is modelled after constraints opt for simpler structures and states that the parser builds simpler scope structures whenever possible (Anderson 2004, Frazier & Fodor 1978, Gibson 1998, Wurmbrand 2018). As such GSE and PSE make different predictions about when raising occurs in the processing of a sentence and for what reason raising is performed. Under PSE, QR is avoided whenever possible, while under GSE the parser should perform QR more frequently and even in situations where the result is ultimately ungrammatical. In tsection 5 we present the results of a maze task reading study which is designed to test the predictions of these economy conditions (Boyce Futrell & Levy 2020, Forster Guerrera & Elliot 2009, Gulordava Bojanowski Grave Linzen & Baroni 2018; Witzel Witzel & Forster 2012). We find reading time penalities at positions where QR is prohibited. This suggests that the parser has performed QR and faced a penalty for producing an ungrammatical structure rather than simply avoiding QR. These results strongly align with the predictions of GSE and suggest that the parser frequently attempts QR online.

2. Quantifier raising and scope economy. While there are other proposed theories of scope-taking, this paper is focuses on Quantifier Raising (May 1977, May 1985). QR is a syntactic operation which covertly raises a quantifier to a higher position in the sentence, originally proposed as an adjoin to S. This version of QR allows for the construction of structures with different scope interpretations relative to the surface structure. The example sentence in (1a) holds two possible interpretations. The surface scope interpretation is paraphrased as in (1b), while the QR derived interpretation is paraphrased as in (1c).

(1) a. A teacher saw every student
  b. For a teacher that teacher saw every student
  c. For each student there is some teacher who saw that student

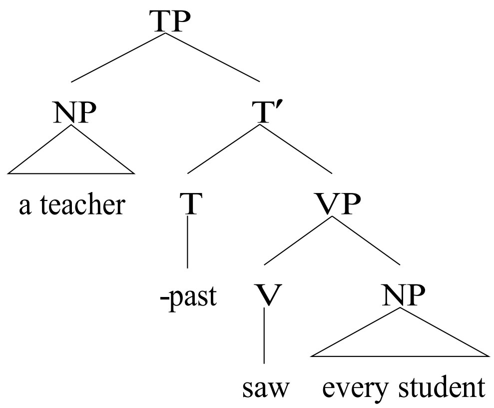
The critical difference between the two representations of the sentence (1a) can be seen in the simplified tree structures (2a) and (2b), which correspond to the surface and inverse scope interpretations (1b) and (1c).[1] In contrast to (2a), the quantifying expression *every student* has been attached to a higher position in (2b). This allows the quantifiers to combine with the verb in a different order and results in a different interpretation for the same surface string.

While we have applied a particular formalism for the trees in (2), there is considerable debate as to the specific syntactic properties of QR. In its original formulation, QR was defined by the simple description "adjoin Q (to S)," where Q stands in for the quantifying expression and S can be understood as the root node of a syntactic structure (May 1977, May 1985). This

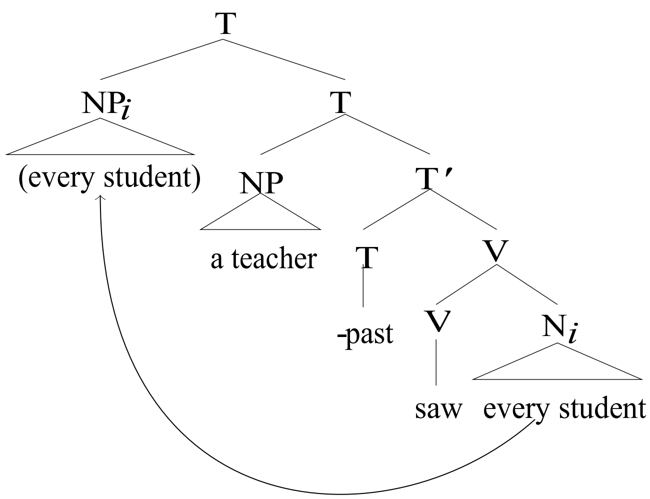
[1] For the purposes of this discussion, we adopt a tree notion consistent with X-bar theory style tree structures which appear in a range of syntax related textbooks (e.g., Adger 2003, Carnie 2021, Koopman Sportiche & Stabler 2014). However, we do not wish to commit ourselves to the particulars of any specific structural formalism. Our approach rests on the critical features of hierarchical structure and movement, along with the generally accepted ordering of regularly occurring elements in the syntactic spine, which can be consistent with a wide range of analyses.

movement was additionally assumed to leave behind a trace at the original position which would be bound by the moved quantifying expression.

(2) a.



b.



However, this formulation of QR was found to have two primary issues. The first is that the operation was relatively unbounded and the second is that there was no limit preventing redundant applications of QR. To the first point, the unboundedness of QR predicted a larger number of available scope interpretations than were thought to be attested, including cases like raising quantifiers out of relative clauses. Thus, the question is whether QR obeys typical restrictions on movement. In some approaches QR was analyzed as being restricted by locality conditions on movement such as the Empty Category Principle or the successive cyclic movement condition (Cecchetto 2004, Chomsky 1986, Chomsky 1977, Hornstein 1994, Hornstein 1995, Huang, 1995, Syrett 2015a, Syrett 2015b). Some proposals opted to limit QR further beyond conditions on other types of movement, including a condition on clause boundedness such that QR cannot occur out of a tensed clause (Farkas 1981, Hornstein 1995, Lasnik & Stowell 1991, Lasnik 1993). Some studied have proposed that quantifiers move to predefined structural positions where conditions on raising are product of the differing properties of movement to and from these positions (Beghelli & Stowell 1997). More recent studies have pointed out QR is relatively unconstrained with respect to locality conditions on movement and the restrictions we see are related to independent factors about the structures in question (Barker 2021, Syrett 2015a, Syrett 2015b). With this in mind, we set aside the question of how QR relates to other movement operations in syntax, and focus our investigation the second issue of limiting redundant QR.

3. Grammatical economy. QR formulated as in (2) predicts that QR can be applied any number of times. The formulation includes no condition on when how QR can be applied. Under a movement theory of QR, there isn't anything that prevents the repeated QR in the same structure, raising the quantifier to arbitrarily high positions. To prune the grammar of this surfeit of possible grammatical scope structures, it was proposed that QR is subject to a condition called scope economy, which we reference as grammatical scope economy (GSE) in this paper for the sake of clarity (Fox 1995, Fox 2000). Under GSE a post-QR structure is not licensed unless it results in a scope informative structure. A structure is scope informative when it results in new interpretation for the sentence relative to the previous steps in the derivation. Under GSE, examples such as (1a) go from being many ways ambiguous to being constrained to the single structure which is of the form of (2b). This rule also ensures that structures such as those in for the examples in (3) are structurally unambiguous, as no application of QR would be scope informative.

(3) a. Every climber visited every mountain
    b. A climber visited a mountain
    c. Jane saw a mountain

Additional work on the formulation of GSE has proposed combining some of the movement constraints with economy conditions. One approach suggests that QR is a short movement and each component movement in a long distance QR operation must be independently motivated by GSE, except for QR applications necessary to resolve type mismatches in logical form (Reinhart 2006, Syrett 2015b, Tanaka 2015). However, in each proposed variant of scope economy the comparison of structure's interpretations through scope-informativity is the critical computation that underlies the constraint. While the impact that GSE has on the grammar is straightforward, the impact such a condition would have on the parser is more complex. As the condition is defined as a comparison of the interpretations read off pre-QR and post-QR structures, both structures must be available to the parser for it to check if a QR application is scope informative. In other words, the fact that GSE relies on a comparison means that the parser must perform QR to evaluate that instance of QR's grammaticality with respect to the GSE condition.

This poses a challenge for parsing. If it is the case that the parser has no indication of whether QR is licensed before performing QR, it becomes impossible to exhaustively pre-filter structures which would be ungrammatical with respect to GSE. When the parser encounters a quantifier, the parser cannot know whether QR is possible because the possibility of QR depends on whether the post-QR structure is scopally informative or not. Here the parser is placed in a difficult situation. If the parser can build the structures allowed in the grammar, it should be able to apply QR in the presence of quantifiers. However, with no ability to pre-calculate scope informativity the parser will produce structures that are not scope informative and ungrammatical.

This is a serious problem for the idea that the human sentence parser does not produce ungrammatical structures, a regularly proposed constraint on parsing (Crain & Fodor 1985, Fodor 1978, Frazier Ackerman Baumann Potter & Yoshida 2015, Kazanina Lau Lieberman Yoshida & Phillips 2007, Miller & Chomsky 1963, Phillips 1996, Phillips Wagers & Lau 2011, Steedman 1996, Stowe 1992). If we hold both that the parser has no ability to know if a structure is scope informative without performing QR and that ungrammatical structures are produced, what remains is a parser that only performs QR when the output will be scope informative, despite the explicit condition that the parser cannot possibly know this before performing QR. Thus, the parser becomes an oracle which only produces valid structures despite there being no way to compute validity before performing QR. However, there are many findings which suggest there is a place for building potentially ungrammatical structures in human sentence processing (Clifton & Frazier 1989, Fodor 1978, Fodor 1981, Freedman & Forster 1985, Phillips 1996, Phillips 2006). If we allow that the parser can build potentially ungrammatical structures, we can imagine a several strategies that the parser could adopt to ensure all grammatical structures are buildable while minimizing QR that is not scope informative and the associated processing penalties.

Let us imagine a scenario where the parser needs to perform QR. Starting from the end of the procedure, the parser must evaluate if the structure it has generated by QR is scope informative. At this point structures which are scope informative and do not violate GSE remain, and parsing continues. On the other hand, structures which are not scope informative and violate GSE must be reverted to their pre-QR configuration. This means that any failed application of QR will incur some penalty as we assume the reversion to a previous structure is not cost-free (Fodor & Ferreira 2013, Frazier 1979, Frazier & Clifton 1998, Schneider & Phillips 200, Sturt 1996).

Thus, in defining a parsing strategy to exhaustively produce the scope structures that are permitted by the grammar, the critical trade-offs appear to be incrementality and the rate of production of ungrammatical QR structures. The parser must strike a balance between how quickly it produces good QR structures and the number times it creates a QR structure that is not scope-informative. A parser which makes few mistakes but waits to perform QR does not conform with the idea that processing is largely incremental, while a parser that performs QR quickly would build structures that aren't scope informative that could be avoided with a more sophisticated strategy. An ideal solution would minimize penalties while retaining perfect generative ability with respect to the grammar, alongside a reasonable degree of incrementality to be in line with existing observations in sentence processing. As this optimization problem contains multiple parameters, we can imagine a system that favors these components to various degrees.

Beginning with the extreme approaches we could imagine a parser that strongly favors incrementality. Under such an approach the parser would immediately perform QR to the maximal position possible when encountering any quantifier. This would ensure the parser doesn't miss any potentially valid scope structures. Processing (1a) with this strongly incremental system the parser would immediately raise the quantifier *a teacher* before any other words are encountered. This QR is not scope informative based on the preceding material and must be reverted. When the parser arrives at the second quantifier *every student*, it would attempt to raise this element, generating a scope informative inverse scope structure. It would then again attempt to raise *a teacher*, as the structure is different now and the scope informativity of this movement needs to be tested, and once again will need to revert this QR application. This approach is highly reckless, as QR is attempted three times and reverted twice. It is important to note that the initial raising of *a* teacher will never be scope-informative, as the absence of any other material means there is no possible change in interpretation. The examples in (3) present a similar problem as QR applications in sentences like these never be scope informative. Thus, it appears the parser could avoid many costs by relaxing incrementality – a point we will return to later.

On the other extreme we could imagine an approach that completely forgoes incrementality to minimize the creation of ungrammatical structures. Under this sort of approach, the parser would wait until the input is exhausted and perform QR on quantifiers proceeding right to left through the string. Under this approach the parser would initially parse (1a) in its surface scope structure and, upon reaching the end of the input begin raising backwards through the input. This would first produce a scope informative inverse scope structure with the raising of the lower quantifier *every student*. Much like in the strongly incremental case, this minimally incremental parser would next need to raise *a teacher* to test this movements scope informativity. As was the case with the strongly incremental parser, this QR is not scope informative as it generates the same interpretation as the surface scope structure initially constructed. The parser would need to revert this movement, yielding a single ungrammatical raising and reversion for this strategy. Thus, this parser has an advantage over the strongly incremental parser in that it needs to perform QR one less time and only needs to revert a single instance of QR. However, this parser's total lack of incrementality is at odds with findings that quantifiers pose processing costs before the wrap-up positions in online processing, as well as the broader idea that human sentence processing is incremental (Anderson 2004, Wurmbrand 2018).[2]

We can now ask if there is some optimal midpoint between these extremes which features more incrementality than the non-incremental parser while performing fewer QR applications and reversions than the strongly incremental parser. Between these two extreme positions there is considerable space for the formulation of different parsers. What follows is a proposal for one such parser and a brief discussion of what other possible parsers exist between the extremes.

We propose that, rather than performing QR immediately upon encountering a quantifier or waiting until the end of the sentence, QR is a triggered operation with a single condition. The trigger itself is the detection of a scope sensitive element in the input. An element is scope sensitive if it acts as an operator or a variable. This means that scope sensitive elements include but are not limited to pronouns, polarity items, and quantifiers themselves. When such a scope sensitive element is encountered parser then must determine if a quantifier exists in memory, as a condition for performing QR. If yes, the parser performs QR in a right to left fashion as described in the extreme parsers. If there is no quantifier in memory, the parser does not perform QR and continues to process new input as normal. Note that existence in memory makes the parser's QR application quite powerful, as it would mean the parser may raise quantifiers in memory that do not exist in a c-command relationship with the input scope sensitive item. While this is on example of possible modification, we believe that the recent work indicating that QR is relatively unconstrained in the grammar and the recent work showing quantifiers interfere with NPI licensing from within relative clauses suggests that such an unconstrainted condition is appropriate (Barker 2021, Orth Slogget & Yoshida 2020, Syrett 2015a, Syrett 2015b).

Let us walk through some of the examples to illustrate how this triggered version of QR combines the best features of the more extreme parsers. For example (1a) we would predict that the parser would not perform raising until the point of the second quantifier *every student*. This is because at the point of *a teacher*, a scope sensitive element is present, but no quantifier yet exists in memory, thus allowing the parser to avoid performing QR here. At the point of *every student* the parser would perform QR for *every student*, moving from the scope structure in (2a) to the structure in (2b). After this the parser would perform raising on *a teacher*, yielding a single instance of QR that is not scope informative and must be reverted. With this procedure, the parser achieves the same QR performance profile of the non-incremental parser, two QR applications and a single revision, but with a higher degree of incrementality.

[2] It should be noted that this does not mean that there are no wrap-up effects related to scope, as sentences with complex scope structures in German incur processing penalties in wrap-up regions (Bott & Schlotterbeck 2015).

If we look at the example (3c) this parser continues to have advantages over the extreme alternatives. For the strongly incremental parser, this sentence would entail one instance of QR and one reverted structure at the point when *a mountain* is encountered, as we might expect. The non-incremental version in this case has the same profile, as the parser would attempt QR for the quantifier *a mountain* upon reaching the end of the sentence and revert the movement as the result is not scope informative. On the other hand, the triggered parser will never attempt QR for this example. While the trigger occurs at the point of *a mountain*, the condition that a quantifier exists in memory is not met, and thus no QR is attempted. In this way this triggered parser possesses a unique advantage for sentences with only a single quantifier.

That isn't to say this triggered parser performs better than the extreme versions in every case. If we explore the examples in (3a) and (3b) the three parsers perform similarly poorly. The strongly incremental parser will attempt QR three times and revert all three cases, while both the non-incremental and triggered version will attempt QR twice and revert both instances. This could desirable if empirical data suggests these sentences are more difficult than their counterparts with non-commutative quantifiers. However, it is also possible that the condition on the triggered parser could be modified such that there must be a quantifier present in memory that is not commutative with the input scope sensitive element, if it is the case that sentences such as (3a) and (3b) are not more difficult to process than their scope ambiguous counterparts. As mentioned previously, a similar constraint could be applied to structural factors like c-command, and the identification of such constraints would be the central task of developing this parsing theory.

4. Processing scope economy. There are still alternative approaches to managing cases of unnecessary QR. Anderson 2004 proposed that GSE could be replaced with a notion of processing scope economy (PSE). This is presented as a complexity-based theory of QR in processing and is based on the observation that sentences with more complex scope structures tend to be less acceptable than their simpler counterpart. Thus, PSE was defined as a parsing preference to construct the scope configuration with the simplest syntactic representation, and that the computation of more complex configurations is possible, if supported by pragmatic factors, but incurs a processing cost (Anderson 2004, Pritchett & Whitman 1995).

This alternative faces a few issues. The first being that this definition doesn't prevent the overgeneration issues in the grammar which GSE was proposed to solve. While the parser won't build structures with arbitrary raising online, PSE does not prune structures with arbitrary raising from the grammar. As such a grammar for a language with PSE would admit QR for the sentences in (3), which neither serve to bind variables or produce alternative scope interpretations.

The second issue is that this definition is somewhat squishy. Under PSE, the parser prefers structures with less complexity. As this condition is stated as a comparative we are faced with the same issue as in GSE. The complexity of two structures cannot be directly compared without access to both structures. While we can model complexity of a structure based on the number of steps in movement chains, this information is uninformative without a second structure with which to compare complexity (Anderson 2004, Pritchett & Whitman 1995). Thus, we must take this condition as a statement that the parser avoids performing QR. The problem becomes stating how strict the avoidance is, or in what circumstances the parser opts to perform QR. If we simply say avoidance fails when inverse scope is pragmatically supported, we land in an uneasy conspiracy where QR is only applied in cases where it pragmatically supported because those are the cases under which it is the interpretation we observe.

Finally, one of stated advantages of PSE is that it allows for pragmatic factors to influence structures more directly, by encouraging QR when context encourages QR (Anderson 2004, Tanaka 2015, Wurmbrand 2018). In addition to the fact this once again verges on the uneasy conspiracy of an oracular parser, it is also not actually an advantage over GSE. In the parsing procedure for GSE we have outlined, there is a clear opportunity for pragmatic factors to influence the final interpretation. After raising has been completed based on a particular trigger, the structure can be changed based on pragmatic factors. The structure will necessarily be the most complex grammatical scope structure, and thus it is possible to either keep this structure or revert to simple structure depending on which interpretation would be more felicitous in context.

With this understanding of GSE as a system where the parser frequently performs QR and of PSE as a system where QR is generally avoided, we wish to test the predictions of these two approaches in online processing. In the following Experiment we tested verb phrases ellipsis configurations with either scopally ambiguous or scopally rigid first and second conjuncts. With the parallelism constraint on elliptical structures as a probe, we can see in this measure how eager the parser is to compute scope relationships. Evidence of raising at the ellipsis site when the first conjunct blocks inverse scope would indicate that the parser performs QR as predicted by GSE, even when another constraint like scope parallelism is present. If we see that the parser doesn't appear to be performing QR in cases where these moves violate parallelism, this could support the idea that the parser avoids QR based on wider context as suggested by PSE.

5. Experiment. The goal of this experiment was to observe the parser's behavior when there are other constraints on the scope structure beyond an economy condition. The sentences tested contain a verb phrase ellipsis (VPE) structure, which is subject to a Scope Parallelism constraint (Johnson 2001, Fox 2002, May 1977, Merchant 2001, Sag 1976). The Scope Parallelism constraint requires that the conjuncts have the same scope structure. Thus, both conjuncts must have surface scope, or both must have inverse scope. In addition to the parallelism constraint, the structures in this study either promote or disallow inverse scope at the first and second conjunct, meaning that the processing of the quantifier containing ellipsis site is subject to two different constraints: the local clause scope preference and parallelism with the first conjunct. PSE and GSE make different predictions about how the ellipsis site should be processed. For PSE, the parser should attempt to avoid raising as much as possible and should face the greatest costs when raising must be performed. Thus, we would anticipate only observing a penalty in the case where inverse scope is permitted in in both conjuncts, as this is the only condition where QR application will ultimately be grammatical under the Scope Parallelism constraint. GSE would instead predict costs in the conditions where inverse scope is blocked in the second conjunct, as the QR performed in these cases would not be scope informative. An additional cost would be predicted after the ellipsis site for the case where inverse scope is licensed by the first conjunct but prohibited by the second conjunct, as the parser must undo the QR application which is not licensed under GSE which results in a new violation of Scope Parallelism. With these hypotheses in mind, we used the grammatical maze task with Automaze generated alternatives to examine reading times at the ellipsis site as a measure of real time parsing behavior (Boyce et al., 2020; Forster et al., 2009; Gulordava et al., 2018; Witzel et al., 2012). We also examined the reading time at earlier quantifier regions throughout these constructions to provide insight into when and how the parser calculates scope prior to reaching the ellipsis site.

5.1. MATERIALS AND METHODS. Participants were 80 English speakers contracted through Prolific, a platform for the recruitment of participants in web-based experiments. An initial 81 participants were recruited, though 1 was dropped from the analysis for failure to comply with task directions. The target recruitment number for all experiments was selected to ensure sufficient statistical power for the experiments and fillers respecting the number of items and conditions. All participants were compensated $6 for their participation. The task itself was estimated to take around 30 minutes though participants were allowed 1 hour to complete the task with the ability to pause during that time. The average completion time was just over 30 minutes, including the reading of forms and any breaks taken.

The task materials consisted of 24 items with 4 conditions each, which were structured as conjoined sentences with VPE. In this study we employed a 2x2 design manipulating the potential for inverse scope readings in the first conjunct and manipulating the potential for inverse scope readings of the second conjunct. In the first conjunct the subject was either composed of the string *At least one NP* or a length matched string of the form *That NP*. The string *At least one NP* allows for inverse scope interpretations of a lower quantifier, as it is a quantifier itself, and has even been suggested to promote inverse scope interpretations when in subject position (Beghelli & Stowell, 1997). On the other hand, the demonstrative *That NP* should block inverse scope interpretations as raising a lower quantifier would not be scope informative. In the second conjunct we perform a similar manipulation of the subject being either of the form *the Ns* or *some Ns*. Like the first conjunct the presence of the quantifying phrase *some Ns* allows for inverse scope, while the form *the Ns* renders QR not scope informative. An example item is presented in (5). In all cases, care was taken to construct the items such that the sentences shared a similar plausibility and structure outside of these critical manipulations.

(5) a. At least one executive will evaluate each hire, and some managers will too because the company is being careful
  b. At least one executive will evaluate each hire, and the managers will too because the company is being careful
  c. That extremely serious executive will evaluate each hire, and some managers will too because the company is being careful
  d. That extremely serious executive will evaluate each hire, and the managers will too because the company is being careful

Each participant rated a total of 96 sentences, consisting of 24 verb phrases ellipsis sentences and 72 filler sentences. The filler sentences consisted of three subsets with similar length and complexity to the VPE sentences. Fillers examined the processing of ambiguous participle attachment, sluicing constructions and polarity illusion sentences.

Sentences were presented using Ibex Farm (Drummond 2013). Sentences were presented one word in the style of the maze task, where the participant needs to select between the grammatical continuation of the sentence and an ungrammatical foil which was generated using Automaze (Boyce Futrell & Levy 2020, Gulordava Bojanowski Grave Linzen & Baroni 2018). Participants indicated their selection by pressing either the "e" or "i" key depending on which side of the screen their choice was on. If the participant selected the ungrammatical foil the trial ended immediately, otherwise the next pair was immediately presented. After each sentence participants could begin the next sentence at any time by pressing the space key. Participants completed a short practice section, received a restatement of the instructions, and then began the experiment. Presentation order was randomized across participants using Ibex Farm's Latin squaring and randomization method.

5.2. RESULTS. Data was analyzed using mixed-effects models which were estimated using the lme4 package for each region of interested (R Core Team 2013). Both the fixed effects were coded with the inverse scope permitting structure weighted 1/2 and the inverse scope blocking structure weighted -1/2. Each model was estimated over log transformed reading times with random intercepts for subjects and items. Reading times greater than 10,000 milliseconds were removed as were reading times lower than 100 milliseconds. Three trials were removed after analysis of Cook's Distance indicated they were influential outliers in either the ellipsis site or wrap-up region. Looking broadly at the pattern of reading times in the full sentence presented in Figure 1, we can see several areas where significant differences emerge between conditions. Setting aside those in the very beginning of the sentence which owe to the difference in the processing of the relatively predictable string *at least one NP* and the less predictable string *That NP*, the remaining areas appear to be related to quantifiers and or the ellipsis site.

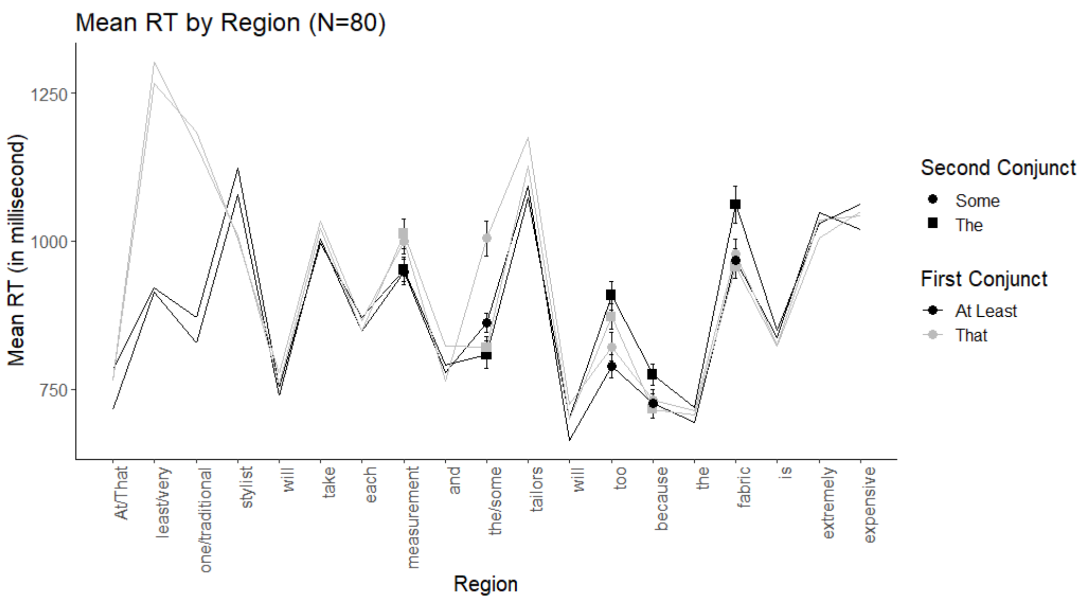


Figure 1. Mean RT by Region

Starting with the region marked by *too*, the ellipsis site, we notice a significant main effect for second conjunct *the*, such that reading times were slower when *the* was present over *some* ($p<0.001$). This can be seen visually in the right panel of Figure 2 and the full model output is presented in Table 1. Looking at the region marked by *because*, the spillover for the ellipsis site, we see a significant main effect for first conjunct *at least* such that reading times are slower for conditions with *at least* than for those with *that* ($p<0.05$). This effect was accompanied by an interaction between first conjunct *at least* and second conjunct *the,* such that the condition with both *at least* and *the* was read slower ($p < 0.05$). This can be seen visually in the right panel of Figure 2 and the full model output is presented in Table 2. In the sentence wrap up regions a visual trend can be seen on the first noun in the because clause, *fabric,* as indicated in Figure 1. There appears to be a main effect such that sentences with first conjunct *at least* are read slower ($p < 0.05$) as well as a weak interaction such that sentences with first conjunct *at least* and second conjunct *the* are read slower ($p < 0.10$). Full model output is presented in Table 3.

Too	Estimate	Std. Error	T Value	P Value
Intercept	-0.005	0.024	0.211	0.834
FC That	-0.001	0.019	-0.043	0.966
SC The	0.085	0.018	**4.639**	**0.000**
That * The	-0.049	0.037	-1.316	0.188

Table 1. Model output for "Too" Region

Because	Estimate	Std. Error	T Value	P Value
Intercept	0.005	0.023	0.238	0.812
FC That	-0.032	0.015	**-2.190**	**0.028**
SC The	0.009	0.015	0.538	0.533
That * The	0.059	0.030	**-1.972**	**0.048**

Table 2. Model output for "Because" Region

Wrap Up	Estimate	Std. Error	T Value	P Value
Intercept	-0.001	0.031	-0.034	0.973
FC That	-0.032	0.016	**-2.020**	**0.043**
SC The	0.016	0.016	1.043	0.297
That * The	-0.055	0.031	*-1.749*	*0.081*

Table 3. Model output for Wrap Up Region Noun

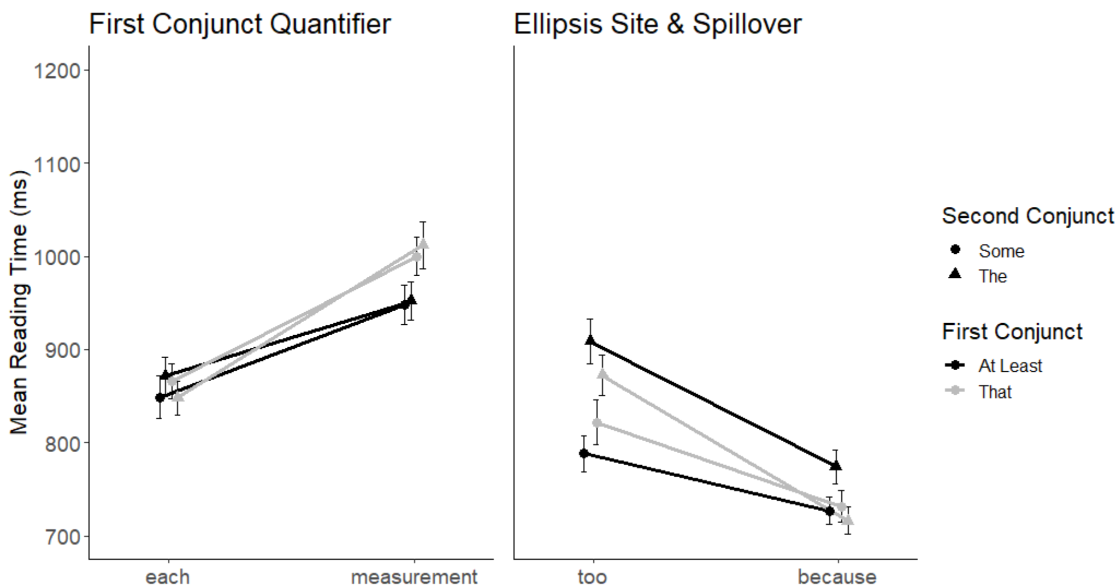


Figure 2. Mean RT in First Conjunct Quantifier and Ellipsis Site and Spillover

Moving to the quantifier regions, we also see significant effects in the quantifiers in the first conjunct. We observe a main effect for first conjunct *that* at the noun in the phrase *each N*, such that the nouns in the first conjunct *that* condition are read slower than the *at least* sentences ($p<0.05$). This can be seen visually in the left panel of Figure 2 and full model output is presented in Table 4.

FC Noun	Estimate	Std. Error	T Value	P Value
Intercept	-0.010	0.034	0.284	0.778
FC That	-0.037	0.015	**2.450**	**0.014**
SC The	0.003	0.015	-0.201	0.841
That * The	-0.003	0.030	0.086	0.093

Table 4. Model output for First Conjunct Noun

We also observe significant effects at the subject positions in the second conjunct. At the point of the determiner, *some* or *the*, we observe main effects for both first conjunct *that* ($p<0.001$) and second conjunct *some* ($p< 0.001$), as well as an interaction ($p< 0.05$). Full model output is presented in Table 5. The effect of first *that* extends into the noun region of the second conjunct subject ($p< 0.001$). Full model output for this region appears in Table 6.

SC Quantifier	Estimate	Std. Error	T Value	P Value
Intercept	0.003	0.024	0.142	0.888
FC That	0.070	0.015	**4.516**	**0.000**
SC The	-0.143	0.015	**-9.288**	**0.000**
That * The	-0.063	0.031	**-2.031**	**0.043**

Table 5. Model output for Second Conjunct Quantifier

SC Quantifier	Estimate	Std. Error	T Value	P Value
Intercept	-0.008	0.036	0.242	0.810
FC That	-0.068	0.018	**3.864**	**0.000**
SC The	0.018	0.017	-1.048	0.294
That * The	-0.015	0.035	-0.426	0.670

Table 6. Model output for Second Conjunct Noun

5.3. DISCUSSION. At the ellipsis site we observed a penalty only for the conditions which have *the* in the subject position of the second conjunct, as predicted by GSE. In these two conditions QR in the second conjunct does not create a scope informative structure and should be ungrammatical. Thus, observing a penalty in the ellipsis site suggests that the parser has attempted to raise the quantifier *each* within the ellipsis site, and reverted the structure as it violates GSE. If the parser did not attempt QR, as predicted by PSE, we would have no reason to expect a penalty in this position, and in fact would expect it to be ready quickly in comparison to the other conditions as the other conditions provide a stronger context for performing QR.

In the spillover region for ellipsis site, we see a penalty in the second conjunct *the* condition when *at least* is in the first conjunct. This appears to be an additional penalty that occurs when the parser must undo the raising of the first conjunct, as the reversion of the second conjunct to surface scope requires the first conjunct be reverted under the parallelism constraint. This is because undoing the raising in the second conjunct forces a conflict with the first conjunct.

What can these results tell us? First, performing QR appears to simply require a quantifier in memory, not necessarily local context, as costs are observed at the ellipsis site when the second conjunct subject has a configuration, subject *the*, that prevents any QR application from being scope informative. Second, parallelism appears to be a secondary constraint, as a penalty is uniquely observed in the condition which violates parallelism at the spillover region, after the ellipsis site where the cost for GSE violating QR is observed. Third, the spillover penalty for the *at least* by *the* condition extends into later wrap up regions, suggesting that the work of resolving parallelism and constructing the proper scope configurations remains a challenge for the parser as processing continues.

The predictions of GSE also appear to be borne out by the processing of the antecedent of the ellipsis site. In particular, the penalty observed at the noun position of *each NP* in the first conjunct strongly mirror the finding at the ellipsis site in the second conjunct. In both conjuncts we observe slowdowns when the subject of the conjunct blocks inverse scope, suggesting that the parser is attempting to perform QR for *each* in both the first conjunct and ellipsis positions and facing penalties when the context prohibits inverse scope. Under PSE we would anticipate no difficulty in this region, as QR can simply be avoided.

Finally, in the subject position of the second conjunct subject we observe a penalty in the *that* by *some* condition, with the main effect for first conjunct *that* spilling over into the noun region. These effects could stem from several sources. We suspect that the penalty stems from a different parallelism violation. In these items the first conjunct subject contains multiple adjectival descriptors when in the *that* condition, while the second conjunct lacks additional material beyond the determiner and noun. The uniquely slow reading time in the *that* by *some* condition might similarly stem from a violation of a different notion of parallelism where the penalty owes to the mismatch in the length of the subjects of the clauses.

6. General discussion. In the experiment we observed the parser's behavior in sentences where the possibility QR was influenced both by local context and a global parallelism constraint. We observed reading time penalties when QR would be locally not scope informative in both the first and second conjunct. This result patterns with the expectations of the theory of GSE outlined in section 3, as the penalties appear to be associated with the reversion of ungrammatical scope structures. PSE would have predicted penalties in cases where QR is not avoided, but the condition where QR is most contextually supported, *at least* by *some*, was generally read the fastest in regions of interest. Taken together, the results of this experiment support the idea that scope economy is a condition which is a part of the grammar.

In addition to supporting the idea the parser must be organized to reflect GSE, the results of this study also indicate that the parser adopts a strategy that is neither strongly incremental nor non-incremental. Rather, we observe penalties throughout the processing of the sentence, as well as in wrap-up regions, suggesting that the process by which the parser computes scope relations is active throughout the processing of an input sentence.

Future research should focus on identifying where the parser falls with respect to optimizing for grammatical coverage, incrementality, and minimizing penalties for QR that is not scope informative. As we suggested in section 3, two fruitful areas of research are the investigation of sentences in which quantifiers are commutative, and the investigation of how the structural relations between the quantifier in memory and the scope sensitive item trigger influence QR. As an additional follow-up, we plan to investigate the same paradigm as detailed in the above experiment, with the exception that the ellipsis site be overtly filled with a quantifier. Under such conditions, the Scope Parallelism constraint that applies to ellipsis is not present, and we might expect a different pattern of cost in the second conjunct and wrap-up regions.

References

Adger, David. 2003. *Core syntax: A minimalist approach*. Vol. 20. Oxford: Oxford University Press.

Anderson, Catherine 2004. *The structure and real-time comprehension of quantifier scope ambiguity*. Evanston, IL: Northwestern University dissertation.

Barker, Chris. 2021. Rethinking scope islands. *Linguistic Inquiry* [Early access online-only]. https://doi.org/10.1162/ling_a_00419.

Beghelli, Filippo & Tim Stowell. 1997. Distributivity and negation: The syntax of *each* and *every*. In Anna Szabolsci (ed.), *Ways of scope taking*. 71–107. Dordrecht: Springer. https://doi.org/10.1007/978-94-011-5814-5_3.

Bott, Oliver & Fabian Schlotterbeck. 2015. The processing domain of scope interaction. *Journal of Semantics* 32(1). 39–92. https://doi.org/10.1093/jos/fft015.

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

Brendel, Cole Ira. 2019. An investigation of numeral quantifiers in English. *Glossa: A Journal of General Linguistics.* 4(1). 104. https://doi.org/10.5334/gjgl.391.

Carnie, Andrew. 2021. *Syntax: A generative introduction*. Malden: Wiley-Blackwell.

Cecchetto, Carlo. 2004. Explaining the locality conditions of QR: Consequences for the theory of phases. *Natural Language Semantics*. 12(4). 345–397. https://doi.org/10.1007/s11050-004-1189-x.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1977. On wh-movement. In Peter Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*. New York, NY: Academic Press.

Clifton, Charles & Lyn Frazier. 1989. Comprehending sentences with long-distance dependencies. In Greg N. Carlson & Michael K. Tanenhaus (eds.) *Linguistic structure in language processing*, 273–317. Dordrecht: Springer.

Crain, Stephen & Janet Dean Fodor. 1985. Rules and constraints in sentence processing. *Proceedings of the North East Linguistics Society* 15(1). 8.

Drummond, Alex. 2013. Ibex farm. Online server. http://spellout.net/ibexfarm.

Farkas, Donka. 1981. Quantifier scope and syntactic islands. *Proceedings of the Chicago Linguistic Society* 17. 59–66.

Fodor, Janet Dean & Fernanda Ferreira (eds.). 2013. *Reanalysis in sentence processing* (Studies in Theoretical Psycholinguistics 21). Dordrecht: Springer. https://doi.org/10.1007/978-94-015-9070-9.

Fodor, Janet Dean. 1978. Parsing strategies and constraints on transformations. *Linguistic Inquiry*, 9(3). 427–473.

Fodor, Janet Dean. 1981. Does performance shape competence? *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*. 295(1077). 285–295. https://doi.org/10.1098/rstb.1981.0140.

Forster, Kenneth, Christine Guerrera & Lisa Elliot. 2009. The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods* 41. 163– 171. https://doi.org/10.3758/BRM.41.1.163.

Fox, Danny. 1995. Economy and scope. *Natural Language Semantics* 3. 283–341. https://doi.org/10.1007/BF01248820.

Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, MA: MIT press.

Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33(1). 63–96. https://doi.org/10.1162/002438902317382189.

Frazier, Lyn. 1979. *On comprehending sentences: Syntactic parsing strategies*. Mansfield, CN: University of Connecticut dissertation.

Frazier, Lyn & Charles Clifton. 1998. Sentence reanalysis, and visibility. In Fernanda Ferreira & Janet Dean Fodor (eds.), *Reanalysis in sentence processing*, 143– 176. Dordrecht: Springer.

Frazier, Lyn & Fodor, Janet Dean. 1978. The sausage machine: A new two-stage parsing model. *Cognition* 6(4). 291–325. https://doi.org/10.1016/0010-0277(78)90002-1.

Frazier, Michael, Lauren Ackerman, Peter Baumann, David Potter & Masaya Yoshida. 2015. Wh-filler-gap dependency formation guides reflexive antecedent search. *Frontiers in Psychology* 6. 1504. https://doi.org/10.3389/fpsyg.2015.01504.

Freedman, Sandra & Kenneth Forster. 1985. The psychological status of overgenerated sentences. *Cognition*. 19(2). 101–131. https://doi.org/10.1016/0010-0277(85)90015-0.

Gibson, Edward. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*. 68(1). 1–76. https://doi.org/10.1016/S0010-0277(98)00034-1.

Gulordava, Kristina, Piotr Bojanowski, Edouard Grave, Tal Linzen & Marco Baroni. 2018. Colorless green recurrent networks dream hierarchically. In Marilyn Walker, Heng Ji & Amanda Stent (eds.), *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics* (Human Language Technologies, Vol. 1), 1195–1205. New Orleans: Association for Computational Linguistics.

Hornstein, Norbert. 1994. An argument for minimalism: The case of antecedent-contained deletion. *Linguistic Inquiry* 25(3). 455–480.

Hornstein, Norbert. 1995. *Logical form: From GB to minimalism.* Oxford: Blackwell.

Hornstein, Norbert. 1996. *Minimalism and QR*. Ms. University of Maryland.

Huang, James. 1995. Logical form. In Gert Webelhuth (ed.), *Government and binding theory and the minimalist program*. 125-175.

Johnson, Kyle. 2001. *What VP ellipsis can do, and what it can't, but not why*. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory,* 439–479.

Kazanina, Nina, Ellen Lau, Moti Lieberman, Masaya Yoshida & Colin Phillips. 2007. The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language*. 56(3). 384–409. https://doi.org/10.1016/j.jml.2006.09.003.

Koopman, Hilda, Dominque Sportiche & Edward Stabler. 2013. *An introduction to syntactic analysis and theory.* Malden, MA: Wiley-Blackwell.

Lasnik, Howard & Tim Stowell. 1991. Weakest crossover. *Linguistic Inquiry* 22(4). 687–720.

Lasnik, Howard. 1993. Lectures on minimalist syntax. *University of Connecticut Working Papers in Linguistics Occasional Papers* 1.

May, Robert. 1977. *The grammar of quantification*. Cambridge, MA: MIT dissertation.

May, Robert. 1985. *Logical form: Its structure and derivation.* Cambridge MA: MIT press.

Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.

Miller, George & Noam Chomsky. 1963. Finitary models of language users. In Duncan Luce (ed.), *Handbook of mathematical psychology,* 419–491. Hoboken, NJ: John Wiley.

Orth, Wesley, Masaya Yoshida & Shayne Sloggett. 2020. Negative polarity item (NPI) illusion is a quantification phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 47(6). 906–947. https://doi.org/10.1037/xlm0000957.

Phillips, Colin. 1996. *Order and structure*. Cambridge MA: MIT dissertation.

Phillips, Colin. 2006. The real-time status of island phenomena. *Language*. 82(4). 795–823. https://doi.org/10.1353/lan.2006.0217.

Phillips, Colin, Matthew Wagers & Ellen Lau. 2011. Grammatical illusions and selective fallibility in real-time language comprehension. In Jeffrey Runner (ed.), *Experiments at the interfaces*, 147–180. Leiden: Brill. https://doi.org/10.1163/9781780523750_006.

Pritchett, Bradly & John Whitman. 1995. Syntactic representation and interpretive preference. In Reiko Mazuka & Noriko Nagai (eds.), *Japanese sentence processing,* 65–76. Hillsdale, NJ: Erlbaum.

Reinhart, Tanya. 2006. Scope shift with numeral indefinites. In Svetlana Vogeleer & Liliane Tasmowski (eds.), *Non-definiteness and plurality*, 291–310. Amsterdam. John Benjamins.
Sag, Ivan. 1976. *Deletion and logical form.* Cambridge MA: MIT dissertation.
Schneider, David & Colin Phillips. 2001. Grammatical search and reanalysis. *Journal of Memory and Language* 45(2). 308–336. https://doi.org/10.1006/jmla.2001.2777.
Steedman, Mark. 1996. *Surface structure and interpretation.* Cambridge MA: MIT press.
Stowe, Laurie. 1992. The processing implementation of syntactic constraints: The sentence matching debate. In Helen Goodluck &Michael Rochemont (eds.), *Island constraints: Theory acquisition and processing*, 419–443. Dordrecht: Springer.
Sturt, Patrick & Matthew W. Crocker. 1996. Monotonic syntactic processing: A crosslinguistic study of attachment and reanalysis. *Language and Cognitive Processes* 11(5). 449–494. https://doi.org/10.1080/016909696387123.
Syrett, Kristen. 2015a. Experimental support for inverse scope readings of finite-clause-embedded antecedent-contained-deletion sentences. *Linguistic Inquiry* 46(3). 579–592. https://doi.org/10.1162/LING_a_00194.
Syrett, Kristen. 2015b. QR out of a tensed clause: Evidence from antecedent-contained deletion. *Ratio* 28. 395–421. https://doi.org/10.1111/ratsi.12107.
Tanaka, Misako. 2015. *Scoping out of adjuncts: Evidence for the parallelism between QR and wh-movement*. London: University College London dissertation.
R Core team. 2013. *R: A language and environment for statistical computing.*
Witzel, Naoko, Jeffrey Witzel & Kenneth Forster. 2012. Comparisons of online reading paradigms: Eye tracking, moving window, and maze. *Journal of Psycholinguistic Research* 41. 105–128. https://doi.org/10.1007/s10936-011-9179-x.
Wurmbrand, Susanne. (2018). The cost of raising quantifiers. *Glossa: A Journal of General Linguistics* 3(1). 19. https://doi.org/10.5334/gjgl.329.