University of Groningen

Sentence-internal same and its quantificational licensors

Brasoveanu, Adrian; Dotlacil, Jakub

[Link to publication in University of Groningen/UMCG research database](#)

Sentence-internal *same* and its quantificational licensors:
A new window into the processing of inverse scope *

Adrian Brasoveanu
Department of Linguistics,
University of California Santa Cruz

Jakub Dotlačil
Center for Language and Cognition,
University of Groningen

Abstract This paper investigates the processing of sentence-internal *same* with four licensors (*all*, *each*, *every*, and *the*) in two orders: licensor+*same* (surface scope) and *same*+licensor (inverse scope). Our two self-paced reading studies show that there is no general effect of surface vs. inverse scope, which we take as an argument for a model-oriented view of the processing cost of inverse scope: the inverse scope of quantifiers seems to be costly because of model structure reanalysis, not because of covert scope operations.

The second result is methodological: the psycholinguistic investigation of semantic phenomena like the interaction of quantifiers and sentence-internal readings should generally involve a context that prompts a deep enough processing of the target expressions. In one of our two studies, participants read the target sentences after reading a scenario and they were asked to determine whether the sentence was true or false relative to the background scenario every time. In the other study, the participants read the same sentences without any context and there were fewer follow-up comprehension questions. The relevant effects observed in the study with contexts completely disappeared in the out-of-context study, although the participants in both studies were monitored for their level of attention to the experimental task.

Finally, the results of the first, in-context experiment suggest that the processing of quantifier scope and sentence-internal readings happens in two stages, similar to the way interpretation unfolds in (classical) DRT: there seems to be a shallower level of meaning processing that is parallel to the process of constructing a DRS / mental discourse model for the current sentence / discourse; and there is a deeper level of meaning processing that corresponds to linking this DRS to the actual, "real-world" model, which involves constructing an embedding function that verifies the DRS.

Keywords: distributivity, quantification, scope, sentence-internal *same*, processing

1 Introduction: Sentence-internal readings and inverse scope

Languages have lexical means to compare two elements and express identity, difference or similarity between them. English uses adjectives of comparison (AOCs) like *same*, *different*, and *similar* for this purpose.

AOCs can have both sentence-external and sentence-internal readings. In the case of sentence-external readings, AOCs compare an element in the current sentence and an element mentioned in a previous sentence, as exemplified in (1) below.

(1)   a.   Arnold saw *'Waltz with Bashir'*.
        b.   Heloise saw the *same* movie.

In sentence-internal readings, AOCs make a comparison that is internal to the sentence in which they occur without referring to any previously introduced element. This is exemplified in (2) below.

(2) $\left\{\begin{array}{l}\text{All the students}\\\text{Each student}\\\text{Every student}\\\text{The students}\end{array}\right\}$ saw the *same* movie.

As observed in Carlson 1987, sentence-internal readings of AOCs must be licensed by a semantically (but not necessarily morphologically) plural element. For example, if we replace the semantically plural subjects in (2) above with a proper name, the only available reading is the sentence-external one, as shown by the example below.

(3)   #Sue saw the same movie.

Furthermore, Carlson (1987) argues that AOCs and their sentence-internal licensors must be in the same scope domain. This explains, for example, why *everyone* can license the sentence-internal *same* in (4a) but it cannot do so in (4b).

(4)   a.   The same waiter served everyone.   (Barker 2007)
        b.   #The same waiter whispered that everyone left.

Carlson (1987), Moltmann (1992), Beck (2000), Dotlačil (2010), Brasoveanu (2011) among others discuss further restrictions on licensing sentence-internal readings and various differences between AOCs with respect to these restrictions. In particular, they observe that while any semantically plural element can license sentence-internal readings of *same*, the conditions for licensing sentence-internal *different* are much more stringent. The ability to license sentence-internal *different* greatly depends on the type of determiner that the licensor contains, as shown in (5) below.

(5) $\left\{\begin{array}{l} \text{?All the students} \\ \text{Each student} \\ \text{Every student} \\ \#\ \text{The students} \end{array}\right\}$ saw a *different* movie.

We see that the plural definite *the students* is not a possible licensor of sentence-internal singular *different*, unlike distributive quantifiers with the determiner *every* or *each*. The universal quantifier with the determiner *all* is somewhat worse than the distributive quantifiers but significantly better than the plural definite; see Brasoveanu & Dotlačil 2012 for an acceptability judgment study confirming these intuitive judgments.

These generalizations about the differences between the licensors of *different* are based on native speakers' intuitions about the acceptability and interpretation of sentences with sentence-internal *different*, whether informally or systematically collected. It is possible that using other, finer-grained experimental methodologies can provide data revealing that the situation with sentence-internal *same* is just as complex as the one for *different*. In fact, the questionnaire study in Brasoveanu & Dotlačil 2012 already provides an indication that this might be the case: in that study, distributive quantifiers were shown to be somewhat dispreferred as licensors of sentence-internal *same*.

In this paper, we investigate this issue further by means of experimental methodologies other than acceptability or truth-value judgment tasks. In particular, we examine the incremental processing of sentence-internal *same* using the self-paced reading paradigm of Just, Carpenter & Woolley (1982) in order to ascertain whether any differences between licensors show up during the process of incremental interpretation and if so, at which point.

Most importantly, however, we investigate the extent to which the licensing of sentence-internal *same* depends on its structural position relative to the licensor.[1] This paper investigates how sentence-internal *same* is processed:

  i. with four of its licensors: universal quantifiers like *all the students* (abbreviated as ALL), distributive quantifiers like *each student* (abbreviated as EACH), distributive quantifiers like *every student* (abbreviated as EVERY), and plural definites (abbreviated as THE), and

  ii. in two scopes: SURFACE-SCOPE, exemplified in (2) above, and INVERSE-SCOPE, exemplified in (6) below.

(6) The *same* student saw $\begin{Bmatrix} \text{all the movies} \\ \text{each movie} \\ \text{every movie} \\ \text{the movies} \end{Bmatrix}$.

We will discuss the results of:

  i. a self-paced reading study in which these 8 conditions (4 licensors × 2 orders) were investigated in context, that is, after the presentation of a background scenario relative to which the target sentence was either true or false (truth / falsity was balanced across conditions)

  ii. a self-paced reading study in which the same target sentences were presented out of context.

We use self-paced reading as our experimental method for two reasons. First, it is a common methodology in previous studies of inverse-scope processing (Tunstall 1998, Anderson 2004, Dotlačil & Brasoveanu 2014 among others). Second, the real-time interpretation of sentence-internal *same* and its interaction with scope has not been systematically studied before.

---

1 Brasoveanu & Dotlačil (2012), for example, considers only surface-scope configurations in which the quantificational licensors are always higher than the AOC in the surface structure.

The main results of the experiments are as follows. First, we do not see any across-the-board effect of inverse scope. When licensors occur in object position they need to take inverse scope to license *same*. Yet, this inverse scope does not lead to difficulties observable in an increase of reading times. This provides a novel argument in favor of processing theories of inverse scope that do not assign any inherent cost to the covert syntactic or semantic operations needed to derive inverse scope, as explained in Section 2.

Second, the in-context experiment shows that EACH and THE cause a slowdown when licensing *same*. These findings confirm the results of the acceptability study reported in Brasoveanu & Dotlačil 2012, and increase our confidence that the self-paced reading task was actually able to target the intended interpretive effects. Importantly, the differences between EACH / THE on one hand and ALL / EVERY on the other hand disappear in the second (out-of-context) experiment. We take this to indicate that participants don't interpret *same* deeply enough in out-of-context tasks to really enforce the licensing requirement associated with its sentence-internal reading. This is particularly interesting because we have no independent reasons to think that the participants did not pay attention to the task: most of them correctly answered the majority of the comprehension questions in this second experiment. This is the second important result of our studies: it seems that the experimental investigation of deep interpretive effects that are of interest to formal semanticists (i.e., that mainly arise as a consequence of semantic composition) require the presence of fairly rich and explicit contexts to manifest themselves behaviorally.

Finally, the results suggest that the processing of AOC licensing and quantifier scope happens in two stages, very similar to the way interpretation unfolds in Discourse Representation Theory (DRT, Kamp 1981, Kamp & Reyle 1993). There seems to be a shallower level of meaning processing that is parallel to the process of constructing a DRS / mental discourse model for the current sentence / discourse. And there is a deeper level of meaning processing that corresponds to linking this DRS to the actual, "real-world" background situation; this corresponds to constructing an embedding function (partial variable assignment) that verifies the DRS, which involves linking this DRS (and therefore the mental discourse model the DRS encodes) to the actual, "real-world" model.

The paper is structured as follows. We first show in Section 2 how studying the processing of *same* is relevant for our understanding of scope and of the semantics of quantifiers and AOCs, and we summarize previous studies

on the processing of quantifier scope and AOCs. We introduce the first (in-context) self-paced reading experiment and the resulting generalizations in Section 3. Section 4 introduces the second (out-of-context) self-paced reading experiment and briefly compares its results to the first one. Section 5 puts forth an account of the generalizations extracted from the two experiments and Section 6 concludes.

The experimental items for the two studies are provided in the appendix.

2 Previous theories and their predictions

2.1 Two processing theories of quantifier scope

It is generally assumed that in a sentence with two scopally interacting quantifiers, the inverse-scope interpretation is dispreferred and harder to process (Ioup 1975, Tunstall 1998, Anderson 2004, Reinhart 2006, AnderBois, Brasoveanu & Henderson 2012, among many others). The processing cost is indicated by increased reading times of inverse scope readings, as compared to surface scope interpretations, in on-line studies (Tunstall 1998, Anderson 2004, among others). Consider for example the sentence in (7) below: the most salient and easiest interpretation for this sentence is one in which a single boy climbed every tree (the surface-scope interpretation), as Anderson (2004) shows.

(7)   A boy climbed every tree.

This observation can be explained in two different ways. One approach is to explain the difficulties associated with inverse scope in terms of covert scope operations: inverse scope requires an extra operation (Tunstall 1998, Anderson 2004, Pylkkänen & McElree 2006, Reinhart 2006, among others) to derive the requisite logical form / semantic representation.

One way of fleshing out this approach is to say that the quantified object has to undergo an extra quantifier raising (QR) in the inverse-scope reading, as shown in Figure 1, modeled after Fox 2000, where quantifiers always adjoin to VP in their original order of c-command, and the inverse-scope interpretation requires an extra movement and adjunction of a quantifier.

Another version of this approach appeals to type-shifting instead of QR: an optional type-shifter has to be inserted to derive inverse-scope readings (Hendriks 1993). Either way, an extra operation is necessary, which can explain the processing cost of inverse scope.
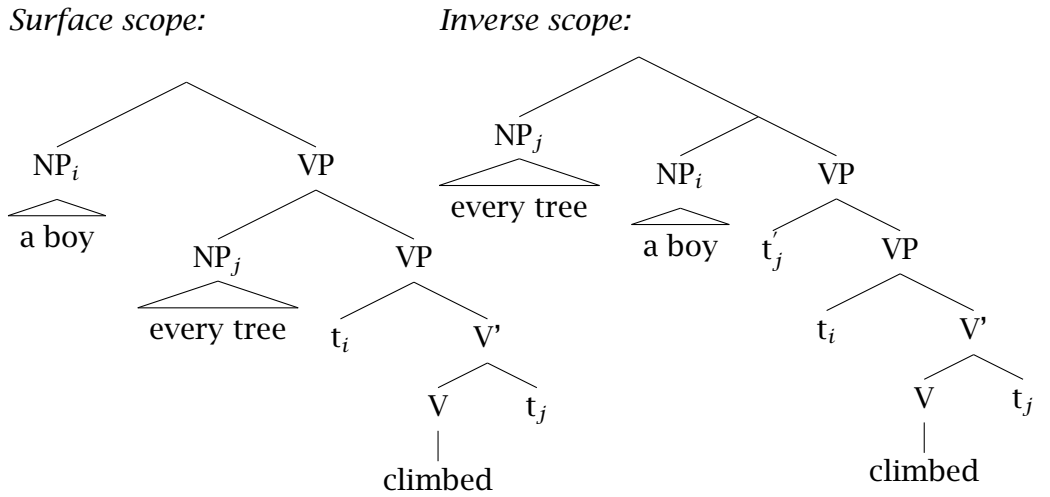
Figure 1 Two readings for *A boy climbed every tree* on a QR approach.

Alternatively, we could explain inverse-scope processing difficulties in terms of changes to the discourse model structure: inverse scope is harder because it requires revising the already built discourse model structure (Fodor 1982; see also Crain & Steedman 1985, Altmann & Steedman 1988).

To see this, consider how sentence (7) is interpreted on-line. We first hear or read *A boy climbed ...*, at which point we add a new entity to our discourse model that is a boy and that stands in the climbing relation to whatever direct object we are about to interpret. Then we hear or read the direct object *... every tree.* If we want the direct object quantifier to take wide scope, we need to revise the current discourse structure and introduce a set of boys, each of which is associated with a possibly distinct tree.[2]

---

2 As presented, this theory seems to predict that the scope of quantifiers should always be first and foremost based on their linear order. Such a simplified viewpoint suffices to understand this paper, but we note that the prediction is more complicated. It is possible that the model structure is not incrementally constrained or specified as each individual word is processed, but only when certain semantically coherent "chunks" / domains are processed (see Radó & Bott 2012 for more on this issue). Furthermore, if the speaker signals dependency (for instance, by using a dependent indefinite), the hearer might use that information and leave the relevant parts of the discourse model unspecified to avoid its subsequent revision. We will come back to this issue in Section 5.

2.2 Sentence-internal *same* and theories of inverse-scope processing

The AOC *same* on its sentence-internal reading enables us to distinguish between these two approaches to inverse scope. Sentence-internal *same* has to be in the scope of a semantically plural noun phrase (Carlson 1987) but because of its meaning, no revision of the discourse model structure is necessary when a quantifier takes inverse scope over it.

Consider, for example, the sentence in (8) below: *every movie* scopes and distributes over *same* to license its sentence-internal reading, but the model structure does not change. It will contain only one student both before and after the processing of *every movie.*

(8)   The same student saw every movie.

Thus, *same* can distinguish between the two theories of inverse-scope processing difficulties. If the covert scope operation itself is costly, the inverse scope needed to license *same* should lead to processing difficulties. If, on the other hand, the observed cost of inverse scope is due to changing / revising the discourse model structure, we should not find such difficulties when inverse scope is necessary to license *same*: whether *same* in sentence (8) is interpreted sentence-internally (which requires the universal *every movie* to take inverse scope) or sentence-externally, the discourse model will have only one student.

2.3 Previous work on the processing of AOCs

Sentence-internal readings of AOCs have been previously studied in the psycholinguistic literature, but there is no systematic study of the on-line interpretation of sentence-internal *same* in both surface-scope and inverse-scope contexts, and with multiple quantificational licensors.

Anderson 2004 studied sentence-internal *different*. Importantly for us, Anderson found out that sentences with *different* incurred processing costs when a semantically plural NP had to undergo QR in order to license the sentence-internal reading of *different*. This provides evidence that inverse-scope difficulties are not exhibited only by sentences in which ordinary indefinites are followed by distributive quantifiers, but they can also be observed with AOCs.

Dwivedi et al. 2010 examined event-related brain potentials in the processing of sentence-internal *same* and *different*. However, they only focused

on surface-scope structures and they only considered one licensor, EVERY. They found a slow negative shift in the *same* condition, which was missing in the case of *different*, and argued that the shift reveals processing difficulties. This is an interesting result, which suggests that in some ways, *same* might be harder to process than *different*, but it is orthogonal to the research reported in this paper.

3   The first self-paced reading experiment

This section describes the experimental methodology for the first self-paced reading experiment (Subsection 3.1) and presents the data analysis of the experimental results (Subsection 3.2). Section 5 provides an account of the generalizations.

3.1   Method, materials, procedure, and participants

We used a self-paced reading task to test how easy it is to process sentence-internal *same* (*i*) with 4 licensors: ALL, EACH, EVERY, and THE, and (*ii*) in 2 scopes: SURFACE-SCOPE (quantifier precedes *same*) and INVERSE-SCOPE (*same* precedes quantifier), for a total of $4 \times 2 = 8$ conditions. Each condition was tested 4 times, 2 times in sentences most likely judged as true relative to the background scenarios and 2 times in sentences most likely judged as false, for a total of 32 items.

Each item consisted of a scenario, the target sentence, and a follow-up *yes/no* comprehension question. The same follow-up question was used in every item. After reading the scenario, the participants moved on to a new screen where they read the target sentence word-by-word with all the words initially hidden (dashes of the appropriate length were displayed where the words should be) and the SPACE bar revealing the next word and hiding the preceding one (self-paced reading task; Just, Carpenter & Woolley 1982). Then, the follow-up *yes/no* question was displayed on a new screen.

An example in which the target sentence contains the licensor EACH taking surface scope to license *same* is provided in (9) below: the scenario is given in (9a), the sentence in (9b), and the follow-up question in (9c). The parallel item that exemplifies inverse scope is provided in (10).

(9)   SURFACE-SCOPE & EACH

    a. To prepare for fieldwork, three researchers — a botanist, a linguist, and an anthropologist — had to learn one of two languages spoken in the eastern Indonesian islands — Bahasa Indonesia or Ternate. The botanist learned Bahasa Indonesia, the linguist learned Bahasa Indonesia, and the anthropologist learned Bahasa Indonesia too.

    b. I think that each researcher learned the same language spoken in the eastern Indonesian islands.

    c. Am I right to think that?

(10)   INVERSE-SCOPE & EACH

    a. To prepare for fieldwork, two researchers — a botanist and an anthropologist — had to learn at least one out of three languages spoken in the eastern Indonesian islands — Bahasa Indonesia, Ternate or Tidore. The botanist learned Bahasa Indonesia, Ternate, and Tidore. The anthropologist learned nothing and used the botanist as his guide and advisor.

    b. I think that the same researcher learned each language spoken in the eastern Indonesian islands.

    c. Am I right to think that?

In general, scenarios consisted of 2 sets of entities, for example, researchers and languages, and a relation between them, for example, the *learn* relation. In true scenarios, it was specified that all the members of one set of entities were related to only one member in the other set. In false scenarios, it was specified that one member of one set of entities was related to a different entity than the other two members (see the appendix for the complete list of items).

There were 43 participants in the experiment, all of them undergraduate students from the University of California Santa Cruz (UCSC). They completed the experiment online on a UCSC hosted installation of the IBEX platform (http://code.google.com/p/webspr/) for course credit or extra-credit.

There were 32 test items and 35 fillers. The fillers had the same structure as the test items: they included a background scenario, a test sentence, and a comprehension question. The background scenario introduced two sets of entities and a relation between them. The target (self-paced reading) sentence started with the words *I think that...* so that its beginning would be indistinguishable from the test items. However, the self-paced sentence did

not include AOCs (with the exception of one filler, which had *same* licensed by an adjunct) and it used pluralities other than the ones tested — negative quantifiers, numerals, coordinated NPs. The structure of the sentences also was more varied, allowing ditransitives or pluralities appearing in adjunct positions.

Each of the test items was passed through all 8 conditions (2 scopes × 4 licensors). 8 lists were created following a Latin square design (in each list, every item appeared only in one condition). Every participant in the experiment responded to one list, consisting of 67 stimuli (32 experimental items + 35 fillers); the order of the stimuli was randomized for every participant (any two experimental items were separated by at least one filler). Every one of these 67 stimuli consisted of a background scenario, a target sentence, and the same follow-up *yes/no* comprehension question.

4 outlier participants were excluded because of their low answer accuracy (they had 15% or more incorrect answers). The final number of participants: 39. All responses ≤ 50 ms and ≥ 2000 ms were removed and the remaining observations were log transformed to mitigate the right-skewness characteristic of reading-time data.

3.2   Data analysis and resulting generalizations

Following Trueswell, Tanenhaus & Garnsey 1994 among others, we factored out the influence of word length and word position by running a linear mixed-effects regression. The regression had intercept-only random effects for subjects and two fixed effects — word length in characters and word position in the sentence. The resulting residualized log reading times (log RTs) were used for all subsequent analyses.

The main regions of interest (ROIs) for the analysis are the four words immediately following *same* / the quantificational licensor in object position. These words are underlined in the examples below. Note that they are identical (modulo sg. / pl. agreement) across all 8 conditions:

(11)   Main ROIs: the words immediately following *same* / the quantificational licensor in object position.

a.   ... $\left\{ \begin{array}{c} \text{all the} \\ \text{each} \\ \text{every} \\ \text{the} \end{array} \right\}$ researcher(s) learned the same <u>language spoken in the</u>
eastern Indonesian islands.

b. …the same researcher learned $\left\{\begin{array}{l}\text{all the}\\ \text{each}\\ \text{every}\\ \text{the}\end{array}\right\}$ <u>language(s) spoken in the</u>

eastern Indonesian islands.

These are the ROIs that follow the full experimental manipulation (sentence-internal *same* in combination with its licensors), so it is here that we expect to see processing differences (if any) between the 2 scopes and the 4 quantificational licensors. We examine only the four words following *same* / the quantifier in object position rather than the following five or six words because some of the experimental items were shorter than the one we used in the examples above (see the appendix for the full list of items), so considering the fifth or sixth word after the object would only be possible for a small subset of experimental items.

There are two other ROIs that are important for our overall argument: we want to examine the two words immediately following the quantificational licensors when they occur in subject position (i.e., in the surface scope order). These words are underlined in the example below.

(12) … $\left\{\begin{array}{l}\text{all the}\\ \text{each}\\ \text{every}\\ \text{the}\end{array}\right\}$ <u>researcher(s) learned</u> the same language spoken in the

eastern Indonesian islands.

The reason for this is as follows. Suppose we observe that EACH is slower than EVERY when we examine the main ROIs exemplified in (11a) and (11b) above (this will actually turn out to be true). This slowness might be a consequence of the semantic combination of *same* and the licensors, for example, it might be due to the fact that EACH is a worse licensor of sentence-internal *same* than EVERY, or it might simply be a consequence of the fact that EACH is inherently more difficult to process than EVERY.

We will be able to rule out the latter possibility if we examine the early regions exemplified in (12) above and we see that there are no significant differences between EACH and EVERY there (again, this will turn out to be true). If EACH is inherently more difficult to process than EVERY, we expect slowness in both the early and the main regions. But if EACH is more difficult to process than EVERY only when sentence-internal *same* needs to be licensed, we expect to see slowness only in the late regions.

Figures 2 and 3 plot the mean reading times (RTs) and the associated standard errors (SEs) for all these 6 ROIs, that is, both the early ones (the two words immediately following the quantifier / *same* in subject position) and the late / main ones (the four words following the quantifier / *same* in object position), in SURFACE and INVERSE scope respectively. The figures also plot the quantifier / *same* in subject and object position for completeness.
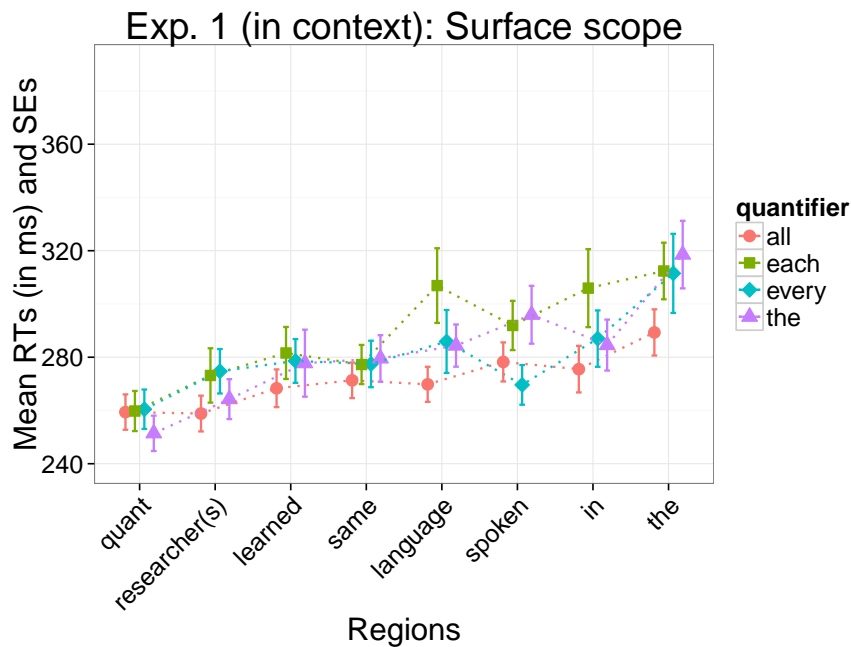
Figure 2   Mean RTs and SEs for all ROIs in Experiment 1 (in context); surface scope only.

Following the order of the words in these figures, we will henceforth refer to the two early ROIs (*researcher(s)* and *learned*) as Word 2 and Word 3, and to the four late ROIs (*language(s)*, *spoken*, *in*, and *the*) as Word 5, Word 6, Word 7, and Word 8.

Note that we plotted the quantifiers / *same* in the two plots (see Word 1 and Word 4) only for completeness. The quantifiers / *same* differ in several respects (frequency, quantificational nature for the licensors vs. anaphoric nature for *same*), all of which are possible confounds for our experimental manipulation — so the measurements in these two regions cannot tell us anything. In contrast, the other regions are identical in all respects (except for singular / plural number in certain cases), which minimizes the issue of confounds.
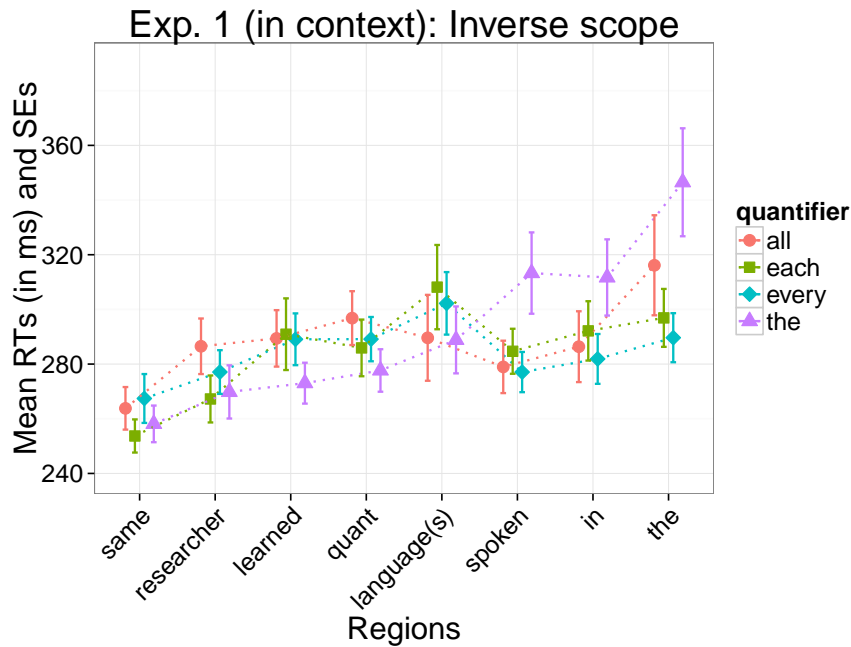
Figure 3 Mean RTs and SEs for all ROIs in Experiment 1 (in context); inverse scope only.

We can already see in these plots that there is no clear difference between the SURFACE and the INVERSE scope conditions: there is no systematic slowness, that is, upward shift, associated with all the inverse scope lines and indicative of processing difficulty. We will examine this data in much more detail when we turn to its statistical analysis in the next subsection.

Finally, in addition to the above six ROIs, we are interested in the RTs for full sentences: we will examine the sum of the residualized log RTs for full sentences because they have been previously argued to reveal the processing cost of inverse scope (Anderson 2004).

3.2.1 The statistical analysis of the six ROIs and resulting generalizations

For each ROI, we analyze the data by means of a linear mixed-effects regression model. The fixed effects are the ones associated with our experimental manipulation: quantifier type (EVERY as the reference level vs. ALL, EACH, and THE) and order (SURFACE-SCOPE as the reference level vs. INVERSE-SCOPE). We only report the models with main effects since these were the best models, that is, the models that optimally balanced parsimony and data fit according

to Likelihood Ratio (LR) tests that compared models with quantifier type $\times$ order interactions and models without interactions (with main effects only). The LR tests showed that adding interactions did not reduce deviance significantly (at the usual $\alpha = 0.05$ level). The only time the interaction model was better ($p = 0.035, \chi^2 = 8.6, df = 3$) was in the Word 2 region, and that was because the ALL & SURFACE condition was significantly faster than the ALL & INVERSE condition as well as the other quantifiers in the SURFACE condition. We return to this issue below.

We selected SURFACE-SCOPE as the reference level for the order factor since it has been consistently argued to be the easier one of the two in the previous psycholinguistic literature. We selected EVERY as the reference level for quantifier type because of its relative "blandness" as a universal quantifier:

    i. relative to EACH, which is more context-dependent (see Beghelli & Stowell 1997, particularly Section 5, and Dayal 2012),

    ii. relative to ALL, which has been argued to be primarily an exhaustivity marker in Brisson 2003,

    iii. and finally, relative to THE, which has been argued to have a collective reading by default rather than a universal distributive one (see Dotlačil 2010 and literature therein).

There is another reason for selecting EVERY as the reference level rather than ALL. From a purely semantic point of view, ALL would have been a good reference level in view of the fact that its various readings (distributive, cumulative, and collective) seem to be equally prominent / available by default. But there is a systematic, non-semantic difference between the experimental items in the ALL & INVERSE-SCOPE condition and all the other 7 conditions. Consider, for example, the word *language(s)* in (11a) and (11b) above. This word is of primary interest to us: it is the first main ROI that we want to examine. This ROI immediately follows EACH, EVERY and THE in the INVERSE-SCOPE order, and also *same* in the SURFACE-SCOPE order. In contrast, this ROI is separated from ALL in the INVERSE-SCOPE order by one word (namely *the*).

While this difference might prove to be relatively inconsequential, it would introduce a possible confound if ALL was selected as the reference level and consequently, every other quantifier was compared to it.

Following one of the recommendations in Barr et al. 2013, all our models included the maximal random effect structure for subject and items justified by our data. More precisely, we did backward model selection for random effect structures and we report here the model with the maximum random-effect structure that converges and that has the smallest deviance.[3] The confidence intervals (CIs) for all the models reported here are profile-likelihood CIs if the profile-likelihood CIs could be computed with the `lme4` package[4] without errors, but possibly with warnings; if the profile-likelihood CIs could not be computed without errors, we report the (less reliable) Wald CIs.

Table 1 displays the coefficients of the linear mixed-effects regression models for all the six words / ROIs.[5] Recall that the response variable is residualized log RTs (word length and word position have already been factored out), so the intercept coefficients can be negative. The other coefficients can also be negative because they represent differences relative to the intercept, which is the mean residualized log RT for the reference cell EVERY & SURFACE. All the coefficients whose 95% CIs exclude 0 (i.e., all those that are statistically significant at the $\alpha = 0.05$ level) are boldfaced.

The first important observation is that there is no slowdown associated with INVERSE-SCOPE in any of the ROIs, that is, no processing difficulties of INVERSE-SCOPE are detected. This is summarized in (13) below:

(13)   *Generalization 1*   INVERSE-SCOPE is not inherently slower / more difficult than SURFACE-SCOPE.

---

3 The maximal main-effects models that converge are not nested with respect to their random-effect structures, so we do not select based on Likelihood Ratio tests but rather by simply examining their deviance.

4 See Bates et al. (2013) and R Core Team (2013).

5 The models whose coefficients are reported in Table 1 had the following random-effect structure:

> Word 2: INTERCEPT + ORDER for subjects and items

> Word 3: INTERCEPT + ORDER for subjects; INTERCEPT + ORDER + QUANTIFIER TYPE for items

> Word 5: INTERCEPT + QUANTIFIER TYPE for subjects; INTERCEPT + ORDER for items

> Word 6: INTERCEPT + ORDER + QUANTIFIER TYPE for subjects; INTERCEPT for items

> Word 7&8: INTERCEPT + ORDER for subjects; INTERCEPT + QUANTIFIER TYPE for items

Exp. 1 (in context)	Word 2 [researcher/s]	Word 3 [learned]	Word 5 [language/s]	Word 6 [spoken]	Word 7 [in]	Word 8 [the]
INTERCEPT (EVERY&SURFACE)	$-0.007(-0.04, 0.03)$	$-0.002(-0.04, 0.04)$	$-0.03(-0.07, 0.01)$	$\mathbf{-0.08(-0.12, -0.04)}$	$\mathbf{-0.07(-0.12, -0.03)}$	$-0.04(-0.10, 0.02)$
ALL	$-0.02(-0.06, 0.01)$	$-0.02(-0.07, 0.02)$	$\mathbf{-0.06(-0.11, -0.01)}$	$-0.004(-0.05, 0.04)$	$-0.02(-0.08, 0.03)$	$-0.03(-0.08, 0.03)$
EACH	$-0.02(-0.06, 0.01)$	$0.006(-0.04, 0.05)$	$0.03(-0.01, 0.08)$	$\mathbf{0.04(0.002, 0.08)}$	$0.04(-0.11, 0.19)$	$0.02(-0.03, 0.07)$
THE	$-0.03(-0.07, 0.004)$	$-0.03(-0.07, 0.01)$	$-0.01(-0.06, 0.03)$	$\mathbf{0.07(0.02, 0.13)}$	$0.04(-0.10, 0.18)$	$\mathbf{0.06(0.01, 0.12)}$
INVERSE	$0.01(-0.02, 0.05)$	$0.01(-0.02, 0.05)$	$0.01(-0.02, 0.04)$	$0.01(-0.02, 0.04)$	$0.02(-0.02, 0.05)$	$-0.007(-0.05, 0.04)$

Table 1   Experiment 1 (in context): Coefficients & 95% CIs for the linear mixed models of the 6 ROIs.

The second observation is that some of the licensors are slower / more difficult to process than others. In particular, EACH and THE are slower than EVERY and ALL in the late (object) regions, but crucially not in the early (subject) regions. This is summarized below:

(14)   *Generalization 2*   EACH and THE are slower / more difficult than EVERY and ALL in the object but not the subject regions.

As we already noted in the previous subsection, the fact that EACH and THE are slower in the late (object) regions, but not in the early (subject) regions indicates that they are not inherently more difficult to process than EVERY and ALL. The increased difficulty is associated with their semantic combination with sentence-internal *same*. We will propose an explanation for this issue and adduce independent evidence to support it in Section 5 below.

Finally, we observe that ALL is as fast as EVERY in all regions except in the Word 5 region, where it is even faster. But this facilitation is orthogonal to our experimental manipulation. In fact, we already observe it in Word 2 (i.e., in the early region counterpart of Word 5) which, as we already noted above, is the only region where the interaction model is better than the main-effects model according to an LR test. The interaction is significant in Word 2 because the ALL & SURFACE condition was significantly faster than the ALL & INVERSE condition, as well as the other quantifiers in the SURFACE condition. This is clearly visible in Figure 4, which plots the mean RTs and SEs for Word 2 for all 8 conditions: 4 licensors × 2 orders, SURFACE-SCOPE, plotted with continuous lines, and INVERSE-SCOPE, plotted with dot-dash lines.

The difference between ALL and the other quantifiers we observe in Word 2 is orthogonal to our concerns since the word preceding Word 2 in the SURFACE case is *the* (as in *all the researchers*) and the one preceding Word 2 in the INVERSE case is *same* (as in *the same researcher*). So the slowness of the latter simply reflects the increased processing difficulty associated with
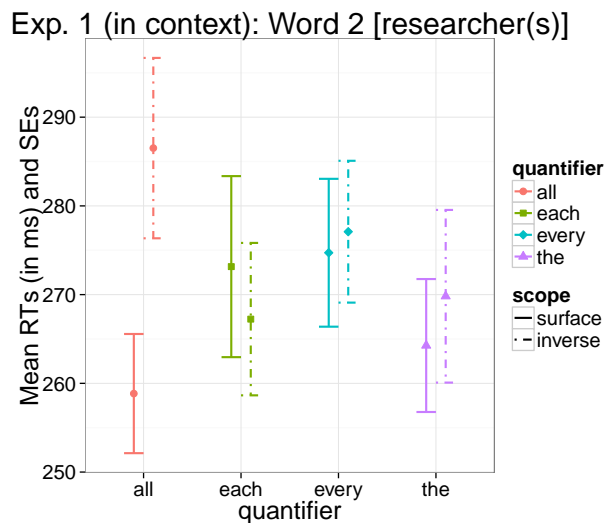
Figure 4   Mean RTs and SEs for Word 2 in Experiment 1 (in context).

anaphoric *same* in subject position, as well as a speed-up associated with *all* relative to the other quantifiers due to the fact that readers had an extra word, namely *the*, to process the meaning of the quantifier, while for EACH, EVERY and THE, Word 2 immediately follows the quantifier.[6]

The facilitation we observe with ALL in Word 5 seems to be the exact same issue as in Word 2. It might be due to the fact that ALL is an even better licensor of sentence-internal *same* than EVERY, but it is also possible that it is simply due to the fact that by the time readers reached Word 5, they had an extra word (namely *the*) to process ALL. Thus, the fact that we see the same kind of speed-up for ALL in the early Word 2 as well as in its late counterpart Word 5 is an indication that the speed-up is an inherent property of ALL rather than an effect of the semantic combination of ALL and sentence-internal *same*. Either way, this does not affect our main point about INVERSE VS. SURFACE scope, so we will not discuss this further.

---

6 On its own, the relative difficulty associated with *same* in subject position is not significant, and the same holds for the relative lack of difficulty associated with ALL. But together, these opposite effects produce a significant interaction.

3.2.2 The analysis of reading times for full sentences

We will now examine the RTs for full sentences by summing the residualized log RTs for every word in a sentence. This is an essential measurement in Anderson 2004, for example, which identified a slowdown for INVERSE-SCOPE only in the examination of complete sentences.

We remove 3 outlier observations out of a total of 1248 observations (see Baayen & Milin 2010 for more discussion of *a posteriori* trimming of reaction time data). Just as before, we fit a linear mixed-effects regression model to the resulting data set with fixed effects for quantifier type and order and intercept random effects for subjects and items. Again, the interaction of quantifier type and order does not significantly improve the data fit. The maximum likelihood estimates (MLEs) and associated 95% CIs are provided in Table 2:[7]

INTERCEPT (EVERY&SURFACE)	**−0.45**(**−0.80**, **−0.11**)
ALL	−0.08(−0.46, 0.30)
EACH	0.21(−0.16, 0.59)
THE	0.28(−0.09, 0.69)
INVERSE	0.05(−0.24, 0.34)

Table 2 Experiment 1 (in context): Coefficients & 95% CIs for the linear mixed model of the full-sentence (summed) residualized log RTs.

These results reinforce Generalizations 1 and 2 in (13) and (14) above: INVERSE and SURFACE scope seem to be indistinguishable (the 95% CI for INVERSE includes 0). The CIs of THE and EACH also include 0, and thus do not provide strong support for Generalization 2, but they are numerically slower than EVERY (and ALL) and the 95% CI for THE almost excludes 0.

3.2.3 The analysis of answer times and probabilities of giving correct answers

For completeness, we will also analyze the answer times and the pattern of (in)correct answers provided by participants. Recall that the last part (out of three) for every item and filler sequence was the same *yes/no* comprehension question which was basically asking whether the target sentence was true or

---

7 The model whose coefficients are reported in Table 2 had the following random-effect structure: INTERCEPT for subjects; INTERCEPT + ORDER + QUANTIFIER TYPE for items.

false relative to the background scenario (the items were balanced for truth and falsity).

We examine the (log transformed) data of the same 39 participants (answer times ≤ 50 ms or ≥ 10000 ms were trimmed). The best (minimal deviance) model with maximal random-effect structure that converged is the model with main effects for quantifier type and scope and without any interaction between quantifier type and scope (just as it was the case for RTs); in addition, the model has a main effect for (IN)CORRECT answers (reference level: CORRECT), and two-way interactions between (IN)CORRECT answers and quantifier type, as well as (IN)CORRECT answers and scope. The MLEs and associated 95% CIs are provided in Table 3:[8]

INTERCEPT (EVERY&SURFACE&CORRECT)	**7.31(7.23, 7.38)**
ALL	−0.0001(−0.04, 0.04)
EACH	0.01(−0.03, 0.06)
THE	**0.06(0.02, 0.10)**
INVERSE	0.02(−0.02, 0.05)
INCORRECT	0.25(−0.02, 0.51)
INCORRECT×ALL	−0.03(−0.28, 0.22)
INCORRECT×EACH	−0.06(−0.30, 0.18)
INCORRECT×THE	**−0.26(−0.48, −0.03)**
INCORRECT×INVERSE	0.09(−0.06, 0.25)

Table 3 Experiment 1: Coefficients & 95% CIs of the linear mixed model for log answer times.

Although most of the coefficients involving answer correctness in Table 3 are not significant, answer correctness is a significant predictor of answer time: the Likelihood Ratio test comparing the model reported in Table 3 and the model with the same random-effect structure but with answer correctness dropped as a fixed effect is significant ($p = 0.01, \chi^2 = 14.46, df = 5$). We see that INCORRECT answers take longer than CORRECT ones. This is as expected: lack of certainty about the answer should translate into extra processing time.

We also see that the processing difficulty associated with THE (but not EACH) is visible this late in the processing of the sentence and background scenario — when the given answer is CORRECT. However, when the given answer is INCORRECT, we actually notice a significant speed-up for THE

8 The model whose coefficients are reported in Table 3 had the following random effect structure: INTERCEPT + (IN)CORRECT ANSWER + ORDER for subjects and items.

relative to the other quantifiers. We think that this speed-up is due to the fact that THE is much harder than the other conditions and participants give up trying to find the right answer for stimuli occurring in this condition, hence the speed-up.

These effects are all on the log scale because we transformed the data to better satisfy the assumption of normality associated with linear mixed-effects models. But they can be visualized more easily if we examine the plot of the estimated answer times for the various conditions, provided in Figure 5. The top panel plots the estimated answer times for all 8 quantifier-type & order combinations when the answer was correct, while the bottom panel plots the answer times for the same 8 conditions when the answer was incorrect.
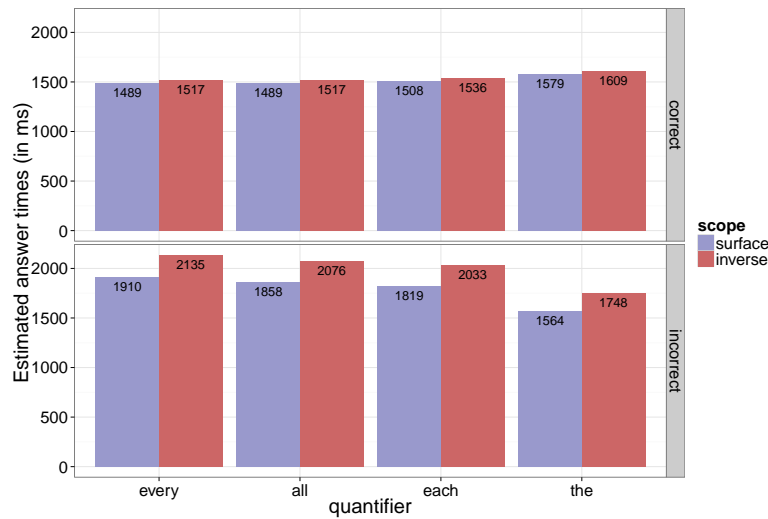
Figure 5 Experiment 1: Estimates for answer times based on the model in Table 3.

Both the plot and the model coefficients show that INVERSE-SCOPE seems to be slightly more difficult than SURFACE-SCOPE, but the difference does not reach significance (the 95% CI includes 0). This very late point in processing — which is after the point where sentence-internal *same* is licensed — is the first point where we actually see any indication that INVERSE-SCOPE might incur a processing cost.

The effect of INVERSE-SCOPE comes into sharper focus if we examine the pattern of CORRECT / INCORRECT answers conditional on our quantifier-type & scope experimental manipulation. Since the response variable (answer cor-

rectness) is binary categorical, we will use mixed-effects logistic regression models to analyze the data. Once again, the best model (according to Likelihood Ratio tests) is the one without interactions. The MLEs and associated p-values are listed in Table 4.[9]

INTERCEPT (EVERY&SURFACE)	$4.16(3.36, 4.96), p < 2 \times 10^{-16}$
ALL	$-0.32(-1.02, 0.37)$
EACH	$-0.47(-1.16, 0.21)$
THE	$-0.88(-1.53, -0.22), p = 0.008$
INVERSE	$-1.44(-2.02, -0.86), p = 1.2 \times 10^{-6}$

Table 4 Experiment 1: Coefficients, 95% CIs and p-values (for the significant coefficients) for probabilities of giving a correct answer (logit scale).

We see that THE is associated with a lower probability of a correct answer and most importantly, that INVERSE-SCOPE has a sizeable effect that is very highly significant. The model and resulting generalizations are easier to intuit if we examine the estimates on the probability scale rather than the logit scale — see Figure 6.
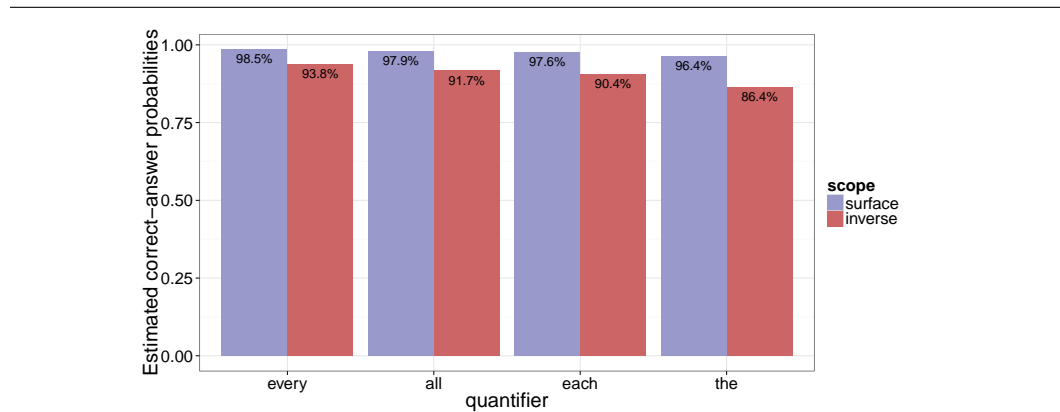
Figure 6 Experiment 1: Estimates for probabilities of giving a correct answer based on the model in Table 4.

We see there that INVERSE-SCOPE systematically decreases the chance of giving a CORRECT answer and this effect is most visible for THE and less so, for EACH.

An important consequence of the correlation between INVERSE scope and INCORRECT answers identified by this logistic regression model is that the

9 The model whose coefficients are reported in Table 4 had the following random effect structure: INTERCEPT + ORDER for subjects and items.

effect on answer times associated with INCORRECT answers that is observable in Table 3 and Figure 5 might be partly due to INVERSE scope. That is, the uncertainty associated with INCORRECT answers causes participants to answer more slowly, but part of that uncertainty might be due to the increased processing load associated with INVERSE scope. This is summarized in Generalization 3 below.

(15)   *Generalization 3*

    a.   INVERSE-SCOPE significantly reduces the probability of giving a CORRECT answer relative to SURFACE-SCOPE.

    b.   INCORRECT answers take significantly longer than CORRECT ones and this might be partly due to the increased processing load associated with INVERSE-SCOPE.

3.3   Interim summary

In this section, we examined three distinct points in the processing of quantifiers, scope, and the licensing of sentence-internal *same*.

The earliest regions we examined were Word 2 and Word 3. The results indicate that of the four quantificational licensors we considered, EACH, EVERY, and THE do not exhibit inherent processing differences. ALL, in contrast, was read faster in Word 2 in the surface-scope condition, which might be due to the fact that Word 2 immediately followed the three quantifiers, but was separated by the word *the* in the case of ALL. It is likely that the extra word that was present only in the case of ALL gave readers extra time to process the quantifier interpretation before Word 2 appeared.

The second set of regions we examined consisted of Word 5 through Word 8 (late regions, in object position). These regions show no effect of INVERSE-SCOPE, not even a numerical indication that INVERSE-SCOPE might be harder to process than SURFACE-SCOPE (see Generalization 1 in (13) above). But they do show an effect of quantifier type: EACH and THE are more difficult to process than EVERY and ALL (see Generalization 2 in (14) above). Given that we observe no difficulty with EACH and THE in the early (subject) regions, we conclude that these processing difficulties are due to the semantic combination of quantificational licensors and sentence-internal *same*. EACH and THE are not more difficult to process on their own; instead, it is the licensing requirement contributed by *same* that is more difficult to satisfy when EACH and THE license sentence-internal readings than when EVERY and ALL do.

Examining the full-sentence reading times yielded no statistically significant results but numerically, the full-sentence reading times provide additional support for Generalizations 1 and 2 above.

The third and final processing stage we examined was answer times and answer accuracy. The effect of THE was visible in both: THE caused increased answer latencies and also diminished answer accuracy. But we also observed a highly significant effect of INVERSE-SCOPE on answer accuracy: INVERSE-SCOPE reduced the probability of giving a CORRECT answer for all quantifier types.

Importantly, we also observed a significant effect of giving an INCORRECT answer: INCORRECT answers took longer than CORRECT answers. This is not surprising: we expect higher uncertainty about the answer to cause lengthier decision times. But part of that uncertainty might be due to the increased processing load associated with INVERSE scope (see Generalization 3 in (15) above).

Given the results we obtained by examining these three different temporal slices, namely:

 i. no effect in the early (subject) self-paced reading regions

 ii. an effect of quantifier only in the late (object) self-paced reading regions

 iii. effects for both quantifier and scope in the answer part

we might wonder how exactly to interpret the quantifier effect in the late self-paced reading regions (see (*ii*) above).

In particular, we might think at this point that the quantifier effect in (*ii*) is just a reflection of INVERSE-SCOPE: for some reason, we see the processing load associated with the INVERSE-SCOPE of EACH and THE before the answer stage, that is, already in the late self-paced regions. Or maybe there is an extra processing load associated only with the INVERSE-SCOPE of EACH and THE (over and above the load we see for all quantifiers in the answer part) that is visible in the late self-paced reading regions.

This cannot be an appropriate interpretation of the experimental result because it crucially relies on an *interaction* of quantifier type and order: it predicts that we should see effects of INVERSE-SCOPE only for EACH and THE. That is, it predicts that we should see a significant interaction between INVERSE and EACH / THE in the late regions—but no interactions were significant in these regions.

This is why we took the results in the late self-paced reading regions to be indicative of the relative difficulty of licensing sentence-internal *same* associated with EACH / THE relative to EVERY / ALL. This difficulty surfaces irrespective of whether the quantifier needs to take INVERSE scope to license *same* or can do the licensing from its SURFACE position.

To reiterate, the contrast between the lack of quantifier effects in the early self-paced reading regions and the presence of such effects in the late regions is an argument that the effects in the late regions are not due to the inherent processing difficulty associated with EACH / THE (relative to EVERY / ALL) but are a consequence of the semantic combination of EACH / THE and the licensing requirement contributed by sentence-internal *same*.

Additional supporting evidence for this hypothesis is provided by the contrast between the study we just discussed (Experiment 1, sentences in context) and a minimally different study in which the same sentences were presented out of context and to which we will henceforth refer as Experiment 2 (no context). The following section will briefly discuss Experiment 2 and will show that the late-region quantifier effects completely disappear: the results in both the early and the late regions are null. We will take this as an indication that processing out of context is not deep enough to get at semantic effects that involve longer distance composition and integration of semantic representations of the kind needed to license sentence-internal *same*.

4   The second self-paced reading experiment

4.1   Method, materials, procedure, and participants

The method, materials, and procedure for Experiment 2 were very similar to Experiment 1. The experimental manipulation (4 quantifiers $\times$ 2 orders) and the 32 self-paced reading target sentences were identical. The only difference was that the sentences were read out of context.

There were 62 participants in this experiment, all of them undergraduate students from UCSC. They completed the experiment online on a UCSC hosted installation of the IBEX platform (http://code.google.com/p/webspr/) for course credit or extra-credit. Experiment 1 and Experiment 2 were administered in two different quarters at a 2-3 month time interval.[10]

---

10 Given the fact that the participant pool is largely renewed every quarter, the chance of common participants was very small, so no explicit effort was undertaken to identify if any

Just as in Experiment 1, 8 lists were created following a Latin square design (in each list, every item appeared only in one condition). Every participant in Experiment 2 responded to 135 stimuli (32 experimental items + 103 fillers[11]), the order of which was randomized for every participant (any two experimental items were separated by at least one filler).

The fillers consisted of one sentence, just as the experimental items. Furthermore, their structure was similar to the structure of the experimental items (transitives with adjunct modifiers). Around 5 fillers included *all the* or *each*, and 11 fillers included *every*.

Some of the experimental items and fillers were followed by *yes/no* comprehension questions. The total number of comprehension questions was 61, 16 of which were associated with experimental items. The comprehension questions tested whether participants paid attention to the sentences they read (e.g., the sentence *During the circus show, every clown slipped on the same banana peel on the floor* was followed by the question *Was there a banana peel on the floor in the circus show?*). 6 experimental items tested whether participants interpreted *same* (e.g., *Last night, every nurse comforted the same patient in the emergency room* was followed by the question *Did every nurse comfort a different patient?*).

8 outlier participants were excluded because of their low answer accuracy (more than 17% incorrect answers out of a total of 61). The final number of participants: 54. These participants had, on average, 82% correct answers to the questions asking about the interpretation of *same*; this is basically 1 mistake in 6 answers. Just as before, all responses ≤ 50 ms and ≥ 2000 ms were removed and the remaining observations were log transformed to mitigate the right-skewness characteristic of reading-time data.

4.2 Data analysis and resulting generalizations

Again, we factored out the influence of word length and word position by running a linear mixed-effects regression. The regression had intercept-only random effects for subjects and two fixed effects — word length in characters

---

participants took part in both experiments. Even if that happened, the significant time lag between the two experiments and the fact that participants usually take several experiments every quarter make priming effects a very unlikely possibility.

11 The reason for the smaller number of fillers in the in-context experiment relative to the out-of-context one is that the in-context experiment would have been excessively long had we included 103 fillers, each with its own background scenario, as we did for the out-of-context experiment.

and word position in the sentence. The resulting residualized log reading times (log RTs) were used for all subsequent analyses.

The 6 ROIs were the same as the ones in Experiment 1. Figures 7 and 8 plot the mean RTs and the associated SEs for all these ROIs.

Again, there is no difference between the SURFACE and the INVERSE scope conditions. But we see an overall upward shift associated with both conditions in Experiment 2 relative to Experiment 1. This is as expected: there was no preceding context in Experiment 2, so the words in the target sentence were less predictable than in Experiment 1, which leads to overall higher RTs.

4.2.1 The statistical analysis of the six ROIs

Just as for Experiment 1, we analyze each word / ROI by means of a linear mixed-effects regression model. The fixed effects are the ones associated with our experimental manipulation: quantifier type (EVERY as the reference level vs. ALL, EACH, and THE) and order (SURFACE-SCOPE as the reference level vs. INVERSE-SCOPE).

Table 5 displays the coefficients of the linear mixed-effects regression models for all the six words / ROIs. We only report the models with main effects since these were the models with the best balance between parsimony and data fit according to LR tests comparing models with quantifier type $\times$ order interactions and models without interactions (with main effects only). Adding interactions did not reduce deviance significantly (at the usual $\alpha = 0.05$ level) except in 2 regions, Word 2 ($p = 0.01, \chi^2 = 11.2, df = 3$) and Word 5 ($p = 0.04, \chi^2 = 8.16, df = 3$). We return to this issue in detail below.[12]

The most important thing to note about the results summarized in Table 5 is that the effects of EACH and THE we observed in Experiment 1 in Word 6 and

---

12 The models whose coefficients are reported in Table 5 had the following random-effect structure:

> Word 2: INTERCEPT for subjects; INTERCEPT + ORDER + QUANTIFIER TYPE for items
>
> Word 3: INTERCEPT + QUANTIFIER TYPE for subjects; INTERCEPT + ORDER for items
>
> Word 5: INTERCEPT + ORDER for subjects; INTERCEPT + QUANTIFIER TYPE for items
>
> Word 6: INTERCEPT + ORDER + QUANTIFIER TYPE for subjects; INTERCEPT for items
>
> Word 7: INTERCEPT + ORDER for subjects; INTERCEPT + QUANTIFIER TYPE for items
>
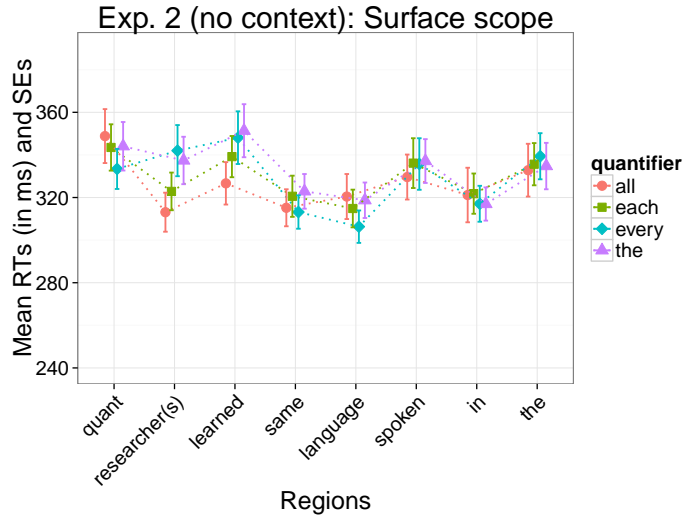> Word 8: INTERCEPT + ORDER + QUANTIFIER TYPE for subjects; INTERCEPT + ORDER for items

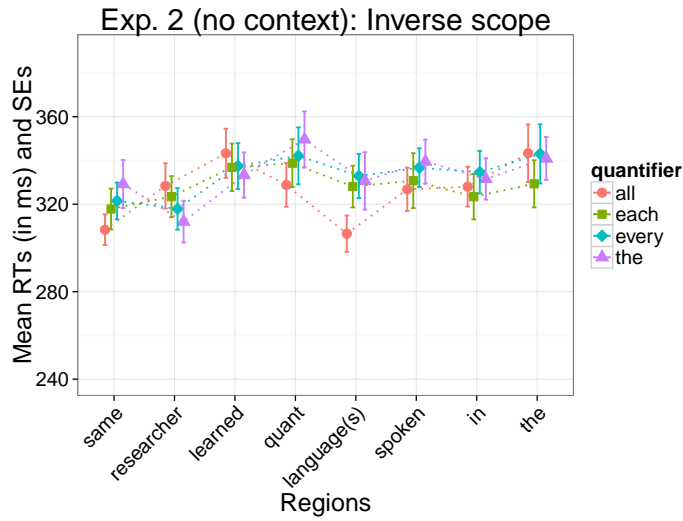**Figure 7**   Mean RTs and SEs for all ROIs in Experiment 2 (no context); surface scope only.

Figure 8   Mean RTs and SEs for all ROIs in Experiment 2 (no context); inverse scope only.

Same and its licensors

Exp. 2 (no context)	Word 2 [researcher/s]	Word 3 [learned]	Word 5 [language/s]	Word 6 [spoken]	Word 7 [in]	Word 8 [the]
INTERCEPT (EVERY&SURFACE)	$-0.02(-0.06, 0.02)$	$-0.005(-0.04, 0.03)$	$\mathbf{-0.06(-0.09, -0.03)}$	$0.003(-0.03, 0.04)$	$-0.02(-0.05, 0.007)$	$0.006(-0.04, 0.05)$
ALL	$-0.02(-0.07, 0.02)$	$-0.015(-0.06, 0.03)$	$-0.02(-0.06, 0.01)$	$-0.02(-0.09, 0.06)$	$-0.02(-0.09, 0.06)$	$-0.01(-0.05, 0.03)$
EACH	$-0.006(-0.04, 0.03)$	$-0.0002(-0.04, 0.04)$	$0.01(-0.03, 0.05)$	$-0.01(-0.05, 0.04)$	$-0.01(-0.07, 0.04)$	$-0.01(-0.05, 0.03)$
THE	$-0.01(-0.06, 0.03)$	$-0.003(-0.04, 0.04)$	$0.007(-0.03, 0.04)$	$0.01(-0.09, 0.11)$	$0.001(-0.05, 0.05)$	$0.004(-0.04, 0.05)$
INVERSE	$-0.02(-0.05, 0.01)$	$-0.004(-0.03, 0.02)$	$0.015(-0.01, 0.04)$	$0.002(-0.04, 0.04)$	$\mathbf{0.03(0.002, 0.06)}$	$0.001(-0.03, 0.03)$

Table 5   Experiment 2 (no context): Coefficients & 95% CIs for the linear mixed models of the 6 ROIs.

Word 8 are completely gone. This suggests that readers do not process *same* deeply enough to (fully) trigger its requirement that the sentence-internal reading needs to be licensed by an appropriate quantificational NP.

This is particularly interesting in view of the fact that we had a fairly large number of participants (54) whose accuracy on comprehension questions was high (at most 17% of the answers were incorrect). That is, participants paid attention to the task and actually read for comprehension. They also processed *same* to a certain extent, given that their responses to the 6 questions targeting it were answered correctly. However, such answers did not require establishing the proper licensing of sentence-internal readings at the point of reading sentences. This interpretation might have appeared only when answering comprehension questions required it, or it might have not appeared at all, since the questions could be answered by noticing the lexical mismatch, in particular, the contrast between *same* in test sentences and *different* in questions.

The phenomenon of properly licensing sentence-internal readings of *same* is a crucially compositional phenomenon: it requires the non-local combination / integration of the semantic representations contributed by both anaphoric *same* and the quantificational licensors. Thus, it seems that background scenarios and comprehension questions explicitly asking for the truth/falsity of the target sentence relative to the background scenario are necessary for participants to semantically process the target sentences deeply enough to reach the level where non-adjacent semantic representations are compositionally integrated.

In addition to this methodological point, the lack of EACH / THE effects in the late regions of Experiment 2 increases our confidence that the effects we observed in the late regions of Experiment 1 are really due to the semantic combination of sentence-internal *same* and its quantificational licensors. That is, Experiment 1 participants processed the target sentence deeply enough

1:29

to trigger the licensing requirement associated with *same*, attempted to satisfy it, and in the process of satisfying it, assigned INVERSE scope to the quantificational licensors as needed.

Let us turn now to the discussion of the significant effect of INVERSE-SCOPE we see in Word 7 in Table 5 above, and also the fact that the interaction models are better than the main-effects models in Word 2 and Word 5. Consider first the data summary for Word 2, provided in Figure 9; just as before, SURFACE-SCOPE is plotted with continuous lines and INVERSE-SCOPE is plotted with dot-dash lines.
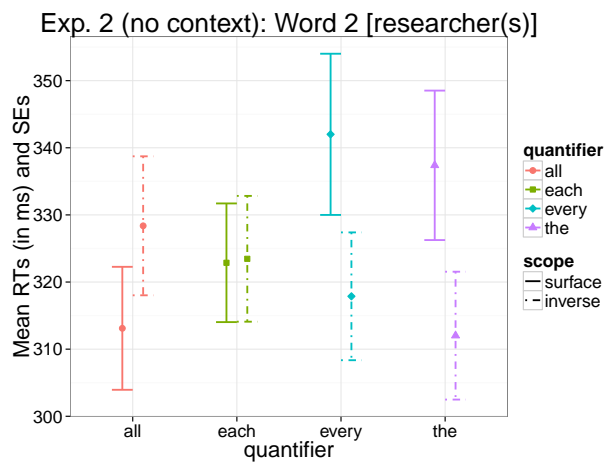
Figure 9   Mean RTs and SEs for Word 2 in Experiment 2 (no context).

We see that the interaction model is better for the Word 2 region primarily because EVERY takes significantly more time in SURFACE-SCOPE than ALL and EACH. We observe the same effect for THE, but this is less surprising given the anaphoric nature of the definite article.

The fact that we observe this effect in an early region indicates that it is inherent to EVERY and is unrelated to our experimental manipulation. We have no explanation for this except to suggest that it might be due to the fact that inadvertently, a higher number of fillers and associated comprehension questions featured EVERY. This might have prompted participants to flag this particular quantifier and pay more attention to the regions immediately following it.

Whether this is due to the fillers or not, the occurrence of this effect for EVERY in an early region strongly suggests that it is orthogonal to our experimental manipulation. And we think that the same effect causes the

interaction model to be better in the Word 5 region (which is the late counterpart of the Word 2 region), as well as the occurrence of significant main effects in Word 5 and Word 7. To see this, consider the data summaries for Word 5 and Word 7 in Figure 10: the only effect they exhibit is the one associated with EVERY that we observed in the early Word 2 region; even the effect associated with THE in Word 2 is mitigated in these late regions.
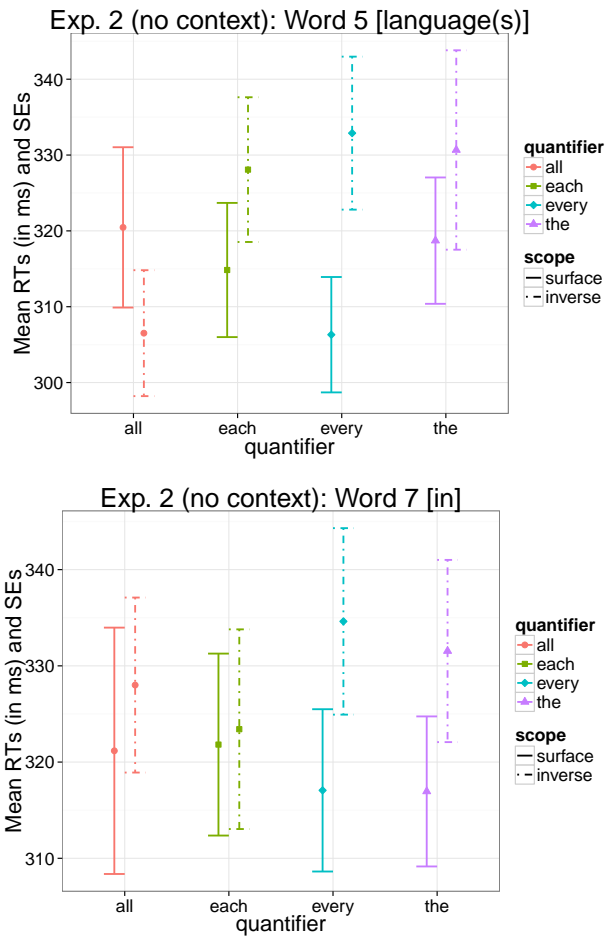

Figure 10   Mean RTs and SEs for Word 5 and Word 7 in Experiment 2 (no context).

We therefore conclude that the significant interactions in Word 2 and Word 5 as well as the INVERSE-SCOPE effect in Word 7 are not consequences of our experimental manipulation.

4.2.2 The analysis of reading times for full sentences

We also analyze the RTs for full sentences by summing the residualized log RTs for every word in a sentence. We fit a linear mixed-effects regression model to this data with fixed effects for quantifier type and order, and the maximal random-effect structure justified by our data (just as before, we retain only convergent models and we select the main-effects-only model with the smallest deviance). Again, the interaction of quantifier type and order does not significantly improve the data fit. The maximum likelihood estimates (MLEs) and associated 95% CIs are provided in Table 6.[13]

INTERCEPT (EVERY&SURFACE)	$-0.01(-0.28, 0.26)$
ALL	$-0.12(-0.53, 0.29)$
EACH	$-0.02(-0.39, 0.35)$
THE	$0.12(-0.25, 0.49)$
INVERSE	$0.10(-0.20, 0.39)$

Table 6 Experiment 2 (no context): Coefficients & 95% CIs for the linear mixed model of the full-sentence (summed) residualized log RTs.

The across-the-board null effects are not even numerically suggestive, and they reinforce the observation that readers did not process the target sentences deeply enough in this out-of-context task to reach the compositional integration of semantic representations that we were targeting.

5 Accounting for the self-paced reading generalizations

We turn now to our account of the three generalizations in (13), (14), and (15) above.

5.1 Generalization 2: EACH and THE are slower than EVERY and ALL

We begin with Generalization 2 (14): EACH and THE are slower than EVERY and ALL in the late (object) but not the early (subject) regions.

We interpreted this as an indication that participants actually process the sentence-internal requirement contributed by *same* and look for a semantically-plural quantificational NP to license it. But as the acceptability study in

---

13 The model whose coefficients are reported in Table 6 had the following random effect structure: INTERCEPT for subjects and INTERCEPT + ORDER + QUANTIFIER TYPE for items.

Brasoveanu & Dotlačil 2012 shows, not all licensors of sentence-internal *same* — or sentence-internal *different* or *similar*, for that matter — are born equal. Some licensors of *same* are better than others, in particular, ALL is a better licensor than THE or EACH.

A plot of the acceptability judgments reported in Brasoveanu & Dotlačil 2012 is provided in Figure 11. We see that EACH is judged as a significantly worse licensor of sentence-internal *same* than ALL.
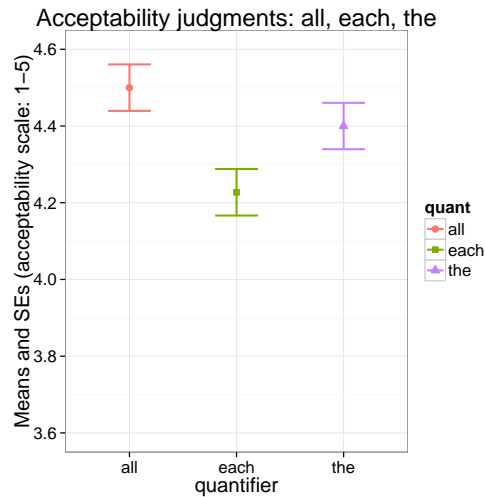
Figure 11 Mean acceptability and SEs for ALL, EACH and THE based on Brasoveanu & Dotlačil 2012 (acceptability scale: 1 (worst), 2, 3, 4, 5 (best))

EACH has been argued in the previous literature to require event differentiation in its scope (Tunstall 1998). In Tunstall's terms, this means that each object in the restrictor set of *each* is associated with its own subevent, and the subevent should be clearly distinguishable from the other subevents. One way to distinguish subevents is to assume that they occurred at different time points or different locations. Alternatively, if other entities appear in the subevents, it suffices to assume that these entities differ from each other. The latter way of satisfying the event differentiation requirement explains why we have a very strong preference for associating different researchers with different languages when we interpret sentences like (16) below (see Anderson 2004, Roeper, Pearson & Grace 2011 for experimental evidence).

(16)   Each researcher learned a language.

The event differentiation requirement contributed by *each* can also explain why the quantifier is a dispreferred licensor of sentence-internal *same*: licensing *same*—as in (17) below—goes against the default tendency to establish event differentiation in terms of a direct object with varying dependent reference. Of course, event differentiation can still be satisfied in (17) but it requires the reader to infer something that was not supplied by the sentence or the background scenario—for example, that each researcher is associated with a subevent whose temporal trace or location is different from the temporal traces or locations of the other subevents.

(17)   Each researcher learned the same language.

Due to the incompatibility of *same* with event differentiation and the necessary extra inference, we expect EACH to take more time than ALL in the late self-paced reading regions, but not in the early ones—which is exactly what Generalization 2 states.

But Generalization 2 also states that there was a slowdown for THE. The acceptability study in Brasoveanu & Dotlačil 2012 does not predict that. In fact, that study did not find a significant difference between ALL and THE, even though there was a numerical tendency along the lines of our Experiment 1 findings (i.e., the acceptability of THE was numerically worse than the acceptability of ALL). We see that, as is common, a real-time task can make subtler distinctions than an off-line (acceptability judgment) task and can uncover distinctions that would otherwise remain hidden.

We submit that this part of Generalization 2 follows from the fact that THE can appear with many readings—collective, cumulative, and distributive—but not all readings are equally acceptable. In particular, THE prefers a collective interpretation over a distributive one (Dotlačil 2010, Pagliarini, Fiorin & Dotlačil 2012). Collective readings, however, cannot license the sentence-internal reading of *same*. To see this, consider (18) below, which is infelicitous because of the collective reading required by *elect*.

(18)   The students elected {Harry / # the same representative}

The incompatibility of collective readings with *same* requires one to consider the dispreferred, distributive interpretation of *the*. This is possible but it is costly, and should lead to increased latencies during reading, as Generalization 2 confirms.

5.2   Generalization 1: No difference between INVERSE and SURFACE scope

Generalization 1 states that INVERSE-SCOPE is not inherently slower / more difficult than SURFACE-SCOPE in the self-paced reading part of the task: we observed no systematic effect of INVERSE scope. We take this generalization to support a model-structure account of the cost of inverse scope (Fodor 1982) rather than an account that assigns processing cost to LF-related operations, be they quantifier raising (QR) or covert type-shifting (Anderson 2004).[14]

The fact that we see no difference in reading times between INVERSE and SURFACE scope indicates that there is no processing cost associated with covert scoping operations. This is expected if the processing cost associated with inverse scope could only arise because of model structure revision: no revision takes place in our experimental items.

We note that another possible explanation for these findings is that AOCs somehow do not cause processing difficulties when requiring inverse scope. More concretely, one might suggest that covert operations are not always costly, they are only costly when their application is optional, driven by interpretive reasons rather than the grammatical system of the language under investigation. This would explain why there is a processing cost associated with universals scoping over indefinites, for example, *A boy climbed every tree*, since in this case the inverse scope operation is fully optional — the sentence is acceptable whether the inverse-scope operation applies or not.

In contrast, in our experiments, the inverse-scope operation was always required to license *same* in subject position. The hypothesis that inverse scope is not costly *per se* but only when it competes with a simpler strategy (surface scope) would therefore explain why no processing cost is observed when the simpler strategy becomes unavailable.[15] However, this interpretation of our results (compatible with the theory in Reinhart 2006, among others)

---

14 Brasoveanu (2011) and Brasoveanu & Dotlačil (2012) hypothesize that there are two ways of licensing sentence-internal *same*: one requires the licensor to scope over *same*, while the other merely requires the licensor to be semantically plural (no scoping is needed). The second licensing route is related to the analysis of plural *different* in Beck 2000. It is worth noting that postulating a licensing ambiguity along these lines does not alter the interpretation of our findings related to scope. This is because the latter interpretation of *same*, that is, the one that does not require the licensor to scope over *same*, is only compatible with non-distributive quantifiers. Thus, inverse scope is still necessary at least for *every* and *each* licensors, and no slowdown was observed with these quantifiers in the INVERSE-SCOPE condition.

15 We are grateful to Eytan Zweig for emphasizing this alternative explanation and for related discussion.

does not seem to be the right one given that Anderson (2004) found that sentences with sentence-internal *different* led to a slowdown in INVERSE-SCOPE despite the fact that inverse scope was the only option to license the sentence-internal reading of *different*, just as it was for *same* in our studies. But the difference between our results with *same* and Anderson's results with *different* is predicted by the hypothesis that model structure revision, not inverse-scope taking, is costly: *different*, unlike *same*, leads to changes in discourse model structure when its quantificational licensors are forced to take inverse scope.

Another possibility is that our experiments did not have enough power to detect the effect of inverse scope, especially if this effect is small. To determine if this is the case, we ran a power analysis with our experimental setup and the magnitude of the effect of inverse scope reported in Anderson 2004. More precisely, we considered the difference between inverse and surface scope reported in Anderson's Experiment 7 (see Anderson 2004: 75–76, and Figure 7 in particular). Experiment 7 investigated the difference between inverse and surface scope for universal quantifiers in two distinct conditions. In the ambiguous condition, the universal took surface or inverse scope relative to an ordinary indefinite in subject position, for example, *A climber scaled every cliff*. In the non-ambiguous condition, the subject contained the sentence-internal AOC *different*, for example, *A different climber scaled every cliff*, for the inverse scope case. Thus, there were two reported differences between inverse and surface scope, one for non-ambiguous stimuli, and one for ambiguous stimuli. The latter is the smaller one and we used this one in our simulations. The reason is that if our power is enough to detect effects of this magnitude, it will most probably be enough to detect the other, larger effect.

The INVERSE − SURFACE difference for ambiguous stimuli is $351 - 97 = 254$ (see Anderson 2004: 76, Figure 7). The SE of the INVERSE-SCOPE mean seems to be around 120 and the SE of the SURFACE-SCOPE mean seems to be at most 90. Therefore, the SE of the INVERSE − SURFACE difference should be $\sqrt{120^2 + 90^2} = 150$.

In our simulation, we used the coefficients we obtained from a random-intercepts only model for our full-sentence residualized RTs.[16] But we re-

---

16 Anderson's residualized RTs were obtained in a partly different manner: the RTs were not log-transformed, and the only predictor was word length (word position in the sentence was not considered). Furthermore, the data was aggregated over participants (see Anderson 2004: 47). Therefore, for the purposes of this simulation, we changed our RT residualization

placed the INVERSE coefficient (i.e., the estimated INVERSE − SURFACE difference) with Anderson's estimate. The reasoning behind this is as follows. If our null effect is simply a consequence of the low power of our experimental setup, we expect data sets simulated based on Anderson's estimate for scope to also yield null effects — despite the fact that the effect for scope was highly significant in Anderson's case (see Anderson 2004: 75: $p < 0.001$). We simulated 20,000 data sets, each consisting of 1248 observations (39 subjects × 32 items).[17] For each of the 20,000 data sets, we estimated a mixed-effects model and we collected the t-value associated with the inverse scope effect. We obtained the percentage of significant effects by using a normal-distribution based cutoff point: $|t| \geq 1.96$. The percentage of significant, non-null results was 86.1%. This means that our experimental setup was powerful enough to detect effects of scope of the magnitude reported in Anderson 2004 more than 85% of the time (the usual threshold being 80%, see Cohen 1992 among others). Thus, it is highly unlikely that our null effects are a consequence of low power.

We also considered the possibility that Experiment 1 did not show any effect of inverse scope in reading times because the participants were able to predict the scope of the upcoming target sentence based on the structure of the background scenario. Because we wanted our background scenarios to introduce the minimal number of entities that would make the target sentences felicitous, we always introduced three entities for the set picked up by the universal quantifier and the definite (two entities would not have been enough since *both* would have probably been the most natural choice rather than *all / every / each*), and two entities for the set picked up by *same* (no need to go beyond two entities in this case). The participants might have been sensitive to this type of regularity. But to be able to predict the scope of the upcoming sentence, noticing this regularity would not have been enough: readers would also have to form correct expectations about the structure of the upcoming sentence (i.e., how the two sets will be related to the two NP

procedure to match Anderson's: we did not log-transform the data, and we used word length as the only predictor in the residualizing regression. We kept random intercepts for subjects to match the aggregation over participants. The resulting residualized RTs for our Experiment 1 (summed for the entire sentence, just as in Anderson — see Anderson 2004: 76, Figure 7) were very similar in magnitude to Anderson's.

17 The fixed effects, except for inverse scope, and the standard deviations for the random intercepts for subjects and items were the ones we estimated from our data. In each of the 20,000 iterations, the effect for inverse scope was a random draw from a normal distribution centered at 254 and with a variance of 150^2.

types). Although we found it unlikely that readers would be able to develop correct expectations on all these counts, we couldn't exclude the possibility that our background scenarios helped them form the right expectations.

To study this issue further, we analyzed the data provided by each participant in the first half of the experiment, since at that point participants could hardly form such detailed predictions about the upcoming target sentences. We wanted to see if the data from the first half of the experiment showed a significant slowdown for the INVERSE-SCOPE condition. If it did, the null effect observed for the whole experiment could have been due to the fact that the effect of INVERSE-SCOPE was washed away in the second half of the experiment because scope was predictable based on scenario structure. But the results from the first half of the experiment paint the same picture as the complete results: the INVERSE-SCOPE condition doesn't exhibit a significant slowdown in reading times, while the effect of ALL remained significant (Word 5) and the effect of THE was borderline significant (Word 6). The effect of EACH was not significant, but this is most likely due to the inevitable decrease in power that is a consequence of halving the number of observations (lower power is also the most probable reason for the weaker effect of THE). We therefore conclude that the structure of the background scenarios did not (inadvertently) wash away the effect of inverse scope. Rather, there was no effect of inverse scope to begin with.

5.3 Generalization 3: The effect of INVERSE scope on answer accuracy/latency

Generalization 3 states that there is a strong negative effect of INVERSE-SCOPE on the probability of giving a CORRECT answer and furthermore, that this effect might be part of the reason INCORRECT answers take significantly longer than CORRECT ones. How should we understand this effect of inverse scope on answer accuracy and answer times?

One possibility could be that participants construct mental models for sentences / discourses only to the extent they really need to for a particular task. The difference between narrow-scope sentence-internal *same* and narrow-scope sentence-internal *different* is already represented at a shallower level: we only need to mentally "index" one individual in the former case, but we need multiple individuals or even a function / dependency between individuals for the latter.

But narrow-scope sentence-internal *same* and narrow-scope sentence-internal *different* might be much more similar at a deeper level, that is, at the level of semantic processing needed to verify the truth/falsity of a sentence. This involves taking the discourse model we built for the sentence and "matching" it against a real-world background model. Whether we verify *The same researcher learned every language* or *A different researcher learned every language*, we need to go through the contextually-specified list of languages and somehow check whether their corresponding researchers are identical or distinct. Either way, this requires us to retrieve the list of languages first and match each of them against the researcher(s).

Crucially, the list of languages is less salient. It consists of inanimate entities, and it is mentioned by the quantifier in the less prominent (non-subject) position. Hence, it is likely that it is more difficult to retrieve this list when the target sentence involves INVERSE-SCOPE, and it is easier to do the same in case of SURFACE-SCOPE.

Note that this truth verification "procedure" is exactly what is encoded by embedding functions in Discourse Representation Theory (DRT; Kamp 1981 and Kamp & Reyle 1993): embedding functions relate Discourse Representations Structures (DRSs, i.e., mental discourse models) and the actual, "real-world" model. The shallower level of discourse model processing would thus correspond to constructing a DRS for the current sentence / discourse. The deeper level of discourse model processing would correspond to linking this DRS to a real-world background situation, that is, to constructing an embedding function (partial variable assignment) that verifies this DRS.

6 Conclusion

The paper presented novel evidence regarding the processing of inverse scope and the interpretation of sentence-internal *same* with four licensors (*all*, *each*, *every*, and *the*), collected in two self-paced reading studies. These real-time / on-line processing studies complement the results currently available in the formal semantics literature, which are exclusively based on offline acceptability judgments, whether formally or informally collected.

The two studies show that there is no general effect of surface vs. inverse scope, which we take as an argument for a model-oriented view of the processing cost of inverse scope: the inverse scope of quantifiers seems to be costly because of model structure reanalysis, not (only) because of covert scope operations. We also observed a slowdown for EACH and THE

relative to EVERY and ALL in the late self-paced reading regions, which we took as a argument for the lexically-specified differences between these licensors of sentence-internal readings. In particular, we argued that the lexical requirement of event differentiation contributed by EACH clashes with the meaning of *same*, and so does the fact that THE is preferably associated with a collective interpretation.

The second main result is methodological: the psycholinguistic investigation of semantic phenomena like the interaction of quantifiers and sentence-internal readings should always involve a context that prompts a deep enough processing of the target expressions. In one of our two studies, participants read the target sentences after reading a scenario introducing the two sets of entities that the quantifier NP and the *same* NP referred back to, and they were always asked to determine whether the sentence was true or false relative to the background scenario. In the other study, the participants read the same sentences without any context and there were fewer follow-up comprehension questions. The relevant effects observed in the in-context study completely disappeared in the out-of-context study, although the participants in both studies were monitored for their level of attention to the experimental task.

Finally, we conjectured that the effect of INVERSE-SCOPE we observed in answer accuracy and answer times is due to the fact that discourse models for sentences / discourses are processed at different levels of depth, depending on the particular task readers / interpreters attend to. This corresponds roughly to first constructing a DRS for a sentence / discourse (shallower semantic processing), and then constructing an embedding function that verifies this DRS by linking it to a background situation / model (a deeper level of semantic processing).

Experimental items

The same target sentences were used in both self-paced reading experiments. Since the first experiment also included background scenarios that were associated with these sentences, we provide here only the items used in the first experiment.

(1)   a. SURFACE-SCOPE

      i. *Scenario.* Three movie critics — Ramon, Sue, and Taylor — work for a journal in Boston. Last week, there were two new movies available for review, 'Blob 2' and 'Will she love me?'. Ramon reviewed 'Blob 2', Sue reviewed 'Blob 2', and Taylor reviewed 'Blob 2' as well.

      ii. *Sentence.* I think that all the critics / each critic / every critic / the critics reviewed the same movie for Boston magazine.

   b. INVERSE-SCOPE

      i. *Scenario.* Two movie critics — Ramon and Taylor — work for a journal in Boston. Last week, there were three new movies available for review, 'A scary cup', 'Horrible death 23' and 'Fire wheels'. Ramon reviewed 'A scary cup', 'Horrible death 23' and 'Fire wheels'. Taylor took a week off and reviewed nothing.

      ii. *Sentence.* I think that the same critic reviewed all the movies / each movie / every movie / the movies for Boston magazine.

(2)   a. SURFACE-SCOPE

      i. *Scenario.* Three professors were invited to the University of Pennsylvania to talk about one of two drama forms — pantomime or comedy — in the local lecture series organized by the Theater Department. The first professor talked about pantomimes. The second professor talked about comedies. The third professor talked about pantomimes.

      ii. *Sentence.* I think that all the speakers / each speaker / every speaker / the speakers discussed the same drama form in the Theater Department lecture series.

   b. INVERSE-SCOPE

      i. *Scenario.* Two professors were invited to the University of Pennsylvania to talk about at least one out of three drama forms — pantomime, comedy, or tragedy — in the local lecture series organized by the Theater Department. The first professor talked about pantomimes and comedies. The second professor talked about tragedies.

      ii. *Sentence.* I think that the same speaker discussed all the drama forms / each drama form / every drama form / the drama forms in the Theater Department lecture series.

(3)   a. SURFACE-SCOPE

      i. *Scenario.* Three children, Jiang, Kramer and Lopez, were asked to make a presentation about their favorite animal in science class. Jiang made a presentation about crocodiles. Kramer presented crocodiles and Lopez presented crocodiles too.

      ii. *Sentence.* I think that all the children / each child / every child / the children presented the same animal in science class.

   b. INVERSE-SCOPE

      i. *Scenario.* Two children, Jiang and Lopez, were asked to make a presentation about at least one out of three animal species — crocodiles, monkeys, or lions — in science class. Jiang made a presentation about crocodiles, monkeys and lions. Lopez presented nothing.

      ii. *Sentence.* I think that the same child presented all the animals / each animal / every animal / the animals in science class.

(4)   a. SURFACE-SCOPE

      i. *Scenario.* Three customers were in an appliance store right before it was about to close. The first customer, a middle-aged man, studied a toaster. The second customer, a young hipster, was examining the toaster too. The third customer, a young woman, closely studied an expensive water cooler.

      ii. *Sentence.* I think that all the customers / each customer / every customer / the customers closely examined the same appliance right before the store closed.

b. INVERSE-SCOPE

    i. *Scenario.* Two customers were in an appliance store right before it was about to close. The first customer, a middle-aged man, studied a toaster and a pasta maker. The second customer, a young woman, was examining an expensive water cooker.

    ii. *Sentence.* I think that the same customer closely examined all the appliances / each appliance / every appliance / the appliances right before the store closed.

(5) a. SURFACE-SCOPE

    i. *Scenario.* Three young women — Sarah, Sue and Madeleine — live in a village that has only two shops, a bakery and a small supermarket. Just before lunch time, Sarah went to the bakery, Sue went to the bakery and Madeleine went to the bakery as well.

    ii. *Sentence.* I think that all the women / each woman / every woman / the women visited the same shop in the village before lunch time.

b. INVERSE-SCOPE

    i. *Scenario.* Two young women — Sarah and Madeleine — live in a village that has only three shops, a fabric store, a bakery and a small supermarket. Last Monday just before lunch time, Sarah went to the fabric store, then to the bakery and finally to the small supermarket, while Madeleine stayed home.

    ii. *Sentence.* I think that the same woman visited all the shops / each shop / every shop / the shops in the village before lunch time.

(6) a. SURFACE-SCOPE

    i. *Scenario.* Grant, Allen and Jack are three students that study at Emporia school. Last September, they went to a store that had two discounted computers, an ACER and a Mac, during the Monday special-deal period. Grant tested the ACER, Allen tested the Mac and Jack tested the ACER.

    ii. *Sentence.* I think that all the students / each student / every student / the students tested the same laptop during the Monday special-deal period.

b. INVERSE-SCOPE

    i. *Scenario.* Grant and Allen are two students that study at Emporia school. Last September, they went to a store that had three discounted computers, an ACER, an IBM and a Mac, during the Monday special-deal period. Grant tested the ACER and the Mac. Allen tested the IBM.

    ii. *Sentence.* I think that the same student tested all the laptops / each laptop / every laptop / the laptops during the Monday special-deal period.

(7) a. SURFACE-SCOPE

    i. *Scenario.* To prepare for fieldwork, three researchers — a botanist, a linguist and an anthropologist — had to learn one of two languages spoken in the eastern Indonesian islands — Bahasa Indonesia or Ternate. The botanist learned Bahasa Indonesia, the linguist learned Bahasa Indonesia and the anthropologist learned Bahasa Indonesia too.

    ii. *Sentence.* I think that all the researchers / each researcher / every researcher / the researchers learned the same language spoken in the eastern Indonesian islands.

b. INVERSE-SCOPE

    i. *Scenario.* To prepare for fieldwork, two researchers — a botanist and an anthropologist — had to learn at least one out of three languages spoken in the eastern Indonesian islands — Bahasa Indonesia, Ternate or Tidore. The botanist learned Bahasa Indonesia, Ternate and Tidore. The anthropologist learned nothing and used the botanist as his guide and advisor.

    ii. *Sentence.* I think that the same researcher learned all the languages / each language / every language / the languages spoken in the eastern Indonesian islands.

(8) a. SURFACE-SCOPE

    i. *Scenario.* Bob, Bill and David are three American tourists travelling in southern Greece. Bob visited Crete, an island in that region. Bill visited Crete too. David visited Milos, another island.

    ii. *Sentence.* I think that all the tourists / each tourist / every tourist / the tourists visited the same island in southern Greece.

b. INVERSE-SCOPE

    i. *Scenario.* Bob and David are two American tourists trying to visit three islands in southern Greece — Crete, Milos and Rhodos. Bob visited Crete and nothing else. David visited Milos and Rhodos but did not manage to visit Crete.

    ii. *Sentence.* I think that the same tourist visited all the islands / each island / every island / the islands in southern Greece.

(9) a. SURFACE-SCOPE

    i. *Scenario.* Three cello pieces by Vivaldi, Cello concerto in C major, G major and D major, were supposed to be played in one of two San Francisco concert halls last month — either in the Conservatory of Music Hall or in the Davies Symphony Hall. In the end, the three cello pieces were played in the Davies Hall, while the Conservatory of Music Hall hosted music by Bach.

    ii. *Sentence.* I think that all the cello pieces / each cello piece / every cello piece / the cello pieces by Vivaldi was / were played in the same concert hall in San Francisco.

b. INVERSE-SCOPE

    i. *Scenario.* Two cello pieces by Vivaldi, Cello concerto in C major and D major, were supposed to be played in at least one out of three San Francisco concert halls last month — in the Conservatory of Music Hall, in the Davies Symphony Hall, or in the Herbst Theatre Hall. The concerto in C major was first played in the Conservatory of Music Hall. Later, it was played in the Davies Hall and after that, it was played in the Herbst Theatre Hall. The other concerto was never played in San Francisco.

    ii. *Sentence.* I think that the same cello piece by Vivaldi was played in all the concert halls / each concert hall / every concert hall / the concert halls in San Francisco.

(10) a. SURFACE-SCOPE

    i. *Scenario.* Three hunters, David, Ron and Mitt, went hunting in the tiny forest by the local lake. Two bears lived there, a brown bear and a grizzly bear. David saw the brown bear. Ron saw the brown bear too. Mitt, however, saw the grizzly bear.

    ii. *Sentence.* I think that all the hunters / each hunter / every hunter / the hunters saw the same bear living in the tiny forest by the lake.

b. INVERSE-SCOPE

    i. *Scenario.* Two hunters, David and Mitt, went hunting in the tiny forest by the local lake. Three bears lived there, a brown bear, a black bear and a grizzly bear. David saw the brown bear and the black bear. Mitt only saw the grizzly bear.

    ii. *Sentence.* I think that the same hunter saw all the bears / each bear / every bear / the bears living in the tiny forest by the lake.

(11) a. SURFACE-SCOPE

    i. *Scenario.* Three millionaires, Andrew, Lisa and William, wanted to sail their boats past two islands in the Caribbean sea, Aruba and Curacao. Andrew sailed only past Aruba. So did Lisa. William sailed past Aruba as well. None of them managed to sail to Curacao.

    ii. *Sentence.* I think that all the boats / each boat / every boat / the boats sailed past the same island in the Caribbean sea.

b. INVERSE-SCOPE

    i. *Scenario.* Two millionaires, Andrew and William, wanted to sail their boats past the islands in the Caribbean sea. Andrew managed to do that. But William had to stay home because his boat had a serious engine problem.

    ii. *Sentence.* I think that the same boat sailed past all the islands / each island / every island / the islands in the Caribbean sea.

(12) a. SURFACE-SCOPE

    i. *Scenario.* Three detectives working in Cherryville, Jane, Nick and Philip, were working on two different cases last year: a homicide and a house break-in. Independently of each other, Jane and Nick solved the homicide. Philip solved the house break-in.

ii. *Sentence.* I think that all the detectives / each detective / every detective / the detectives solved the same crime in Cherryville last year.

b. INVERSE-SCOPE

i. *Scenario.* Two detectives working in Cherryville, Jane and Philip, were working on three different cases last year: a homicide, a house break-in and a fraud. Jane solved the homicide and the house break-in. Philip solved the fraud.

ii. *Sentence.* I think that the same detective solved all the crimes / each crime / every crime / the crimes in Cherryville last year.

(13)   a.   SURFACE-SCOPE

i. *Scenario.* Three maids, Susan, Natalie and Diana, had to dust two shelves in the library every day. Yesterday, Susan dusted the top shelf. Natalie dusted the top shelf too. Unbeknownst to them, Diana had dusted the top shelf that day too.

ii. *Sentence.* I think that all the maids / each maid / every maid / the maids dusted the same shelf in the library.

b.   INVERSE-SCOPE

i. *Scenario.* A library had ten shelves and two maids, Susan and Diana, had to dust them every day. Yesterday, Susan dusted the ten shelves. Diana skipped cleaning the library.

ii. *Sentence.* I think that the same maid dusted all the shelves / each shelf / every shelf / the shelves in the library.

(14)   a.   SURFACE-SCOPE

i. *Scenario.* A whiteface clown, an auguste and a pierrot had a show together. First, the whiteface clown came on stage and slipped on a banana peel. Then, the auguste came and slipped on the banana peel too. Finally, the pierrot came. He acted as if he noticed the banana peel but did not notice an orange peel lying next to it. He carefully walked around the banana peel, only to slip on the orange peel a few seconds later.

ii. *Sentence.* I think that all the clowns / each clown / every clown / the clowns slipped on the same banana peel on the floor.

b.   INVERSE-SCOPE

i. *Scenario.* A whiteface clown and a pierrot had a show together. First, the whiteface clown came on stage and slipped on two banana peels. Then, the pierrot came on stage. He acted as if he noticed the two banana peels but did not notice a third banana peel lying next to them. He carefully walked around the two banana peels, only to slip on the third one a few seconds later.

ii. *Sentence.* I think that the same clown slipped on all the banana peels / each banana peel / every banana peel / the banana peels on the floor.

(15)   a.   SURFACE-SCOPE

i. *Scenario.* Three copy editors, Gillian, Ian and Boris, had to correct mistakes in a manuscript that was about to be published. As usual, they read and made corrections independently of each other. Gillian found only one mistake, on page 20, and corrected it. Ian found and corrected the mistake on page 20 too. Boris found and corrected the mistake on page 20 as well.

ii. *Sentence.* I think that all the copy editors / each copy editor / every copy editor / the copy editors corrected the same mistake in the manuscript.

b.   INVERSE-SCOPE

i. *Scenario.* Two copy editors, Gillian and Boris, had to correct mistakes in a manuscript that was about to be published. As usual, they read and made corrections independently of each other. Gillian found one mistake on page 20, another one on page 25 and yet another one on page 30 and corrected them. Boris read the whole manuscript but did not find any mistakes.

ii. *Sentence.* I think that the same copy editor corrected all the mistakes / each mistake / every mistake / the mistakes in the manuscript.

(16) a. SURFACE-SCOPE

  i. *Scenario.* The best restaurant in Springville employs three cooks — Thomas, Brad and Terence. Yesterday, Thomas prepared ratatouille and tasted it afterwards. He wasn't sure about the seasoning, so Brad tasted the food as well. Meanwhile, Terence prepared and tasted lasagna.

  ii. *Sentence.* I think that all the cooks / each cook / every cook / the cooks tasted the same dish prepared in the restaurant.

  b. INVERSE-SCOPE

  i. *Scenario.* The best restaurant in Springville employs two cooks — Thomas and Terence. Yesterday, Thomas prepared ratatouille and tasted it afterwards. Meanwhile, Terence prepared and tasted lasagna and asparagus soup.

  ii. *Sentence.* I think that the same cook tasted all the dishes / each dish / every dish / the dishes prepared in the restaurant.

(17) a. SURFACE-SCOPE

  i. *Scenario.* Hugh, Karl and Ronald wanted to buy flowers for their wives. There were two flower shops in the town where they lived. One shop specialized in lilies and the other shop sold roses. Hugh went to the lily shop. Karl went to that shop too. And so did Ronald.

  ii. *Sentence.* I think that all the men / each man / every man / the men went to the same flower shop in town to buy flowers.

  b. INVERSE-SCOPE

  i. *Scenario.* Hugh and Ronald wanted to buy flowers for their wives. There were three flower shops in the town where they lived. One shop specialized in lilies, the other shop sold roses and the third shop sold tulips. Hugh went to the lily shop, to the rose shop and to the tulip shop. Ronald got sick and had to stay at home.

  ii. *Sentence.* I think that the same man went to all the flower shops / each flower shop / every flower shop / the flower shops in town to buy flowers.

(18) a. SURFACE-SCOPE

  i. *Scenario.* Three teenagers, Andy, Doug, and Jim, went to the local movie theater, which played Godzilla and Batman. Andy went to see Godzilla. Doug went to see Batman. Jim went to see Godzilla.

  ii. *Sentence.* I think that all the teenagers / each teenager / every teenager / the teenagers went to the same movie playing at the local movie theater.

  b. INVERSE-SCOPE

  i. *Scenario.* Two teenagers, Andy and Jim, went to the local movie theater, which played Godzilla, Batman and Spiderman. Andy went to see Godzilla and then he went to see Spiderman. Jim went to see Batman.

  ii. *Sentence.* I think that the same teenager went to all the movies / each movie / every movie / the movies playing at the local movie theater.

(19) a. SURFACE-SCOPE

  i. *Scenario.* An intern from New York University, an intern from Rutgers University and an intern from Princeton were getting bored at the office where they worked. They found two card games on their computers, hearts and poker. The intern from New York University played hearts. The intern from Rutgers University played hearts too, and so did the intern from Princeton.

  ii. *Sentence.* I think that all the interns / each intern / every intern / the interns played the same card game installed on the computers in the office.

  b. INVERSE-SCOPE

  i. *Scenario.* An intern from New York University and an intern from Princeton were getting bored at the office where they worked. They found three card games on their computers, hearts, solitaire and poker. The intern from New York University played hearts, solitaire and poker. The intern from Princeton was worried that someone might see him play at work so he decided not to play any of the games.

ii. *Sentence.* I think that the same intern played all the card games / each card game / every card game / the card games installed on the computers in the office.

(20) a. SURFACE-SCOPE

    i. *Scenario.* Three servants, Mr. Rice, Mr. Nowak and Mr. Hill, had to paint the chairs in the mansion's attic. Mr. Rice painted one chair. Mr. Nowak did not like the color and painted that chair over. Meanwhile, Mr. Hill painted the other chairs in the attic.

    ii. *Sentence.* I think that all the servants / each servant / every servant / the servants painted the same chair in the mansion's attic.

  b. INVERSE-SCOPE

    i. *Scenario.* Two servants, Mr. Rice and Mr. Hill, had to paint the chairs in the mansion's attic. Mr. Rice painted one chair. Meanwhile, Mr. Hill painted the other chairs in the attic.

    ii. *Sentence.* I think that the same servant painted all the chairs / each chair / every chair / the chairs in the mansion's attic.

(21) a. SURFACE-SCOPE

    i. *Scenario.* There were three rock bands in a town: Cracks, Horses Erased and Monster Lives Downtown. The town had two clubs which were open past midnight: Bond and Streetlight Limited. Cracks, Horses Erased and Monster Lives Downtown played in Bond. Streetlight Limited only hosted hip-hop bands.

    ii. *Sentence.* I think that all the rock bands / each rock band / every rock band / the rock bands played in the same club downtown that was open past midnight.

  b. INVERSE-SCOPE

    i. *Scenario.* There were two rock bands in a town: Cracks and Monster Lives Downtown. The town had three clubs which were open past midnight: Bond, Huge and Streetlight Limited. On Friday, Monster Lives Downtown played in Bond at 10pm. Then, the band played in Huge at midnight. Finally, it moved to Streetlight Limited. Cracks did not have any gigs that evening.

    ii. *Sentence.* I think that the same rock band played in all the clubs / each club / every club / the clubs downtown that was / were open past midnight.

(22) a. SURFACE-SCOPE

    i. *Scenario.* Three fashion models, Violetta, Patricia and Natalie, had to present two luxury dresses on the catwalk. One dress was laced with gold, the other dress was laced with silver. Violetta presented the gold-laced dress. Patricia presented the gold-laced dress too. Natalie presented the silver-laced dress.

    ii. *Sentence.* I think that all the models / each model / every model / the models presented the same luxury dress on the catwalk.

  b. INVERSE-SCOPE

    i. *Scenario.* Two fashion models, Violetta and Natalie, had to present three luxury dresses on the catwalk. One dress was laced with gold, one was laced with platinum and the last one was laced with silver. Violetta presented the gold-laced dress and the platinum-laced dress. Natalie presented the silver-laced dress.

    ii. *Sentence.* I think that the same model presented all the luxury dresses / each luxury dress / every luxury dress / the luxury dresses on the catwalk.

(23) a. SURFACE-SCOPE

    i. *Scenario.* Three kids, Jacob, Wilhelm and Vera, loved the local bakery. Yesterday, Jacob went in to admire a strawberry cake, which was on display in the bakery. Wilhelm stopped by to admire the strawberry cake too and so did Vera.

    ii. *Sentence.* I think that all the kids / each kid / every kid / the kids came to admire the same cake on display in the bakery.

  b. INVERSE-SCOPE

    i. *Scenario.* Two kids, Jacob and Vera, loved the local bakery. Yesterday, a cheese cake, a strawberry cake and a carrot cake were on display in the bakery. Jacob stopped by to admire them. Vera wanted to come by too, but she got sick and had to stay at home.

    ii.   *Sentence.* I think that the same kid came to admire all the cakes / each cake / every cake / the cakes on display in the bakery.

(24)   a.   SURFACE-SCOPE

    i.   *Scenario.* Three producers, Timothy, Chris and Mark, wanted to endorse one of two leading actors, Tom or Alec, in the best-selling show that they produced. Timothy chose Tom. So did Chris. Mark, however, chose Alec.

    ii.   *Sentence.* I think that all the producers / each producer / every producer / the producers endorsed the same leading actor in the best-selling show.

    b.   INVERSE-SCOPE

    i.   *Scenario.* Two producers, Timothy and Mark, wanted to endorse one of the three leading actors, Tom, Chris or Alec, in the best-selling show that they produced. Timothy could not decide, so he endorsed two actors: Tom and Alec. Mark chose to endorse Chris.

    ii.   *Sentence.* I think that the same producer endorsed all the leading actors / each leading actor / every leading actor / the leading actors in the best-selling show.

(25)   a.   SURFACE-SCOPE

    i.   *Scenario.* Three farmers, Maggie, Sabrine and Ulrika, went to the farmers' market downtown. It was spring and they wanted to sell their tomato plants. Maggie advertised her Celebrity tomato plant. So did Sabrine. Ulrika advertised her Celebrity tomato plant too.

    ii.   *Sentence.* I think that all the farmers / each farmer / every farmer / the farmers advertised the same tomato plant at the farmers' market downtown.

    b.   INVERSE-SCOPE

    i.   *Scenario.* Two farmers, Maggie and Ulrika, went to the farmers' market downtown. It was spring and they wanted to sell their Celebrity, Spider and Bush tomato plants. Maggie advertised her Celebrity tomato plant, her Spider tomato plant and her Bush tomato plant. Ulrika advertised nothing.

    ii.   *Sentence.* I think that the same farmer advertised all the tomato plants / each tomato plant / every tomato plant / the tomato plants at the farmers' market downtown.

(26)   a.   SURFACE-SCOPE

    i.   *Scenario.* Two patients, Dante and Gerard, were in the emergency room. A male nurse and two female nurses stopped by to comfort them. The male nurse comforted Dante. The two female nurses comforted Gerard.

    ii.   *Sentence.* I think that all the nurses / each nurse / every nurse / the nurses comforted the same patient in the emergency room.

    b.   INVERSE-SCOPE

    i.   *Scenario.* Three patients, Dante, Billy and Gerard, were in the emergency room. A male nurse and a female nurse stopped by to comfort them. The male nurse comforted Billy and Gerard. The female nurse comforted Dante.

    ii.   *Sentence.* I think that the same nurse comforted all the patients / each patient / every patient / the patients in the emergency room.

(27)   a.   SURFACE-SCOPE

    i.   *Scenario.* Two hair products, a shampoo and a gel, were on display at the local fair. Three women, Lisa, Louise and Miranda, stopped by to try them. Lisa tried the shampoo. So did Louise. Miranda tried the shampoo too.

    ii.   *Sentence.* I think that all the women / each woman / every woman / the women tried the same hair product at the local fair.

    b.   INVERSE-SCOPE

    i.   *Scenario.* Three hair products, a shampoo, a conditioner, and a gel, were on display at the local fair. Two women, Lisa and Miranda, stopped by to try them. Lisa tried the shampoo, the conditioner and the gel. Miranda did not like the sales person and tried nothing.

ii.   *Sentence.* I think that the same woman tried all the hair products / each hair product / every hair product / the hair products at the local fair.

(28)   a.   SURFACE-SCOPE

i.   *Scenario.* Arnim, Rob and Max used to be bartenders, but they were fired last month and are looking for a job now. There are two bars downtown: 417 and Penny's. Yesterday, Arnim went to 417 looking for a job. Rob asked for work at Penny's, and so did Max.

ii.   *Sentence.* I think that all the men / each man / every man / the men went to the same bar downtown asking for work.

b.   INVERSE-SCOPE

i.   *Scenario.* Arnim and Max used to be bartenders, but they were fired last month and are looking for a job now. There are three bars downtown: 417, Penny's, and High and Low. Yesterday, Arnim went to 417 and Penny's looking for a job. Max asked for work at High and Low.

ii.   *Sentence.* I think that the same man went to all the bars / each bar / every bar / the bars downtown asking for work.

(29)   a.   SURFACE-SCOPE

i.   *Scenario.* Three disciples, Urk, Doron and Edith, were asked to recite two morning prayers. They said the first prayer together. Afterwards, they dedicated the second prayer to the rising sun and jointly chanted it.

ii.   *Sentence.* I think that all the disciples / each disciple / every disciple / the disciples dedicated the same morning prayer to the rising sun.

b.   INVERSE-SCOPE

i.   *Scenario.* Two disciples, Urk and Edith, were asked to recite three morning prayers. Urk dedicated the first prayer to the rising sun and chanted it. He then dedicated the second and third prayers to the rising sun as well, after which he and Edith chanted these prayers together.

ii.   *Sentence.* I think that the same disciple dedicated all the morning prayers / each morning prayer / every morning prayer / the morning prayers to the rising sun.

(30)   a.   SURFACE-SCOPE

i.   *Scenario.* In a recent study about the behavior of bees, a team of biologists placed two beehives on a meadow: one in the northern part and one in the southern part. Then, they let the bees fly freely over the meadow. They noticed that after a while some bees flew to the northern beehive, while the rest settled on flying to the southern beehive.

ii.   *Sentence.* I think that all the bees / each bee / every bee / the bees flew to the same beehive on the meadow.

b.   INVERSE-SCOPE

i.   *Scenario.* In a recent study about the behavior of bees, a team of biologists placed three beehives on a meadow: one in the northern part, one in the southern part and one in the western part. Then, they let the bees fly freely over the meadow. They noticed that after a while some bees flew to the northern beehive, some bees flew to the southern beehive and the rest flew to the western beehive. Also, no bee visited more than one beehive.

ii.   *Sentence.* I think that the same bee flew to all the beehives / each beehive / every beehive / the beehives on the meadow.

(31)   a.   SURFACE-SCOPE

i.   *Scenario.* Three designers were developing a new role-playing game featuring two characters in the main quest, a wizard and a knight. The first designer worked on the wizard's background story. The second designer worked on the wizard's appearance. And the third designer worked on the wizard's animation.

ii.   *Sentence.* I think that all the designers / each designer / every designer / the designers worked on the same character in the main quest.

    b. INVERSE-SCOPE

        i. *Scenario.* Two designers were developing a new role-playing game featuring three characters in the main quest, a wizard, a priest and a knight. The first designer worked on the wizard, the priest and the knight. The second designer worked on the monsters that could be encountered in the main quest.

        ii. *Sentence.* I think that the same designer worked on all the characters / each character / every character / the characters in the main quest.

(32) a. SURFACE-SCOPE

        i. *Scenario.* Three journalists, Jason, Ewa and Dorothy, investigated two suspicious transactions authorized by a bank manager, one supporting guerilla groups in Africa and one involving Asian cocaine producers. Jason focused on the transaction to Africa. Ewa and Dorothy investigated the suspicious transaction involving Asian cocaine producers.

        ii. *Sentence.* I think that all the journalists / each journalist / every journalist / the journalists investigated the same suspicious transaction authorized by the bank manager.

    b. INVERSE-SCOPE

        i. *Scenario.* Two journalists, Jason and Dorothy, investigated three suspicious transactions authorized by a bank manager, one supporting guerilla groups in Africa, one involving Asian cocaine producers and one connected to high politics in Columbia. Jason focused on the transactions to Africa and Columbia. Dorothy investigated the suspicious transaction involving Asian cocaine producers.

        ii. *Sentence.* I think that the same journalist investigated all the suspicious transactions / each suspicious transaction / every suspicious transaction / the suspicious transactions authorized by the bank manager.

References

Altmann, Gerry & Mark Steedman. 1988. Interaction with context during human sentence processing. *Cognition* 30(3). 191–238. http://dx.doi.org/10.1016/0010-0277(88)90020-0.

AnderBois, Scott, Adrian Brasoveanu & Robert Henderson. 2012. The pragmatics of quantifier scope: A corpus study. *Proceedings of* Sinn und Bedeutung *(SuB) 16.* A. Aguilar-Guevara, A. Chernilovskaya & R. Nouwen (eds.). 15–28.

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

Baayen, R. Harald & Petar Milin. 2010. Analyzing reaction times. *International Journal of Psychological Research* 3(2). 12–28.

Barker, Chris. 2007. Parasitic scope. *Linguistics and Philosophy* 30(4). 407–444. http://dx.doi.org/10.1007/s10988-007-9021-y.

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278. http://dx.doi.org/10.1016/j.jml.2012.11.001.

Bates, Douglas, Martin Maechler, Ben Bolker & Steven Walker. 2013. *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.0-5. http://CRAN.R-project.org/package=lme4.

Beck, Sigrid. 2000. The semantics of *different*: Comparison operator and relational adjective. *Linguistics and Philosophy* 23(2). 101–139. http://dx.doi.org/10.1023/A:1005566722022.

Beghelli, Filippo & Timothy Stowell. 1997. Distributivity and negation. In Anna Szabolcsi (ed.), *Ways of scope taking*, 71–109. Dordrecht: Kluwer.

Brasoveanu, Adrian. 2011. Sentence-internal *different* as quantifier-internal anaphora. *Linguistics and Philosophy* 34(2). 93–168. http://dx.doi.org/10.1007/s10988-011-9096-3.

Brasoveanu, Adrian & Jakub Dotlačil. 2012. Licensing sentence-internal readings in English: An experimental study. *Logic, Language and Meaning: 18 Amsterdam Colloquium (revised selected papers)*. 122–132.

Brisson, Christine. 2003. Plurals, *all*, and the nonuniformity of collective predication. *Linguistics and Philosophy* 26(2). 129–184. http://dx.doi.org/10.1023/A:1022771705575.

Carlson, Gregory. 1987. *Same* and *Different*: Some consequences for syntax and semantics. *Linguistics and Philosophy* 10(4). 531–565. http://dx.doi.org/10.1007/BF00628069.

Cohen, Jacob. 1992. Statistical power analysis. *Current Directions in Psychological Science* 1(3). 98–101. http://dx.doi.org/10.1111/1467-8721.ep10768783.

Crain, Stephen & Mark Steedman. 1985. On not being led up the garden path: The use of context by the psychological syntax processor. In Lauri Karttunen David Dowty & Arnold Zwicky (eds.), *Natural language parsing: Psychological, computational and theoretical perspectives*, 320–358. Cambridge, UK: Cambridge University Press.

Dayal, Veneeta. 2012. The syntax of scope and quantification. In Marcel den Dikken (ed.), *The Cambridge handbook of generative syntax*, 827–859. Cambridge, UK: Cambridge University Press.

Dotlačil, Jakub. 2010. *Anaphora and distributivity. A study of 'same', 'different', reciprocals and 'others'*. Utrecht: Utrecht University PhD thesis.

Dotlačil, Jakub & Adrian Brasoveanu. 2014. The manner and time course of updating quantifier scope representations in discourse. *Language, Cognition and Neuroscience*. Published online: 14 May 2014. http://dx.doi.org/10.1080/23273798.2014.918631.

Dwivedi, Veena D., Natalie A. Phillips, Stephanie Einagel & Shari R. Baum. 2010. The neural underpinnings of semantic ambiguity and anaphora.

Brain Research 1311(22). 93–109. http://dx.doi.org/10.1016/j.brainres.2009.09.102.

Fodor, Janet Dean. 1982. The mental representation of quantifiers. In Stanley Peters & Esa Saarinen (eds.), *Processes, beliefs and questions*, 129–164. Dordrecht: Reidel.

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

Hendriks, Herman. 1993. *Compositionality and model-theoretic interpretation*. Amsterdam: University of Amsterdam PhD thesis.

Ioup, Georgette. 1975. Some universals for quantifier scope. In John Kimball (ed.), *Syntax and Semantics 4*, 37–58. New York, NY: Academic Press.

Just, Marcel A., Patricia A. Carpenter & Jacqueline D. Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 111(2). 228–238. http://dx.doi.org/10.1037/0096-3445.111.2.228.

Kamp, Hans. 1981. A theory of truth and semantic representation. In Jeroen Groenendijk, Theo Janssen & Martin Stokhof (eds.), *Formal methods in the study of language*, 277–322. Amsterdam: Mathematical Centre Tracts.

Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic: Introduction to model theoretic semantics of natural language, formal logic and Discourse Representation Theory*. Dordrecht: Kluwer.

Moltmann, Friederike. 1992. Reciprocals and *same/different*: Towards a semantic analysis. *Linguistics and Philosophy* 15(4). 411–462. http://dx.doi.org/10.1007/BF00627683.

Pagliarini, Elena, Gaetano Fiorin & Jakub Dotlačil. 2012. The acquisition of distributivity in pluralities. *Proceedings of the 36th annual Boston University Conference on Language Development*. Alia K. Biller, Esther Y. Chung & Amelia E. Kimball (eds.). 387–399.

Pylkkänen, Liina & Brian McElree. 2006. The syntax-semantic interface: Online composition of sentence meaning. In *Handbook of psycholinguistics*, 537–577. New York, NY: Elsevier.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org/.

Radó, Janina & Oliver Bott. 2012. Underspecified representations of scope ambiguity? *Logic, Language and Meaning: 18th Amsterdam Colloquium (revised selected papers)*. Maria Aloni, Vadim Kimmelman, Floris Roelofsen, Galit Weidman Sassoon, Katrin Schulz & Matthijs Westera (eds.). 180–189.

Reinhart, Tanya. 2006. *Interface strategies: Optimal and costly computations.* Cambridge, MA: MIT Press.

Roeper, Thomas, Barbara Z. Pearson & Margaret Grace. 2011. Quantifier spreading is not distributive. *Boston University Conference on Language Development (BUCLD) 35.* Nick Danis, Kate Mesh & Hyunsuk Sung (eds.). 526–539.

Trueswell, John, Michael Tanenhaus & Susan Garnsey. 1994. Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language* 33(3). 285–318. http://dx.doi.org/10.1006/jmla.1994.1014.

Tunstall, Susanne. 1998. *The interpretation of quantifiers: Semantics and processing.* University of Massachusetts, Amherst PhD thesis.

Adrian Brasoveanu
Stevenson Faculty Services / Linguistics
University of California Santa Cruz
1156 High Street
Santa Cruz, CA 95064
abrsvn@{gmail.com,ucsc.edu}

Jakub Dotlačil
Faculteit der Letteren, Algemene
taalwetenschap — Leerstoelgroep
Oude Kijk in 't Jatstraat 26
9712 EK Groningen
Netherlands
j.dotlacil@gmail.com