

Interpretation of Quantifier Scope Ambiguities

Howard S. Kurtzman
Cornell University

Maryellen C. MacDonald
Carnegie Mellon University

A sentence like (1) exhibits a quantifier scope ambiguity:

(1) Every kid climbed a tree.

Its meaning may correspond to either of the logical structures represented in (2) and (3).

(2) $(\forall x) (\exists y) (x \text{ is a kid} \ \& \ y \text{ is a tree} \ \& \ x \text{ climbed } y)$
(=For every kid x , there is a tree y , such that x climbed y .)

(3) $(\exists y) (\forall x) (x \text{ is a kid} \ \& \ y \text{ is a tree} \ \& \ x \text{ climbed } y)$
(=There is a tree y , such that for every kid x , x climbed y .)

These logical structures differ only in the relative scopes of their quantified terms. In the interpretation in (3), there is one particular tree which every kid climbed. In (2), each kid did not necessarily climb the same tree; rather, there may have been more than one tree, with different kids climbing each tree (or perhaps just one kid per tree). The first quantified term in each ordering is said to have wide scope, the second narrow scope.

Some theoretical and computational linguists have claimed, on the basis of intuitions, that the preferred interpretation of sentences like (1) corresponds to the structure in (2) (VanLehn, 1978; Kempson & Cormack, 1981; Fodor, 1982). By "preferred" it is meant that this interpretation is the one perceivers tend to arrive at first and/or to consider more likely or appropriate. Similar claims have been made for the preferred interpretation of (4) corresponding to the logical structure represented in (5) rather than in (6):

(4) A kid climbed every tree.

(5) $(\exists x) (\forall y) (x \text{ is a kid} \ \& \ y \text{ is a tree} \ \& \ x \text{ climbed } y)$
(=There is a kid x , such that for every tree y , x climbed y .)

(6) $(\forall y) (\exists x) (x \text{ is a kid} \ \& \ y \text{ is a tree} \ \& \ x \text{ climbed } y)$
(For every tree y , there is a kid x , such that x climbed y .)

Fodor (1982) has suggested that these preferences stem from a processing strategy according to which the preferred scope ordering of quantified noun phrases (NPs) is the same as the surface ordering of the NPs in the sentence. Such a strategy could be plausibly implemented by the construction of the initial meaning representation proceeding incrementally as the sentence is continuously processed word by word or phrase by phrase. This surface ordering strategy is essentially a claim concerning on-line processing, yet the only available data in support of this claim are off-line intuitions. There is thus no evidence concerning how strong these preferences are, whether factors other than surface

ordering influences interpretation, or whether the ultimately-preferred interpretations are the only ones available during initial processing of these sentences.

The present experiment seeks to obtain more direct evidence concerning which interpretations are available during processing of a scope ambiguity. The approach is to have subjects first read a scope ambiguous sentence and then read a sentence which is consistent with one interpretation of the ambiguity. The subjects' task is to judge whether the second sentence is compatible with the first. For example, following (1), (7) is (more) consistent with the interpretation corresponding to (2), and (8) is consistent with the interpretation corresponding to (3).

(7) The trees were full of apples.

(8) The tree was full of apples.

Similarly, following (4), (9) is consistent with the interpretation corresponding to (5), while (10) is (more) consistent with the interpretation corresponding to (6).

(9) The kid was full of energy.

(10) The kids were full of energy.

By examining subjects' responses and response times to these second sentences, it is possible to infer which interpretation(s) subjects give the first sentence. By encouraging subjects to decide as soon as they have read both sentences, data can be obtained in a more on-line task, without calling attention to the ambiguities.

METHOD

Subjects

Twenty-four college students in the Pittsburgh area were paid to participate in the experiment. All subjects were native speakers of English.

Materials

All 32 experimental passages contained 2 sentences. The first sentence was five or six words in length and contained two quantifiers. The second sentence contained six words and described one of the nouns mentioned in the first sentence.

Four variables were manipulated in the passages, each with two levels. The levels of Quantifier order in the first sentence were "every...a", as in sentence (1), and "a...every" as in (4). The first sentence Verb described either an action such as "climbed," in which case the subject of the sentence receives an agent thematic role, or a perception such as "saw", where the subject has an experiencer thematic role. The scope Interpretation variable was manipulated in the second sentence of the passage: the noun in this sentence was either singular, as in (8-9) or plural, as in (7) and (10). Finally, Ambiguity was manipulated in the first sentences by the substitution of phrases "the same" or "a different" for the quantifier "a". The unambiguous versions of (1) are illustrated in (11-12) with their

appropriate second sentence. Note that the singular/plural levels of the Interpretation variable in the second sentence dictate whether "the same" or "a different" is the disambiguating phrase in the first sentence.

- (11) Every kid climbed the same tree.
The tree was full of apples.
- (12) Every kid climbed a different tree.
The trees were full of apples.

The "a...every" structure illustrated in (4) was made unambiguous in similar fashion: "The same/a different kid climbed every tree".

Most of the 80 practice and filler passages contained a second sentence incompatible with the first, so that overall, the correct response on the compatibility judgment was "yes" on 60% of the trials (all experimental passages were compatible). A number of syntactic, semantic, and pragmatic violations were used to create incompatible fillers.

Procedure. Passages were presented one sentence at a time on a CRT. Subjects pressed a key once to display the first sentence and again to remove this sentence and display the second. Subjects pressed keys marked YES or NO to indicate whether the second sentence was compatible with the first. Both speed and accuracy were stressed in reading and for the compatibility judgment. Following 20 practice trials, all experimental and filler trials were presented in random order.

RESULTS

Compatibility judgments. As interpretation of the ambiguous sentences is of primary interest here, compatibility judgments for only the ambiguous sentences are reported in Table 1. Passages with unambiguous sentences were judged to be consistent 82.3% of the time, significantly more often than passages with ambiguous first sentences, $\min F'(1, 44) = 23.93$, $p < .001$. This result indicates that subjects did not always have both meanings available for the ambiguous sentences.

The surface ordering strategy (Fodor, 1982) would predict that for ambiguous "every...a" sentences, the plural continuation would be preferred, while the singular continuation should be preferred for the ambiguous "a...every" sentences.¹ The data in Table 1 strongly support this prediction. For the ambiguous sentences, there was a significant Quantifier Order x Interpretation interaction, $\min F'(1, 51) = 62.08$, $p < .0001$. In contrast, there was no evidence that the surface ordering of the quantified NPs affected the interpretation of the unambiguous sentences: this interaction was not reliable for unambiguous sentences, $\min F'(1, 47) = 2.69$, $p > .10$. The disambiguating information in these sentences apparently overrode any surface ordering strategy.

TABLE 1
 Percentage of Ambiguous Passages Judged Compatible. Experiment 1

Interpretation	"Every...a" sentences	
	Action Verb	Perception Verb
Singular	25.0	62.5
Plural	81.3	72.9

Interpretation	"A...every" sentences	
	Action Verb	Perception Verb
Singular	87.5	85.4
Plural	14.5	14.5

Further examination of Table 1 reveals that for the "every...a" sentences, the action/perception verb manipulation dramatically affected acceptance rates. Subjects judged the singular interpretation of these sentences to be more acceptable when the first sentence had contained a perception verb, compared to when it had contained an action verb, $\min F(1, 21) = 6.87, p < .025$. We return to this point in the discussion.

Because the acceptance rates for the non-preferred interpretations of ambiguous sentences are so low, response times to make this compatibility judgment are difficult to interpret. However, we should note that on those trials where subjects judged the second sentence to be compatible with the first, response times for passages with an ambiguous first sentence (3752 ms) were longer than for unambiguous passages (3073 ms). This result suggests that the disambiguating information helped subjects arrive at an interpretation more rapidly. Despite the fact that these response times (which include time to read the second sentence) are rather long, this paradigm still provides a much more immediate measure of how perceivers interpret these ambiguities than can be gleaned from intuitions. Additionally, few subjects reported that they noticed any ambiguities in the stimuli, indicating that the task successfully prevented subjects from consciously searching for multiple meanings for the sentences.

DISCUSSION

This experiment demonstrated that subjects could rapidly interpret scope ambiguous sentences, and that their preferred interpretations were the ones where the first-encountered NP had wide scope over the second NP. Fodor's (1982) surface ordering proposal is clearly compatible with these data, but it is not the only interpretation available. Kempson & Cormack (1981), for example, have suggested that the preferred interpretation is the one

where the sentence topic (usually the first NP) has wide scope.

Pragmatic Influences. The data also suggest that, while the surface-based scope ordering is preferred, it is not always the only one that subjects initially assign. In particular, for the "every...a" sentences with perception verbs, subjects judged both singular and plural continuations to be compatible over 50% of the time. For this sentence type, then, it appears that subjects assigned both interpretations, although the interpretation incorporating the surface-based scope ordering (the plural continuation) was still preferred.

Why would subjects be more likely to assign both interpretations for this sentence type? First, compare the "every...a" sentences with action verbs and with perception verbs. It is quite plausible, within a single event (taken to be the referent of each of these sentences), for more than one perceiver to perceive the same entity, but it is generally somewhat less plausible for more than one agent to perform an action on the same entity. To take an extreme example not used in the experiment, it is plausible for many children to see one unique cookie, but it is pragmatically very awkward to interpret "Every child ate a cookie" as meaning that all the children jointly ate one unique cookie. Our action verb sentences (e.g. (1) and (4)) were intended to be plausible with both singular and multiple entities (e.g. trees), but the data suggest that subjects did not find them equally plausible. Thus, at least for the stimuli tested here, the interpretation in which the second NP (with the determiner "a") has wide scope is pragmatically more appropriate for sentences with perception verbs than with action verbs.

Now compare the "every...a" sentences with perception verbs with the "a...every" sentences with either type of verb. When "a...every" sentences are presented with no preceding context, as in this experiment, the subject, upon receiving the first NP (with the determiner "a"), has no reason to consider the phrase as referring to more than one entity: in fact, it would be pragmatically unparsimonious for the subject to do so (cf. Crain & Steedman, 1985). It is only when the second NP (with the determiner "every") is received that the real potential arises for the first NP to refer to multiple entities. This potential is correlated with the first NP being assigned narrow scope (i.e., with the scope ordering which does not match surface ordering). But a changed reference for the first NP, from singular to potentially multiple, is pragmatically quite significant. It is reasonable to assume that a change in the reference tends not to be determined except when all possible interpretations require it, due to the semantic processing complexities which the new reference necessitates (cf. Gillon, 1984; Crain & Steedman, 1985). Therefore, it is unlikely that the alternative interpretation for the "a...every" sentences--in which the first NP has narrow scope--is routinely assigned.

By contrast, for the "every...a" sentences, the first NP received (with the determiner "every") inherently refers to multiple entities. This is the case regardless of whether the second NP (with the determiner "a") is assigned narrow or wide scope. Because no

pragmatically significant changed reference for the first NP needs to be determined even if the alternative interpretation is assigned. Multiple interpretations can be more readily assigned to the "every...a" sentences than to the "a...every" sentences.

A Processing Model. Precisely how do the pragmatic influences interact with the basic scope ordering preference? One possibility is that for all scope ambiguous sentences, both possible interpretations are determined immediately (i.e., upon receiving the second NP, which is the point where the ambiguity arises). These interpretations vary in their levels of accessibility, based primarily upon the preference for surface-based scope ordering. Thus, following Fodor (1982), the more accessible interpretation would be the one whose scope ordering is more quickly or directly identified in continuous processing of the sentence. Or, following Kempson & Cormack (1981), it would be the one in which the sentence topic has wide scope.

The less accessible interpretation can, however, have its accessibility increased somewhat, on pragmatic grounds. Thus, if the less accessible interpretation is highly plausible and if it does not involve a changed reference for the first NP (with respect to the reference in the other, more accessible interpretation), then its accessibility can be increased, although not to the level of the other interpretation. This is the case for the "every...a" sentences with perception verbs.

Next, all interpretations whose accessibility is below a criterion level are deleted. This deletion would occur sometime between processing of the second NP and evaluation of the continuation sentence. The remaining interpretations and their accessibility levels determine how frequently subjects respond that each of the continuation sentences is reasonable. (See Kurtzman, 1985; Gorrell, 1987, for similar views of the processing of syntactic ambiguities.)

Obviously, further work is required to elaborate and experimentally test this model. Also, further work is needed to determine the nature of the mental representation of scope ordering. Among the options are: standard logical formulae (Hobbs & Shieber, 1987), indexed syntactic structure (Williams, 1988), a syntactic level of LF (May, 1985), mental and discourse-representation models (Fodor, 1982; Kamp, 1984), and propositional networks (Anderson, 1983). In a first step toward further specification of this model, we are currently conducting an experiment with passive versions of the stimuli, such as:

(13) A tree was climbed by every kid.

(14) Every tree was climbed by a kid.

Actives and passives share the same propositional meaning, but differ in the surface ordering of their NPs. Therefore, passives provide a sharp test of the claim that surface-based scope ordering is preferred. The experiment also allows further examination of pragmatic influences on interpretation, in that it provides additional opportunity for examination of the verb effects found in the present experiment. The results will be

reported in a future paper.

FOOTNOTE

1. Strictly, the interpretation in which the phrase with the determiner "a" has narrow scope is compatible with either the singular or the plural continuation sentence. That is, the logical structures in (2) and (6) specify only that it is possible--not that it is necessary--for there to be more than one tree and more than one kid. Thus, one might challenge our claim that preference for the singular interpretation for the "a...every" sentences indicates a preference for wide scope of the first NP. However, the low percentage of YES responses for the plural continuation sentences suggests that the other interpretation (narrow scope on the first NP) was not preferred.

Focussing now on the "every...a" sentences, the high percentage of YES responses to the plural continuation sentences is compatible only with a preference for narrow scope for the "a" phrase. However, as just noted, a singular reference for the "a" phrase is possible with narrow scope. Why then is the percentage of YES responses for the singular continuations lower (especially for the action verb sentences)? Apparently, when an interpretation permits either multiple or singular reference, multiple reference is preferred. Fodor (1982) has suggested how this could be accounted for if quantifiers are represented within the format of mental models. Clearly this is an area for additional investigation.

REFERENCES

- Anderson, J.R. (1983). The architecture of cognition. Cambridge, MA: Harvard Univ. Press.
- Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), Natural language processing. Cambridge: Cambridge Univ. Press.
- Fodor, J.D. (1982). The mental representation of quantifiers. In S. Peters & E. Saarinen (Eds.), Processes, beliefs, and questions. Dordrecht: Reidel.
- Gillon, B.S. (1984). The logical form of plurality and quantification in natural language. Ph.D. Dissertation, M.I.T.
- Corrill, P.G. (1987). Studies of human syntactic processing: Ranked-parallel versus serial models. Ph.D. Dissertation, University of Connecticut.
- Hobbs, J.R., & Shieber, S.M. (1987). An algorithm for generating quantifier scopings. Computational Linguistics, 13, 47-63.
- Kamp, H. (1984). A theory of truth and semantic representation. In J. Groenendijk, T.M.V. Janssen, & M. Stokhof (Eds.), Truth, interpretation, and information. Dordrecht: Foris.
- Kempson, R.M., & Cormack, A. (1981). Ambiguity and quantification. Linguistics and Philosophy, 4, 259-309.
- Kurtzman, H.S. (1985). Studies in syntactic ambiguity resolution. Ph.D. Dissertation, M.I.T. (Distributed by Indiana University Linguistics Club.)
- May, R. (1985). Logical form. Cambridge, MA: MIT Press.
- VanLehn, K.A. (1978). Determining the scope of English quantifiers. Artificial Intelligence Laboratory Report AI-TR-483. M.I.T.
- Williams, E.S. (1988). Is LF distinct from S-Structure? A reply to May. Linguistic Inquiry, 19, 135-146.