

Integrity: A Syntactic Constraint on Quantificational Scoping

Chris Barker
University of California

This paper proposes that quantifier scope alternations respect syntactic constituency in a certain way:

- (1) **The syntactic constituent integrity scoping constraint** (**'Integrity'**): if there is a syntactic constituent that contains B and C but not A, then A must take scope over both B and C or neither.

The claim is that the connection between syntactic structure and quantifier scope is considerably tighter than most current theories guarantee.

One notable exception is the version of Combinatory Categorical Grammar developed in Park (1995) and Steedman (2000). They argue that the connection between syntax and quantifier scope is even closer, in fact, as close as possible: that quantifier scope is completely and uniquely determined by syntactic structure (up to but not including certain types of indefinite; see sections 1 and 3 below). Integrity strikes an intermediate position, stricter than most theories, but, unlike Park and Steedman, allowing a certain limited range of alternation for a given syntactic structure.

For instance, Integrity predicts that the following sentence will have four possible scope construals rather than the traditional six:

- (2) Most subjects [put an object in every box].
 - a. most > an > every
 - b. most > every > an
 - c. every > an > most
 - d. an > every > most
 - e. an > most > every (violates Integrity)
 - f. every > most > an (violates Integrity)

*Thanks for comments to Mark Gawron, Svetlana Godjevac, and Mark Steedman.

Assuming that the bracketed substring is a constituent, Integrity predicts that the quantifier denoted by the subject cannot intervene scope-wise between the quantifiers denoted by the direct object and the locative. As a result, the scoping possibilities indicated in (2e) and (2f) are incompatible with Integrity. I will argue that with one important caveat concerning indefinites (see section 1), this prediction is borne out.

In addition to the work of Park and Steedman, Integrity effects have been noticed before, but not in their full generality. Section 3 discusses how May (1985) and Hobbs and Shieber (1987) approximate Integrity for a restricted set of cases.

The need to represent quantifier scope relations that diverge from surface syntactic relationships remains one of the main motivations for positing a level of Logical Form distinct from surface syntax. The thrust of Integrity is that quantifier scoping sticks much closer to surface constituency than suggested by traditional theories of quantifier scope. To the extent that recognizing Integrity enables in-situ theories of quantifier scope to correctly describe the facts, we are one step closer to eliminating Logical Form as an indispensable level of grammatical representation.

1. Wide-scope indefinites are different

Before discussing Integrity per se, it is necessary to provide a disclaimer concerning certain types of indefinites. There is a long tradition of treating all uses of indefinites as a species of scope displacement, despite well-known troublesome examples like (3), due to Fodor and Sag (1982).

(3) Each dean heard a rumor that a student of mine cheated.

The standard judgment is that there is a reading of (3) on which the choice of a student seems to take scope over *each*, despite the fact that quantifiers in embedded clauses normally cannot take scope outside of their minimal clause. However, work of Farkas, Fodor and Sag, Abush, Park, Steedman, Reinhart, Kratzer, Winter, and others converge on the conclusion that a class of uses including the so-called de dicto/de re ambiguity, specific indefinites, and other indefinites with unusually wide scope should not be lumped together with normal quantifier scope. Kratzer's (1998:167) assessment is that "indefinite NPs are ambiguous between a specific and a quantificational interpretation. If they are quantificational, their scope is local, and they are interpreted as generalized quantifiers, like any other quantifier phrase." If specific, they are treated dif-

ferently, either as referential (e.g., Fodor and Sag 1982), or as a context-controlled choice function (e.g., Kratzer 1998). I will refer to all of these special indefinites using the cover term ‘wide-scope indefinites’, and I will assume that wide-scope indefinites receive their interpretation through mechanisms other than normal quantifier scope, and therefore do not fall under the purview of Integrity.

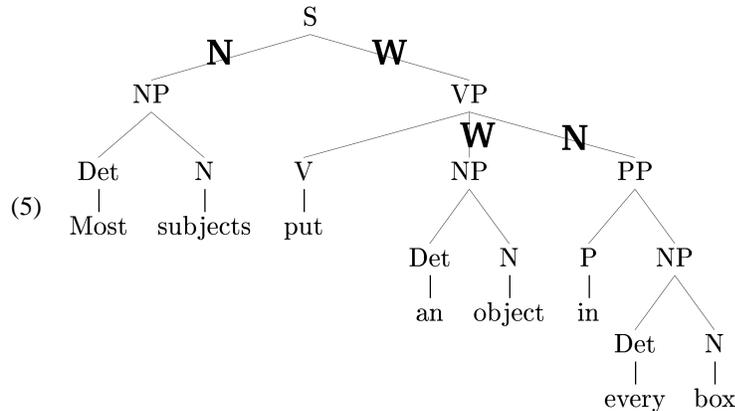
As a practical matter in this paper, I will avoid using examples that depend on giving indefinites wide scope, since such examples can always be explained away by appeal to whatever mechanisms explain wide-scope indefinites. Readings on which an indefinite takes narrow scope remain pertinent, however, since by hypothesis indefinites can only take narrow scope by participating in normal scope construal.

2. The basic prediction: four readings, not six

In order to make the nature of a theory that obeys Integrity more concrete, consider the following simplistic method for predicting possible quantifier scopings:

- (4) For each node in a (surface) syntactic tree, order the daughters in terms of scoping priority.

For instance, simplifying the syntax for exposition, for sentence (2) we might have:



In the tree in (5), the subject branch is marked ‘N’ (for ‘Narrow’), so the subject takes narrow scope with respect to the verb phrase. But since the the di-

rect object branch is marked ‘W’ (for ‘Wide’), it takes wide scope with respect to the locative. Only quantificational elements matter for purposes of quantifier scope, of course, so the resulting scoping for the sentence as a whole is: *an object* takes scope over *every box*, and both take scope over *most subjects*. Obviously, if more than two sister branches contain quantificational elements, a more elaborate labelling system than ‘N’ and ‘W’ will have to be developed. This algorithm may remind some people of the method for determining primary and secondary stress in metrical phonology, as in, e.g., Liberman and Prince (1977).

This method provides exactly four possible scopings for the sentence in (2). The four readings correspond to different ways of answering the questions ‘Does the subject take scope over the VP?’ and ‘Does the locative take scope over the direct object?’. Most theories of quantification, including all Quantifier Raising theories, provide at least 6 scopings, the full factorial range for distinct orderings of sets containing 3 elements. The reason is that the three quantificational NPs can raise to adjoin to S in any order, yielding at least 6 logically distinct LF structures. See, e.g., Heim and Kratzer (1998), chapters 6 and 7, for details.

Note that the predictions of traditional theories and of Integrity coincide when there are only two quantificational expressions. No matter what the syntactic constituency, it will always be possible for a branch containing one element to take priority over some branch containing the other element, and vice versa. Therefore Integrity can only be distinguished from traditional theories when there are three or more quantificational elements.

The scopings ruled out by Integrity are the scopings in which the subject intervenes in scope between the direct and locative. The claim at hand is that (2) has neither of the truth conditions given in (6).

- (6) a. $\exists z.\mathbf{object}(z) \& \mathbf{most}(\mathbf{subjects}, \lambda y \forall x.\mathbf{box}(x) \rightarrow \mathbf{put}(y, z, x))$
 b. $\forall x.\mathbf{box}(x) \rightarrow \mathbf{most}(\mathbf{subjects}, \lambda y \exists z.\mathbf{object}(z) \& \mathbf{put}(y, z, x))$

These logical expressions simply give the truth conditions corresponding to the scopings indicated in (2e) and (2f).

The truth conditions for (6a) give the existential widest scope. As explained in section 1, even if we observe that (2) has a reading with truth conditions equivalent to (6a) (and it probably does), we would have to wonder whether this reading arises through whatever special mechanism gives rise to wide-scope indefinites.

Therefore concentrate on the truth conditions in (6b). Imagine using (2)

to describe an experiment in which three experimental subjects (call them S1, S2, and S3) are required to place each of two objects (an apple and a pen) into any one of three boxes (labelled B1, B2, or B3). The truth conditions in (2b) are satisfied in a situation in which the put relation is as given in (7).

- (7) <S1, apple, B1> <S2, apple, B2> <S3, apple, B3>
 <S1, pen, B2> <S2, pen, B3> <S3, pen, B1>

The truth conditions for the alleged scoping are satisfied, since for every box, most subjects placed an object in that box. Note that since there are only two objects but three boxes, it is impossible for any individual subject to have placed an object in every box.

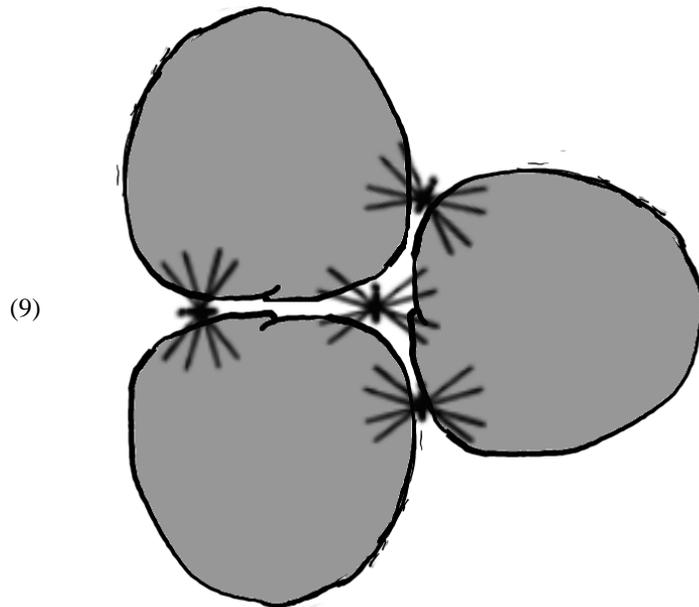
Native speakers consistently deny that (2) can be true in such a situation.

Another relevant example that can be considered with respect to the same set of circumstances is:

- (8) a. *Every subject put an object in no box.*
 b. $\neg \exists b : \mathbf{box}(b) \& \forall s : \mathbf{subject}(s) \rightarrow (\exists o : \mathbf{object}(o) \& \mathbf{put}(s, o, b))$

Integrity predicts that (8a) cannot have the truth conditions described in (8b). If (8b) were a possible reading, (8a) should be capable of being true in the situation given in (7). Once again, most native speakers confirm this prediction.

Presenting the question visually can help sharpen intuitions:



- (10) a. Most spiders [put a foot on every lily pad].
 b. Most spiders have [a foot on every lily pad].

The issue is whether the information in the picture in (9) by itself is sufficient to prove that (10a) is true. That is, does (10a) have a reading that can be paraphrased as *every lily pad is such that most spiders put a foot on it*? After all, for each lily pad, three out of four spiders have at least one foot resting on that lily pad. But if *put a foot on every lily pad* is a constituent, and if English obeys Integrity, then there should be no construal of (10a) that is verified by (9). By way of comparison, all theories, including Integrity, agree in predicting that (10b) cannot be true, since the NP *a foot on every lily pad* is a scope island (see section 3 below).

I must report that a small minority of native speakers claim that the disputed truth conditions are possible, though perhaps dispreferred. Such speakers are invariably highly-trained semanticists with extensive experience calculating the predictions of the theories that insist on six scopings. If there are in-

deed native speakers who competently judge (10a) felicitous and true but (10b) false in the situation depicted in (9), then there are a number of possibilities:

- (11)
- Integrity doesn't exist.
 - Integrity does exist, but it is merely a strong preference, and does not align perfectly with grammatical interpretations.
 - Integrity only governs one type of quantifier scope, and there may be other mechanisms that can provide extra scope readings (as in section 1).
 - There is an alternative syntactic structure for (10a) that groups the subject and the direct object together but excludes the locative, in which case the disputed interpretation is compatible with Integrity (as in section 4).

A word of warning is in order concerning the difficulty of constructing counterexamples to Integrity. First, as explained above in section 1, in order for a putative counterexample to be compelling, it must not depend on indefinites taking wide scope, since there is good reason to believe that indefinites are able to take wide scope through a separate mechanism. In the same spirit, though perhaps even more controversially, other quantificational elements including cardinals, *few*, *many*, *most*, and others now have analyses on which they can denote choice functions that supply a mereological sum as a value (often called a 'witness set'—see, e.g., Beghelli et al. 1997). This complex individual then enters into entailments governed by the semantics of plurals and distributivity without requiring scope alternations. Therefore the persuasiveness of a counterexample to Integrity depends on the degree to which such alternative non-scope explanations can be ruled out.

3. Special cases of Integrity from previous literature

Integrity effects have in fact been noticed, but not in their full generality. The most widely accepted result involves quantifiers embedded inside NPs. May 1985 and Larson (in unpublished work described in Heim and Kratzer 1998:233) conclude that NP is a scope island: an NP embedded within an NP may move only as far as its containing NP. In addition to preventing unbound traces from arising due to the technical details of quantifier raising, they note that this assumption makes some good predictions with respect to the observed

range of scope orderings:

- (12) a. Two politicians spy on [someone from every city].
 b. *every city > two politicians > someone

The sentence (12a) does not have the scoping indicated in (12b). Preventing *every city* from escaping from the object NP prevents the subject from intervening between the indefinite and the universal. Obviously, since *someone from every city* is a constituent on any syntactic analysis, this fact follows directly from Integrity as well.

As a second example of a scope theory that embodies a part of Integrity, Hobbs and Shieber (1987) provide an LF-based scoping algorithm which gives (12a) 5 scopings instead of the full factorial 6. The missing reading is the one where *someone* outscopes *two* which outscopes *every*. As Hobbs and Shieber 1987:48 put it, “in general, a quantifier from elsewhere in the sentence cannot come after the quantifier associated with a head noun and before the quantifier associated with a noun phrase in the head noun’s complement.” This characterization is strikingly similar to the Integrity constraint in spirit, and the class of scopings ruled out for structural reasons by Hobbs and Shieber’s algorithm seems to be a subset of the scopings ruled out by Integrity.

However, when none of the quantificational elements contains another (e.g., as in (2)), Hobbs and Shieber’s algorithm can deliver the full factorial range of scopings. In particular, they provide all 6 possible scopings for (2) instead of the 4 consistent with Integrity. In fact, even when one quantificational NP contains another, Hobbs and Shieber’s algorithm generates some readings that violate Integrity, including the reading given above as (12b). Thus Integrity makes significantly stronger empirical claims about structurally illicit scopings than Hobbs and Shieber’s algorithm.

4. Hyperintegrity: Combinatory Categorical Grammar

Working with Combinatory Categorical Grammar (CCG), Park (1995) sets out to describe a way in which quantifier scopings are ‘constrained by surface constituency’. As Steedman (2000:72) puts it, building on Park’s work, ‘semantic quantifier scope is limited by syntactic derivational scope’. Park (1996b:4) gives a precise characterization of possible scope alternations that resembles but does not coincide with Integrity (his characterization is sufficiently complex that I will not attempt to reproduce it here). Nevertheless, it is clear that

Park and Steedman's system obeys Integrity, and that their approach is very much in the same spirit as Integrity.

In fact, in addition to obeying Integrity, their analysis is more restrictive than required by Integrity. Note that the simple theory given by (4) is the most liberal theory consistent with Integrity, in the sense that it will generate every possible scoping that is consistent with Integrity. As shown for (4), Integrity allows for more than one scoping for a given syntactic structure. In Park and Steedman's system, however, syntactic structure completely determines quantifier scoping (again, with a major exception concerning indefinites, see immediately below). More specifically, in their system expressions that function semantically as functors take scope over expressions that serve as arguments.

However, because there is considerable flexibility in the way in which the elements of a sentence combine to form larger phrases in CCG, sentences can in effect be syntactically ambiguous in a way that affects quantifier scope relations. For instance, the sentence *Every boy saw a girl* can have the traditional syntactic/functional analysis on which the transitive verb combines first with the direct object and then with the subject, namely [[every boy] [saw [a girl]]]; or else the verb can combine first with its subject and then with its direct object, to give the syntactic structure [[[every boy] saw] [a girl]]. In the first case, the subject denotes a function that takes the verb phrase meaning as an argument, resulting in wide scope for the subject. In the second case, the direct object denotes a function that takes the property denoted by the subject composed with the verb as its argument, resulting in wide scope for the direct object. See Steedman (2000, especially chapter 4) for details.

Thus returning to the example in (2), CCG provides at least six distinct syntactic analyses, depending on the order in which functions are composed and applied. Therefore despite obeying Integrity, CCG allows for all six logically distinct quantifier scopings, precisely because CCG contemplates an unusually wide array of syntactic structures.

One place where Park and Steedman call into question their commitment to Integrity concerns examples like (13). Once again, the trouble has to do with indefinites taking unexpected scopes, but interestingly, Park and Steedman have the opposite problem of wide-scope indefinites discussed in section 1, i.e., it is the narrow scope reading that gives them trouble:

(13) [Every boy likes and every girl hates] some saxophonist.

This is an example of so-called non-constituent coordination, in which *every boy likes* coordinates with *every girl hates*. Of course, in CCG, these expressions are in fact constituents, which is precisely why they are able to coordi-

nate. In addition, the coordinate structure itself must be a constituent, which means that there is no room for manipulating syntactic structure to provide alternate scopings. Their analysis further specifies that the object NP *some saxophonist* serves as a functor taking the property denoted by *every boy likes and every girl hates* as an argument, predicting that (13) should only have the reading on which *some saxophonist* takes widest scope. There is such a reading, but there is also a reading on which *some saxophonist* takes narrow scope, in which case there may be a different saxophonist hated by each boy and girl.

Building on work of Fine, Park and Steedman propose that *some saxophonist* denotes an ‘arbitrary object’, modeled as a Skolem term. According to Steedman (2000:76), Skolem terms are closely related to, but distinct from choice functions, and are also similar to but distinct from E-type pronouns as formalized using Skolem functions by Heim (1990). On the arbitrary object analysis, the variability in the interpretation of (13) stems from the point in the derivation at which the direct object is Skolemized. If the direct object is Skolemized before combination with the coordinate structure, object wide scope reading results; if it Skolemizes after combining with the coordinate structure, it is within the scope of the universal operators, in which case the narrow scope reading results.

Whatever the virtues of the arbitrary object analysis of (13), it is worth noting that both sets of truth conditions as described by Park and Steedman fall out from Integrity without needing to appeal to arbitrary objects: the direct object either takes scope over both of the quantificational expressions in the coordinate structure, or neither.

5. Theoretical implications

The need to represent the full range of quantifier scopings is one of the leading motivations for theories that posit a level of Logical Form distinct from surface structure. Unfortunately, allowing LF to differ significantly from surface syntax severely weakens the empirical force of the Principle of Compositionality (this point is developed at some length in Barker 2000), so all else equal, we should strongly prefer a theory that does without LF. If Integrity is correct, and if there are compositional theories that respect Integrity without resorting to LF, then recognizing Integrity calls into question the need for LF as a distinct level of grammatical representation.

As discussed above, CCG is one compositional theory that respect Integrity. In addition, in a recent manuscript (Barker 2000), I propose a theory of quantification based on continuations, a technique adapted from the formal

semantics of programming languages. That theory has Integrity for a theorem, at least for normal quantifier scope interactions (there is a separate mechanism for wide-scope indefinites). Unlike CCG, the continuations theory allows for the full range of scopings compatible with Integrity for each syntactic analysis.

It is no accident that the two theories of quantification that approximate Integrity most closely—my continuation-based theory and Park and Steedman’s CCG—are “in-situ” theories, i.e., quantifiers are interpreted in their syntactic position, without recourse to quantifier raising, quantifying in, storage, or other mechanisms for constructing a level of Logical Form (LF) distinct from surface structure.

Interestingly, however, not all in-situ theories obey Integrity. Hendriks’ (1988, 1993) Flexible Types theory is in-situ in the relevant sense, and strictly compositional (no level of LF), though it does rely on powerful type-changing schemata. However, even assuming a single syntactic structure for (2), the Flexible Types system is powerful enough to provide all six scopings even holding syntactic structure constant (see, e.g., (12c) in Hendriks 1988:113).

Therefore if further research supports the claim that Integrity is empirically valid, it not only argues in favor of in-situ theories of quantification over LF theories, it argues in favor of a particular class of in-situ theories including at least CCG and continuations.

References

- Barker, Chris. 2000. Continuations and the nature of quantification. UCSD manuscript.
- Beghelli, Fillippo, Dorit Ben-Shalom, and Anna Szabolcsi. 1997. Variation, distributivity, and the illusion of branching. In Anna Szabolcsi, ed. *Ways of Scope Taking*, 29–69.
- Fodor, J., and I. Sag. 1982. Referential and Quantificational Indefinites. *Linguistic Inquiry* 5:355–398.
- Heim, Irene. 1990. E-Type Pronouns and Donkey Anaphora. *Linguistics and Philosophy* 13:137–177.
- Heim, Irene and Angelika Kratzer. 1998. *Semantics in Generative Grammar*, Blackwell, Oxford.
- Hendriks, Herman. 1988. Type Change in Semantics: the Scope of Quantification and Coordination, in E. Klein and J. van Benthem, eds., *Categories, Polymorphism and Unification*, ITLI, Amsterdam, 96–119.
- Hendriks, Herman. 1993. *Studied Flexibility*, ILLC Dissertation Series, Amsterdam.
- Hobbs, Jerry and Stuart M. Shieber. 1987. An algorithm for generating quantifier scopings. *Computational Linguistics* 13:47–63.
- Kratzer, Angelika. 1998. Scope or Pseudoscope? Are there wide-scope indefinites? In

- Susan Rothstein, ed., *Events and Grammar*, Kluwer, Dordrecht, 163–196.
- Liberman, M. and A. Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* **8.2**: 249–336.
- May, Robert. 1985. *Logical Form: Its Structure and Derivation*, MIT Press, Cambridge, Massachusetts.
- Park, Jong Cheol. 1995. Quantifier Scopep and Constituency. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*. Cambridge, MA, 205–212.
- Park, Jong Cheol. 1996a. *A Lexical Theory of Quantification in Ambiguous Query Interpretation*. PhD Thesis, University of Pennsylvania, Philadelphia, PA.
- Park, Jong Cheol. 1996b. Quantifier Scopep, Lexical Semantics, and Surface Structure Constituency. IRCS Report 96-28.
- Steedman, Mark. 2000. *The syntactic process*. MIT Press.