Competition during the processing of quantifier scope ambiguities:
Evidence from eye movements during reading

Kevin B. Paterson¹, Ruth Filik², & Simon P. Liversedge³

¹School of Psychology,
University of Leicester, UK

²Department of Psychology,
University of Glasgow, UK

³School of Psychology,
University of Southampton, UK

Address correspondence to:
Kevin Paterson
School of Psychology
Henry Wellcome Building
University of Leicester
Leicester, UK.
Email: kbp3@le.ac.uk
Abstract

We investigated the processing of sentences containing a quantifier scope ambiguity, such as *Kelly showed a photo to each critic*, which is ambiguous between the indefinite phrase (*a photo*) having one or many referents. Ambiguity resolution requires the computation of relative quantifier scope, with either *a photo* or *each critic* taking wide scope, thereby determining the number of referents. Using eye tracking, we established that multiple factors, including the grammatical function and surface linear order of quantified phrases, along with their lexical characteristics, interact during the processing of relative quantifier scope, with conflict between factors incurring a processing cost. We discuss the results in terms of theoretical accounts attributing sentence-processing difficulty to either reanalysis (e.g., Fodor, 1982) or to competition between rival analyses (e.g., Kurtzman & MacDonald, 1993).
Quantifiers (e.g., *some x, every x*) provide quantity information, such as the number of entities in a situation. However, sentences containing two or more quantifiers can be ambiguous, for example:

(1). Kelly showed a photo to every critic.

This sentence contains an indefinite noun-phrase (i.e., *a photo*) and a universal quantifier (i.e., *every critic*). It can mean that the critics collectively viewed one photo or that they each viewed a different photo. The ambiguity arises from the language processor having to compute relative quantifier scope. If the indefinite phrase is assigned wide scope, it will refer to a single entity (i.e., one photo). But if the universal quantifier is assigned wide scope, then the indefinite phrase is likely to be interpreted as referring to multiple entities (i.e., many photos).

For psycholinguists, the interest in ambiguities lies in the nature of the mental processes that enable readers and listeners to resolve them. The study of lexical and syntactic ambiguity resolution has proven fruitful in developing theoretical accounts of such resolution processes (e.g., Duffy, Morris, & Rayner, 1988; Frazier, 1995; Tabor & Tanenhaus, 1999; Van Gompel, Pickering, & Traxler, 2001). Within these domains, an important debate concerns whether the language processor resolves ambiguities by first selecting a single analysis and reanalysing if it proves to be deficient, or by considering the alternative possible analyses in parallel, each of which competes for adoption. It is likely that a similar debate will be relevant when considering the processing of quantifier scope ambiguities. Given this, we examined whether difficulty for these ambiguities is attributable to reanalysis or to competition. We focused on the influence of lexical and structural factors, but note that prior referential context (e.g., Anderson, 2004; Villalta, 2003) and real-world knowledge (e.g., Sanford & Sturt, 2002) also may play a role.

Fodor (1982) proposed a reanalysis-based processing account incorporating the observation that quantifiers usually are interpreted in the same sequence that they occur in a sentence (e.g., Bunt, 1985; Johnson-Laird, 1969; Johnson-Laird, Byrne, & Tabossi, 1989; Lakoff, 1971, Tunstall, 1998) and that quantifiers differ in their propensity for taking wide
scope. She argued that the language processor initially assigns wide scope to the first quantifier in an ambiguous sentence, and that reanalysis is triggered if this analysis is incongruent with subsequent linguistic information, such as the lexical biases of individual quantifiers. Thus, a photo in (1) should be assigned wide scope at first and be interpreted as referring to a single entity. On encountering every critic, the processor may reassign wide scope to this constituent, as universal quantifiers such as each and every have a lexical bias for taking wide scope. In this case, an initial analysis with the indefinite phrase referring to one entity must subsequently be revised to one with it referring to many entities. As revising the number of referents involves modifying the contents of the semantic representation, this incurs a processing cost.

No such difficulty should exist for Kelly showed every critic a photo. For this sentence, the language processor at first should assign wide scope to every critic, and interpret it as referring to multiple entities (e.g., many critics). On encountering a photo, the processor selects between analyses with this phrase referring to one or many entities. As neither option requires revision of the initial semantic representation (with every critic referring to many critics), but simply that it is updated in line with interpretative decisions, no further cost is incurred. Thus, the reanalysis-based account predicts difficulty in processing ambiguities involving an indefinite phrase and a universal quantifier if the indefinite phrase is first but not if the universal quantifier is first.

This approach may be contrasted with accounts in which the syntactic position of quantifiers, rather than their surface position, influences scope processing (e.g., Ioup, 1975; Reinhart, 1983). Ioup devised a hierarchy of grammatical categories, proposing that a quantifier in one category preferentially takes scope over those in a position lower in the hierarchy, as follows (where ‘>’ indicates ‘takes scope over’): topic > deep and surface subject > deep subject or surface subject > indirect object > prepositional object > direct object. She proposed that the grammatical hierarchy interacts with a hierarchy that ranks quantifiers in terms of their propensity to take wide scope, as follows: each > every > a > all > most > many > several > some > a few. The construction of these hierarchies was not theoretically motivated, and Ioup
instead developed them from judgement data that she collected for a range of syntactic constructions and quantifiers, in a variety of languages.

Ioup (1975) also was not explicit about how the hierarchies might interact, but it can be expressed in terms of competition, following Kurtzman and MacDonald (1993). Kurtzman and MacDonald suggested that the possible analyses of an ambiguity are evaluated in parallel and compete for adoption. They argued that when factors collectively favour one representation then that representation is constructed, but if the factors are in conflict then competition between the alternative representations occurs before one finally is selected, with competition incurring a processing cost. Thus, the account predicts that a cost is incurred if the quantifier and grammatical hierarchies favour different analyses of an ambiguity.

The above accounts make the same processing predictions for many English sentences, including double object sentences such as (2), in which the indirect object (e.g., every critic) precedes the direct object (e.g., a photo).

(2). Kelly showed every critic a photo.

The reanalysis-based account predicts that there is a bias for every critic to take wide scope because it is first, whereas the competition-based account attributes this bias to a preference for a quantified indirect object to take wide scope over the direct object. Both accounts predict that difficulty occurs when structural biases conflict with lexical biases. Thus, double object sentences with ‘a-every’ rather than ‘every-a’ order should cause most difficulty. The accounts make opposing predictions for some sentences, including datives such as (1). The order of the direct and indirect objects is reversed in datives as compared with double object sentences. Whereas the reanalysis-based account makes the same prediction for datives as for double object sentences, the competition-based account predicts difficulty for ‘every-a’ datives, because a bias for assigning wide scope to the indirect object (i.e., the second quantifier) conflicts with a lexical bias for every to take wide scope.

Filik, Paterson, and Liversedge (2004) tested these predictions, using eye-tracking to
investigate the processing of sentences containing a and every. The key finding was that total reading times at a region containing the direct and indirect objects were longer for double object sentences with ‘a-every’ than ‘every-a’ order, but that the effect was reversed for datives, with longer reading times for datives with ‘every-a’ than ‘a-every’ order. Filik et al. took this to support the competition-based account.

There were two other findings at this region. First, Filik et al. (2004) observed that the magnitude of the reading time cost was greater for double object sentences than for datives, and conjectured that this was due to the surface order of quantifiers influencing scope processing, in addition to the quantifier and grammatical hierarchies. Research into lexical processing suggests that more difficulty is experienced for balanced ambiguities, where there is an equal likelihood of selecting each of the alternative analyses, than for ambiguities that have a strongly preferred analysis (e.g., Duffy et al., 1988), although similar claims for syntactic ambiguity resolution are disputed (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; Traxler, Pickering, & Clifton, 1998). It was possible that the quantifier scope ambiguity was more balanced, and incurred a greater cost, when surface order and the grammatical hierarchy jointly opposed the quantifier hierarchy than when these factors were in conflict. The other finding was that datives produced the same pattern of effects in first pass and total reading times, but no first pass effects were obtained for double object sentences. Filik et al. speculated that the additional syntactic complexity of double object sentences (Larson, 1988, 1990; but see Jackendoff, 1990) imposed a syntactic processing load that caused higher order interpretative processes to be delayed.

Filik et al. (2004) also examined reading times for a continuation that made singular or plural reference to the indefinite phrase (e.g., but the photo(s) was / were not very impressive provides a continuation for either (1) or (2)). They found that reading times were longer for plurals, and offered two explanations for this effect. First, readers might have an overriding disposition to assign a singular interpretation to this phrase, because it is referentially simpler. Alternatively, they may have developed a coherent understanding of the text without interpreting
the anaphor (e.g., *the photo*) with respect to a representation specifying relative scope. Instead, as previous studies had suggested (e.g., Cloitre & Bever, 1988), readers may have interpreted the anaphor with respect to a surface representation of the text, with co-reference computed solely on the basis of the match between morphological features of the anaphor and its antecedent.

The present experiment extended this research by investigating the influence of *each* on scope processing. Although *each* and *every* both are universal quantifiers, theoretical accounts claim that they differ in their propensity for taking scope over other constituents (e.g., Beghelli & Stowell, 1997; Ioup, 1975; Vendler, 1967). Ioup ranked *each* higher on the quantifier hierarchy than *every*, predicting that it is more strongly disposed to taking wide scope. Beghelli and Stowell argued that whereas *each* is obligatorily distributive, and must take scope over other constituents, *every* is only optionally distributive. It was possible that because *every* is optionally distributive, the readers in Filik et al.’s (2004) study were not strongly disposed to constructing a representation with the indefinite phrase referring to many entities. This might explain the absence of quantifier scope effects at the continuation region in this study. We expected *each* to cause readers to be more strongly disposed to adopting an analysis with it taking wide scope. This stronger disposition might be reflected in differences in reading times, with effects occurring both at the continuation region, as well as earlier, more robust effects, at the quantified region. Table 1 illustrates our hypotheses concerning effects at these regions.

```
| Effects at the quantified region were expected to be informative about costs incurred in processing the ambiguity, thereby providing a test of reanalysis- and competition-based accounts. Effects at the continuation region should be informative about ambiguity resolution. Because the continuation always referred to the indefinite phrase, reading times should be shortest for singular continuations if this phrase is interpreted as referring to one entity, whereas they should be shortest for plural continuations if it is interpreted as referring to many entities. If the ambiguity is resolved by assigning wide scope to the first quantifier then the indefinite |
```

phrase should refer to one entity when it is first but refer to many entities when each is first. If instead the quantifier hierarchy determines ambiguity resolution, then each should take wide scope, and the indefinite phrase should always refer to many entities. The grammatical hierarchy predicts a three-way interaction of quantifier order, constituent order, and continuation number. For double object sentences, it predicts that the first quantifier takes wide scope. Thus, when the indefinite phrase is first, it should refer to one entity. But when each is first, the indefinite phrase should refer to many entities. As the order of constituents is reversed for datives, the grammatical hierarchy predicts the opposite pattern of effects for these sentences. Our data also may be informative about the putative role of syntactic complexity. If the non-canonical form of double object sentences causes interpretative processes to be delayed, then we may obtain first pass effects at the quantified region for datives only.

EXPERIMENT

METHOD

Participants: Forty-eight native English speakers with normal or corrected vision from the University of Derby community participated.

Materials & Design: We modified 48 dative and double object sentences used by Filik et al. (2004) by substituting each for every (examples in Table 1). An indefinite phrase or each was first. The continuation was singular or plural, but always referred to the indefinite phrase. There were three independent variables: quantifier order, constituent order, and continuation number. The dependent variables were measures of reading time for the sentence regions.

Norming Data: Filik et al. (2004) established that singular and plural forms of the continuation noun (e.g., reporter vs. reporters) did not differ in frequency, using the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995), (singular = 81.8 words/million (sd = 132.3), plural = 67.2 words/million (sd = 125.2), t(47)<1). We also collected norming data to ensure that different versions of our sentences did not differ in plausibility. We noted that, depending on quantifier order, a different noun formed the indefinite phrase (e.g., an interview vs. a reporter).
It was possible that sentences would be more plausibly interpreted as referring to one or many entities when this phrase was formed from one noun than another. Therefore, we examined the perceived plausibility of disambiguated sentences. We tested datives only, as disambiguated versions of dative and double object sentences had the same interpretation.

In one study, the sentences were disambiguated as referring to a single entity by substituting a definite phrase for the indefinite phrase (e.g., the interview). Six participants used a 7-point scale (‘1’ = highly implausible, ‘7’ = highly plausible) to evaluate sentences with this phrase in the direct or indirect object position (e.g., the interview vs. the reporter). As the ratings did not differ (4.7 vs. 4.6), \( t(47) < 1 \), it appeared that participants found the sentences to be equally plausible. In a second study, the indefinite phrase was disambiguated as referring to many entities by including the word different. Another six participants evaluated sentences with the disambiguated phrase in the direct or indirect object position (e.g. a different interview vs. a different reporter). Again, the ratings did not differ (3.4 vs. 3.5), \( t(47) = 1.3, p > .05 \), suggesting that the sentences were perceived to be equally plausible. Thus, it appeared that there was no systematic difference in plausibility for alternative interpretations of the sentences.

We conducted a number judgment task to assess the likelihood of the indefinite phrase being interpreted as referring to one or many entities. Twenty-four participants used a 5-point scale (‘1’ indicated “definitely one”, ‘5’ indicated “definitely more than one”) to evaluate sentences without a continuation (e.g., The celebrity gave an in-depth interview to each reporter from the newspaper.). The results were analysed using two 2(quantifier order) X 2(constituent order) ANOVAs, treating participants (\( F_1 \)) and items (\( F_2 \)) as random variables. Effects were considered significant when both analyses were significant at \( p < .05 \). The difference between dative and double object sentences was unreliable (2.8 vs. 2.5), \( F_1(1, 23) = 8.57, p < .05 \), and \( F_2(1, 47) = 2.0, p > .05 \). However, there was an effect of quantifier order, \( F_1(1, 23) = 18.53, p < .001 \), and \( F_2(1, 47) = 15.66, p < .001 \). Participants rated the indefinite phrase as most likely to refer to many entities when each was first (‘a-each’ = 2.3, ‘each-a’ = 2.8). This effect was
qualified by an interaction of quantifier and constituent order $F_1(1, 23) = 17.11, p<.001$, and $F_2(1, 47) = 32.40, p<.001$. Although the indefinite phrase was rated as more likely to refer to many entities when each was first for double object sentences ('a-each' = 1.9, ‘each-a’ = 3.1), $F_1(1, 23) = 28.28, p<.001$, and $F_2(1, 47) = 32.40, p<.001$, no such effect was observed for datives ('a-each’ = 2.9, ‘each-a’ = 2.6; $F_s<1.8$). Thus, there was a bias for interpreting the first quantifier as taking wide scope in double object sentences but not datives.

**Procedure:** A Fourward Technologies Dual Purkinje Image Generation 6 eye-tracker monitored gaze location and participants' right eye movements during reading. It has an angular resolution of 10 min arc. A PC displayed materials on a monitor 80cm from participants' eyes. Gaze location was monitored every millisecond. Output was sampled to produce a sequence of fixations, recorded as x and y character positions, with start and finish times.

We created 8 lists of sentences, each containing one version of each sentence. No sentence appeared more than once in a list. Sentences were displayed in fixed random order, double-spaced across two lines, with the continuation centred on line two, together with 32 sentences from an unrelated experiment and 32 filler sentences. The experiment was run in two blocks, with two practice sentences beginning each. Participants were instructed to read normally and for comprehension. Once seated at the eye-tracker, they completed a calibration procedure with additional checks performed prior to each trial. Head movements were minimized using forehead restraints and a bite bar. Participants pressed a key after reading each item and a comprehension question was displayed. They responded by pressing 'yes' or 'no' keys, with feedback.

**RESULTS**

*Regions:* Sentences were divided into five analysis regions (shown in Table 1), with reading times reported for Regions 2 - 4. Region 2 was the quantified region. Region 3 was the continuation. Region 4 was a spillover region.

*Analysis:* Short contiguous fixations were pooled automatically. Fixations under 80 msec were incorporated into larger adjacent fixations within one character. Those under 40 msec and
not within three characters of another fixation were deleted, as were fixations over 1200 msec. Before analysing the data, we removed trials where participants had failed to read the text or where there had been tracker loss by deleting trials where two neighbouring regions had zero first-pass reading times, accounting for 4.1% of the data. Due to Experimenter error we failed to record data for one condition of one item, producing a total data loss of 4.4%. The remaining data for the affected item were included in $F_1$ analyses but excluded from $F_2$ analyses. Comprehension was high, with 87% correct responses to comprehension questions, and no significant differences across conditions ($F$s<3.6).

We computed first pass and total reading times for each region. First pass reading time summed the duration of fixations made on first entering a region and before exiting it, taken to reflect difficulty occurring early during text processing. Total reading times summed the duration of all fixations in a region until participants pressed a button to indicate that they had finished reading, providing a measure of overall difficulty in a region. Singular and plural continuations differed in length by up to 4 characters (mean = 2.1). To correct for length differences we calculated residual (deviations from predicted scores) first pass and total reading times for Region 3 (Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994). Residual and raw reading times showed the same effects; therefore only analyses for raw times are reported. Data for each region were subjected to two 2(quantifier order) x 2(constituent order) x 2(continuation number) ANOVAs, for participants ($F_1$) and items ($F_2$). Effects were considered significant when both analyses were significant at $p$.05. Table 3 shows the mean reading times.

Region 2 (Quantified Region): There were no significant main effects in first pass reading times ($F$s<3). However, quantifier and constituent order interacted $F_1(1, 47) = 6.82, p$.05, and $F_2(1, 46) = 5.27, p$.05, as illustrated in Figure 1 Panel A. First pass reading times were longer for double object sentences with ‘a-each’ rather than ‘each-a’ order (1645 vs. 1506 msec), $F_1(1,$
$47 = 11.10, p < .001$, and $F_2(1, 46) = 5.75, p < .05$, with no such difference for datives (‘each-a’ = 1675 msec, ‘a-each’ = 1627 msec, $F s < 1$). There were no other first pass effects ($F s < 1.4$). Thus, quantifier order influenced the first pass processing of double object sentences only.$^1$

Total reading times were longer for double object sentences than for datives (2630 vs. 2499 msec), $F_1(1, 47) = 7.43, p < .01$, and $F_2(1, 46) = 4.78, p < .05$. The continuation effect was not significant ($F s < 3.5$). There was an effect of quantifier order, $F_1(1, 47) = 25.79, p < .001$, and $F_2(1, 46) = 14.88, p < .001$, with longer total reading times when the indefinite phrase was first (2669 vs. 2460 msec). However, this effect was qualified by a quantifier and constituent order interaction, $F_1(1, 47) = 48.90, p < .001$, and $F_2(1, 46) = 38.68, p < .001$, as illustrated in Figure 1 Panel B. Total times were longest for double object sentences when the indefinite phrase was first (2917 vs. 2342 msec), $F_1(1, 47) = 53.32, p < .001$; $F_2(1, 46) = 43.93, p < .001$, but longest for datives when each was first (2578 vs. 2421 msec), $F_1(1, 47) = 9.16, p < .01$; $F_2(1, 46) = 4.06, p < .05$. Thus, quantifier order affected sentence processing, although the effect obtained for double object sentences was reversed for datives, as predicted by the competition-based account.

Region 3 (Continuation Region): First pass reading times were shortest for singular continuations (336 vs. 364 msec), $F_1(1, 47) = 21.34, p < .001$, and $F_2(1, 46) = 16.10, p < .001$, and for datives (341 vs. 359 msec), $F_1(1, 47) = 5.81, p < .05$, and $F_2(1, 46) = 6.50, p < .05$. Constituent and quantifier order did not interact significantly ($F s < 3.5$), with no other effects ($F s < 1.9$). Thus, the early processing of the continuation was impervious to factors affecting scope computation.

Total reading times also were shortest for singular continuations (502 vs. 532 msec), $F_1(1, 47) = 22.64, p < .001$; $F_2(1, 46) = 13.75, p < .001$. Thus, the quantifier hierarchy did not affect anaphor processing. The absence of a quantifier order and continuation interaction ($F s < 1$) or significant three-way interaction ($F s < 2.1$) indicated that its processing was unaffected by quantifier order or the grammatical hierarchy. It therefore appeared that the processing of the continuation was insensitive to factors affecting scope computation.

Other total time effects were consistent with a spillover of effects from the quantified
region. Reading times were longest when the indefinite phrase was first (533 vs. 502 msec), $F_1(1, 47) = 5.98, p<.05$, and $F_2(1, 46) = 6.35, p<.05$, and for datives (530 vs. 505 msec), $F_1(1, 47) = 4.13, p<.05$, and $F_2(1, 46) = 4.77, p<.05$. Quantifier and constituent order interacted, $F_1(1, 47) = 8.56, p<.01$, and $F_2(1, 46) = 5.99, p<.05$, with longer reading times for double object sentences when the indefinite phrase was first (566 vs. 496 msec), $F_1(1, 47) = 14.37, p<.001$, and $F_2(1, 46) = 12.28, p<.001$, but no difference for datives (500 vs. 509 msec) ($F$s<1). There were no other significant effects ($F$s<2.9).

Region 4 (‘not very’): No effects of interest occurred at this region. The first pass continuation effect was unreliable, $F_1<3.2$, and $F_2(1, 46) = 4.76, p<.05$, as was the quantifier order and continuation interaction in total reading times, $F_1<.9$, and $F_2(1, 46) = 4.83, p<.05$, with no other significant effects ($F$s<1.4).

To summarise, the most important aspect of our results was the reliable interactive effect of quantifier and grammatical order for the total reading times in the quantified region. A similar, but less robust, pattern of effects also occurred in the first pass reading times.

DISCUSSION

We contrasted two accounts of relative scope processing. A reanalysis-based account predicted that quantifiers initially are interpreted in the same sequence that they occur in a sentence, irrespective of its structure, and that reading difficulty occurs when this analysis conflicts with subsequent linguistic information. By contrast, a competition-based account predicted that the alternative possible interpretations of an ambiguity compete for adoption, with the degree of competition between these determining the extent of the reading difficulty. We expected reading times for the quantified region of our sentences to be informative about ambiguity processing and those for the continuation to be informative about its interpretation.

We will discuss the reading times for the quantified region first. Effects at this region were clearest in the total reading times, and although first pass reading times produced the same pattern of effects, these were less robust. The total reading times produced an interaction of
quantifier and grammatical order. Whereas readers incurred a cost for double object sentences with ‘*a-each*’ rather than ‘*each-a*’ order, the effect for datives was reversed, with readers experiencing most difficulty for sentences with ‘*each-a*’ order. This effect was contrary to the reanalysis-based account, which predicted readers would have difficulty when the indefinite phrase was first, irrespective of the sentence’s construction.

Instead, the data were most consistent with Ioup’s (1975) claim that the relative syntactic position of quantifiers interacts with lexical biases to determine relative quantifier scope. Ioup described this in terms of an interaction between a grammatical hierarchy ranking the tendency for quantifiers in different syntactic position to take wide scope, and a quantifier hierarchy ranking the propensity for individual quantifiers to take wide scope. We reformulated Ioup’s proposition into a competition-based account (following Kurtzman & MacDonald, 1993), predicting that most difficulty occurs when the hierarchies support different interpretations of an ambiguity. The total reading times matched with this prediction, as readers had most difficulty when a grammatical bias (i.e., for the quantifier in the indirect object position to take wide scope) conflicted with a lexical bias (i.e., for *each* to take wide scope). Of concern to some may be the lack of theoretical underpinning for Ioup’s hierarchies, which she constructed from empirical data concerning the preferred interpretation of quantifier scope ambiguities. Nevertheless, despite their provenance, an interaction between these hierarchies accurately predicted the effects obtained in this study, although clearly further research is needed to verify their psychological reality.

Like Filik et al. (2004), we also observed a larger total reading time cost for double object sentences than for datives when the hierarchies were in conflict. As Filik et al. contended, this difference may be due to the surface order of quantifiers contributing to scope processing, albeit in addition to the influence of the hierarchies. The degree of difficulty in processing the ambiguity may depend on the relative balance of these factors, such that readers have most difficulty when surface order and the grammatical hierarchy jointly oppose the quantifier
hierarchy, but have less difficulty when these factors support different analyses.

First pass reading times for the quantified region were expected to be informative about early ambiguity processing. Like the total reading times, these produced an interaction of quantifier and grammatical order, but the difference between ‘a-each’ and ‘each-a’ orders was reliable for double object sentences only. This could have been due to quantifier scope effects occurring earlier during the course of processing for double object sentences than for datives. However, in contrast to the present data, Filik et al. (2004) obtained first pass effects for datives, but not for double object sentences containing an indefinite phrase and the quantifier every. Filik et al. speculated that the absence of first pass effects for double object sentences in their experiment was due to the syntactic complexity of these sentences causing interpretative processes to be delayed. Our results are not consistent with this explanation. Therefore, another explanation is needed for the contrasting first pass effects in the two experiments.

One possibility is that relative quantifier scope is computed quite late in sentence comprehension, and its computation is not reflected as immediately in the eye movement record as computations associated with other forms of linguistic processing (e.g., lexical identification, syntactic parsing). Thus, although some linguistic effects often manifest in early reading time measures, relative scope effects might be observed reliably only in later measures. Additionally, it is possible that the regions that we selected for our analyses prevented us from detecting short-lived effects that occurred during the first pass processing of the ambiguity. Note that the quantified region in our study was large, averaging 8.9 words (9.4 words for datives and 8.4 words for double object sentences). Regions of this length contrast with those often used in research into syntactic ambiguity resolution (e.g., Traxler & Pickering, 1996), which are typically short (usually one word), and are used to gain detailed information of the immediacy of disambiguation effects. Such an approach has been vital in evaluating theoretical accounts (e.g., Garden Path theory; Frazier, 1987) positing that an initially-selected analysis may undergo rapid reanalysis. Unless subjected to appropriately fine grained analyses, early reanalysis effects can
be indistinguishable from the effects of competition.

Note, however, that in conducting our analyses, we avoided partitioning the quantified region into smaller regions, as this would have involved comparing reading times across regions that differed markedly in content. Effects obtained for such comparisons might then have been attributable to content differences between sentences rather than to the specific variables that we manipulated. Nevertheless, we recognise that future research must involve conducting detailed examinations of ambiguity processing using small regions and comparing the processing of ambiguous sentences, such as those used in this experiment, with unambiguous counterparts, perhaps ones rendered unambiguous by including a definite noun-phrase in place of the indefinite phrase or the universal quantifier. This would enable comparison of first pass reading times for identical text regions across ambiguous and unambiguous sentences, thereby potentially permitting the examination of very early effects. Such an approach would allow additional theoretical questions to be assessed too, including if there is a greater cost in constructing semantic representations involving many-to-many mapping (e.g., with many critics and reporters) rather than one-to-many mappings (e.g., with one critic and many reporters).

Before considering effects at the continuation region, we must address possible counter-explanations of the total time effects at this region. First consider that readers had most difficulty with double object sentences when each quantified the direct object, which we attributed to the influence of the quantifier and grammatical hierarchies on relative scope processing. It could be argued that this effect was instead due to readers having difficulty in establishing the scope of each. Each can, in principle, take scope over a constituent that precedes it in a sentence, or one that follows it. Thus, when each precedes the direct object in a double object sentence, such as The celebrity gave a reporter from the newspaper each in-depth interview, it is temporarily ambiguous between quantifying the direct object (e.g., in-depth interview) or the indirect object (e.g., a reporter from the newspaper). Perhaps, then, the inflated reading times for these sentences were due to readers experiencing temporary difficulty in establishing the constituent
that *each* quantifies. However, the same ambiguity cannot occur for sentences with *every*, as it cannot take scope over a preceding constituent. As Filik et al. (2004) obtained the same effects for sentences with *every* as we did for *each*, it is unlikely that the effect obtained in our study was due to readers having difficulty in establishing the direction of quantifier scope.

One intriguing possibility is that our reading time effects are due to preferences regarding the ordering of given and new information rather than to relative scope computation. Although it is widely argued that there is a preference for given information to precede new (e.g., Clark & Clark, 1977; Clark & Haviland, 1977), Clifton and Frazier (2004) reported that this preference does not hold for all sentence constructions. They examined the latencies of acceptability judgements for dative and double object sentences that differed in the order of definite and indefinite post-verbal noun-phrases, and found that the latencies were shorter for double object sentences when the definite phrase was first (e.g., *The senator mailed the woman a report.* vs. *The senator mailed a woman the report.*), reflecting a preference for given information (as signalled by definiteness) to precede new. By contrast, judgement latencies for datives were unaffected by the order of the phrases. Clifton and Frazier took these results as evidence against a general preference for given information to precede new, although they were unable to determine precisely why the effect differed for dative and double object sentences.

We used the same sentence constructions as Clifton and Frazier (2004), but whereas they manipulated the order of definite and indefinite phrases, we manipulated the ordering of a quantified phrase and an indefinite phrase. It could be argued that phrases quantified by *each* (and possibly *every*) have similar referential properties to definite phrases (i.e., they both are discourse-linked, meaning that they are interpreted with respect to context and therefore usually supply given information, e.g., Frazier, 1999). If a comparison is made between the two sets of results, treating the quantified and definite phrases as equivalents, then there is a broadly similar pattern (with a cost for double object sentences when either the quantified or indefinite phrase is first, and a reversed but smaller effect for datives). Thus, it could be argued that our results...
reflect processing preferences for given and new information rather than the processing of relative quantifier scope. Such an argument is not supported by the evidence from the number judgement task, however, which showed that the ordering of quantified and indefinite phrases influenced judgements concerning the likelihood of the indefinite phrase referring to one or more entities. These data instead suggest that our manipulation affected the computation of relative quantifier scope in the judgment task, and it is reasonable to conclude on the basis of these data that the effects observed in on-line reading times reflect its computation too.

We turn now to the continuation region, where reading times were longer for plural noun phrases, with no modulating influence of relative scope computation. Filik et al. (2004) argued that the absence of scope effects at this region might be due to readers either having a default preference for the singular interpretation of an indefinite phrase, or else interpreting the anaphor by computing co-reference exclusively on the basis of the match between morphological features of the anaphor and its antecedent (e.g., Cloitre & Bever, 1988). Also, when a universal quantifier takes wide scope over an indefinite phrase, it can quantify over events rather than entities (e.g., Geurts, 2002; Poesio, 1996). Thus, a sentence such as Kelly showed a photo to every critic can mean that the same photo was shown to different critics on separate occasions. It was possible, therefore, that participants interpreted each as quantifying over events and understood the indefinite to refer to a single entity. A final possibility is that relative scope remained underspecified at least until past the region containing the anaphor. Our eye-tracking data do not discriminate between these possibilities. But our off-line judgement data do give credence to explanations claiming that readers have a strong preference for the singular analysis of an indefinite phrase. These data indicated that although quantifier order modulated number judgements for this phrase, its influence was weak, and participants did not judge the phrase to definitely refer to many entities even under the most propitious conditions.

To conclude, in designing our experiment we expected that the reputedly strong disposition for each to take scope over other constituents (e.g., Ioup, 1975, Beghelli & Stowell,
1997; Vendler, 1967) would produce quantifier scope effects at the continuation, and robust first pass effects at the quantified region. However, this was not the case, and we obtained similar effects to Filik et al. (2004), who used every rather than each. The claim for a difference in the scope-taking propensity of these quantifiers is based on linguistic judgments. In Ioup’s case, participants in an experiment formed the judgments, but in other cases, the theorists themselves formed them. When considering the process of scope computation it is important to take seriously the possibility that the cognitive processes involved in performing off-line number judgements need not be the same as those involved in on-line sentence comprehension. Thus, our results do not rule out the possibility that universal quantifiers differ in scope-taking propensity. It may be that these differences only become apparent to a language user when consciously reflecting on sentence meaning or attempting to reason with quantifiers, but without influencing the rapid interpretative processes that occur during on-line sentence comprehension.

Just as the study of lexical and syntactic ambiguity resolution has given an insight into these aspects of language processing, the study of quantifier scope ambiguity resolution may be informative about the cognitive mechanisms underlying sentence interpretation. We have presented evidence that we believe to be consistent with competition effects during relative scope processing, yet many issues remain to be addressed. Not least, future research must determine if relative quantifier scope is computed early or late in sentence comprehension, and how its computation affects the interpretation of referential expressions.
References


Frazier, L. (1987). Sentence processing: a tutorial review. In M. Coltheart (Ed.), *Attention and


Footnote

1. We conducted further analyses in which we divided the quantified region into two regions containing each of the quantified phrases and examined first pass reading times for the second phrase in isolation, since this was the first point in the sentence at which we might observe effects of the reader evaluating the relationship between the two phrases. This phrase differed in length, content, and grammatical function across experimental conditions (e.g., a / each reporter from the newspaper vs. a / each in depth interview), which we compensated for by computing millisecond per character reading times. The analyses showed that first-pass reading times were longer for phrases quantified by each than for indefinites (24.7 vs. 21.6 ms/char, $F_1(1, 47) = 6.86, p<.05$, and, $F_2(1, 46) = 15.27, p<.001$). There is more than one possible interpretation of this effect. It could provide evidence in support of the reanalysis-based account, since this account predicts difficulty at the second phrase if it is quantified by each and the first phrase is indefinite. However, the effect might equally indicate that there is a processing cost associated with universal quantification that is not experienced when processing an indefinite phrase.
Acknowledgments

We thank Chuck Clifton and Matt Traxler for providing helpful reviews. Requests for reprints should be addressed to Kevin Paterson, School of Psychology, Henry Wellcome Building, Lancaster Road, University of Leicester, Leicester, LE1 9HN, UK. Email: kbp3@le.ac.uk.
Appendix

The following are dative versions of the sentence materials used in the experiment. Slashes denote alternatives and brackets denote modifications made when forming plural nouns. Deleting the preposition and reversing the order of the direct and indirect objects created double object versions of the sentences.

1. The estate agent showed [a / each] city-centre apartment to [a / each] new client, although the [apartment(s) / client(s)] [was / were] not very impressive / impressed.

2. The party host offered [a / each] unattached girl to [a / each] eligible bachelor, although the [girl(s) bachelor(s)] [was / were] not very attractive.

3. The hostess sang [a / each] karaoke song to [a / each] middle-aged man in the bar, although the [song(s) / man (men)] [was / were] not very popular.

4. The landlady served [a / each] evening meal to [a / each] long-term resident in the guesthouse, although the [meal(s) / resident(s)] [was / were] not very generous.

5. The project manager assigned [a / each] clerical assistant to [a / each] researcher in the department, although the [assistant(s) / researcher(s)] [was / were] not very helpful.

6. The music teacher taught [a / each] sea shanty to [a / each] girl in the class, although the [sea shanty (shanties) / girl(s)] [was / were] not very melodic.

7. The librarian lent [a / each] storybook to [a / each] boy in the class, although [the book(s) / boy(s)] [was / were] not very interesting.

8. The Headmaster sent [a / each] teaching plan to [a / each] teacher in the school, but the [plan(s) / teacher(s)] [was / were] not very organised.

9. The teacher recited [an / each] old nursery rhyme to [a / each] child in the infant class, although the [rhyme(s) / child (children)] [was / were] not very interesting / interested.

10. The fashion editor showed [a / each] new dress design to [a / each] clothing manufacturer, although the [design(s) / manufacturer(s)] [was / were] not very interested.
11. The government minister sent [a / each] informal letter to [a / each] high court judge, however the [letter(s) / judge(s)] [was / were] not very friendly.

12. The PhD student read [a / each] thesis chapter to [a / each] professor in the department, although the [chapter(s) / professor(s)] [was / were] not very impressive.

13. The editor assigned [a / each] murder story to [a / each] news reporter on the team, although the [story (stories) / reporter(s)] [was / were] not very remarkable.

14. The President sent [an / each] important broadcast to [a / each] country in Africa, although the [broadcast(s) / country (countries)] [was / were] not very threatening.

15. The teacher gave [a / each] maths problem to [a / each] clever pupil in the class, although the [problem(s) / pupil(s)] [was / were] not very interesting / interested.

16. The prisoner posted [a / each] threatening letter to [a / each] senior officer in the prison, although the [letter(s) / officer(s)] was / were not very nasty.

17. The publisher offered [a / each] study grant to [a / each] children's author, although the [grant(s) / author(s)] [was / were] not very well known.

18. The student emailed [a / each] experimental analysis to [a / each] statistics lecturer in the department, although the [analysis (analyses) / lecturer(s)] [was / were] not very thorough.

19. The manager allocated [a / each] new task to [a / each] junior secretary in the department, although the [task(s) / secretary (secretaries)] [was / were] not very economical.

20. The technician sent [a / each] computer program to [a / each] specialist in the company, but the [program(s) / specialist(s)] [was / were] not very reliable.

21. The celebrity gave [an / each] in-depth interview to [a / each] reporter from the newspaper, but the [interview(s) / reporter(s)] [was / were] not very interesting / interested.

22. The manager demonstrated [a / each] new computer to [a / each] technician in the laboratory, although the [technician(s) / computer(s)] [was / were] not very reliable.

23. The farmer offered [a / each] poorly fed cow to [a / each] butcher in the town, although the [cow(s) / butcher(s)] [was / were] not very impressive.
24. The designer offered [a / each] construction contract to [a / each] reliable builder in the district although the [contract(s) / builder(s)] [was / were] not very cheap.

25. The doctor sent [an / each] unusual patient to [a / each] consultant at the local hospital but the [patient(s) / consultant(s)] [was / were] not very pleased.

26. The tutor offered [a / each] project title to [a / each] student in the practical class although the [student(s) / project(s)] [was / were] not very excited / exciting.

27. The rock star sang [a / each] love song to [an / each] attractive woman in the audience but the [song(s) / woman (women)] [was / were] very sad.

28. The magician taught [a / each] card trick to [a / each] boy in the class but the [boy(s) / card trick(s)] [was / were] not very convinced / convincing.

29. The revolutionary showed [a / each] campaign leaflet to [a / each] passing shopper, although the [leaflet(s) / shopper(s)] [was / were] not very political.

30. The student emailed [a / each] dirty joke to [a / each] lecturer in the Psychology department, but the [joke(s) / lecturers] [was / were] not very amusing / amused.

31. The lecturer taught [a / each] study method to [a / each] student in the class, but the [method(s) / student(s)] [was / were] not very successful.

32. The scientist showed [a / each] mathematical proof to [a / each] student in his class, but the [students / proof(s)] [was / were] not very clever.

33. The Professor gave [a / each] lengthy article to [a / each] project student in his group, but the [article(s) / student(s)] [was / were] very boring / bored.

34. The accountant offered [a / each] investment scheme to [an / each] finance manager in the company, but the [scheme(s) / manager(s)] [was / were] not very reliable.

35. The consultant taught [a / each] new filing system to [a / each] clerk in the factory, but the [system(s) / clerk(s)] [was / were] not very efficient.

36. The researcher posted [a / each] grant proposal to [a / each] funding agency, but the [proposal(s) / agency (agencies)] [was / were] not very inspiring.
37. The senior tutor gave [a / each] marking scheme to [a / each] module leader on the course, but the [marking scheme(s) / module leader(s)] [was / were] not very rigorous.
38. The lecturer told [a / each] filthy story to [a / each] undergraduate student in the department, but the [story (stories) / student(s)] [was / were] not very amusing / amused.
39. The alcoholic told [a / each] lengthy anecdote to [a / each] woman in the bar, but the [anecdote(s) / woman (women)] [was / were] not very interesting / interested.
40. The head chef taught [a / each] recipe to [a / each] junior cook in the hotel, although the [recipe(s) / cook(s)] [was / were] not very adventurous.
41. The General assigned [an / each] important mission to [a / each] experienced commando in the unit, but the [mission(s) / commando(s)] [was / were] not very successful.
42. The consultant sent [a / each] business colleague to [a / each] major client, although the [colleague(s) / client(s)] [was / were] not very appreciative.
43. The University offered [a / each] research grant to [a / each] senior researcher in the department, although the [grant(s) / researcher(s)] [was / were] not very prestigious.
44. The council gave [a / each] large subsidy to [a / each] farmer in the district, although the [subsidy (subsidies) / farmer(s)] [was / were] not every economical.
45. The artist gave [an / each] abstract painting to [a / each] museum curator in London, but the [painting(s) / curator(s)] [was / were] not very remarkable.
46. The Professor assigned [a / each] student to [a / each] tutor in the department, because the [student(s) / tutor(s)] [was / were] not very busy.
47. The Professor sent [a / each] research report to [an / each] academic journal in his discipline, although the [report(s) / journal(s)] [was / were] not very important.
48. The accountant showed [a / each] financial error to [a / each] police detective, and the [detective(s) / error(s)] [was / were] found out to be fraudulent.
Table 1: Predictions concerning reading time effects for the Quantified Region

<table>
<thead>
<tr>
<th>Condition</th>
<th>Surface order</th>
<th>Grammatical Hierarchy</th>
<th>Quantifier Hierarchy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ioup (1975) Grammatical Hierarchy X Quantifier Hierarchy</td>
<td>S &lt; Pl</td>
<td>S &lt; Pl</td>
<td>Pl &lt; S</td>
</tr>
<tr>
<td>‘a reporter . . . each interview’ &gt; ‘each reporter . . . an interview’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘each interview . . . to a reporter’ &gt; ‘an interview . . . to each reporter’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘a reporter . . . each interview’ &gt; ‘each reporter . . . an interview’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘an interview . . . to each reporter’ &gt; ‘each interview . . . to a reporter’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Predictions concerning reading time effects for plural (Pl) and singular (S) noun-phrases at the Continuation Region, including examples of dative and double objects sentences. Vertical lines delimit analysis regions and slashes denote alternatives.
Table 2: Mean first pass and total reading times for Regions 2 to 5, and residual first pass and total reading times for Region 3 of ‘a-each’ and ‘each-a’ dative and double object sentences with singular and plural continuations (standard errors in parentheses)

<table>
<thead>
<tr>
<th>Region</th>
<th>Measure (msec)</th>
<th>‘a-each’</th>
<th>Quantifier order</th>
<th>'each-a'</th>
<th>NP-anaphor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Singular</td>
<td>plural</td>
<td>singular</td>
<td>plural</td>
</tr>
<tr>
<td>Double object sentences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>First pass time</td>
<td>1670 (80)</td>
<td>1620 (77)</td>
<td>1529 (63)</td>
<td>1483 (72)</td>
</tr>
<tr>
<td></td>
<td>Total time</td>
<td>2884 (114)</td>
<td>2951 (145)</td>
<td>2301 (86)</td>
<td>2384 (87)</td>
</tr>
<tr>
<td>3</td>
<td>First pass time</td>
<td>340 (15)</td>
<td>381 (19)</td>
<td>345 (16)</td>
<td>369 (16)</td>
</tr>
<tr>
<td></td>
<td>Residual first pass time</td>
<td>-206 (13)</td>
<td>-171 (14)</td>
<td>-209 (11)</td>
<td>-183 (12)</td>
</tr>
<tr>
<td></td>
<td>Total time</td>
<td>538 (22)</td>
<td>591 (27)</td>
<td>473 (22)</td>
<td>520 (20)</td>
</tr>
<tr>
<td></td>
<td>Residual total time</td>
<td>-267 (23)</td>
<td>-227 (27)</td>
<td>-337 (25)</td>
<td>-293 (28)</td>
</tr>
<tr>
<td>4</td>
<td>First pass time</td>
<td>243 (11)</td>
<td>255 (12)</td>
<td>257 (12)</td>
<td>263 (15)</td>
</tr>
<tr>
<td></td>
<td>Total time</td>
<td>391 (15)</td>
<td>373 (12)</td>
<td>364 (16)</td>
<td>386 (20)</td>
</tr>
<tr>
<td>Dative sentences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>First pass time</td>
<td>1647 (82)</td>
<td>1608 (71)</td>
<td>1632 (58)</td>
<td>1718 (83)</td>
</tr>
<tr>
<td></td>
<td>Total time</td>
<td>2370 (112)</td>
<td>2471 (101)</td>
<td>2546 (114)</td>
<td>2611 (120)</td>
</tr>
<tr>
<td>3</td>
<td>First pass time</td>
<td>327 (15)</td>
<td>371 (17)</td>
<td>332 (17)</td>
<td>334 (16)</td>
</tr>
<tr>
<td></td>
<td>Residual first pass time</td>
<td>-216 (12)</td>
<td>-180 (14)</td>
<td>-213 (12)</td>
<td>-218 (14)</td>
</tr>
<tr>
<td></td>
<td>Total time</td>
<td>457 (24)</td>
<td>544 (24)</td>
<td>501 (25)</td>
<td>518 (25)</td>
</tr>
<tr>
<td></td>
<td>Residual total time</td>
<td>-341 (26)</td>
<td>-268 (24)</td>
<td>-295 (23)</td>
<td>-293 (29)</td>
</tr>
<tr>
<td>4</td>
<td>First pass time</td>
<td>242 (11)</td>
<td>262 (14)</td>
<td>255 (11)</td>
<td>258 (11)</td>
</tr>
<tr>
<td></td>
<td>Total time</td>
<td>384 (20)</td>
<td>384 (17)</td>
<td>382 (18)</td>
<td>404 (21)</td>
</tr>
</tbody>
</table>
Figure 1: First pass reading times (msec) (Panel A) and total reading times (msec) (Panel B) for the Quantified Region of dative and double object sentences with ‘a-each’ and ‘each-a’ quantifier orders.

Panel A

Panel B