The Pragmatics of Quantifier Scope: A Corpus Study

I. Introduction Most investigations of quantifier scope are concerned with the range of *possible* scopes for sentences with multiple quantifiers like (1)-(2) below. Instead, this study examines the *actual* scopes, i.e., the pragmatics of quantifier scope disambiguation. Although actual usage facts fall outside of semantics proper, we are interested in them because they too are facts about natural language interpretation. Moreover, they are facts which provide indirect evidence about the semantics of quantifiers, e.g. if *some* has a stronger preference for narrow scope than *a*, then their semantic representations should arguably be different.

1. Each**&1_S_each#** tape is to be assigned to a different**&2_to_a.different#** time slot.

2. Each**&1_S_each#** professor has one or more**&2_O_one.or.more#** specialities.

Based on introspective judgments and previous work (e.g. [**1**], [**5**], [**4**]), we expect structural factors such as linear order (LIN.ORD) and grammatical function (GRAM.FUN) to influence scope. The lexical realization of quantifiers (LEX.REAL) has been overlooked in computational work (e.g. [**2**]), but we also expect it to contribute to scope disambiguation (e.g. [**3**]).

While introspection can provide evidence for these 3 individual factors, it cannot readily address the interactions *between* them. For example, (***i***) does GRAM.FUN affect preference for wide scope independently of, i.e., while controlling for, LIN.ORD – and vice-versa? This is particularly hard to answer in English since LIN.ORD and GRAM.FUN are highly correlated. In addition, (***ii***) do the two factors LIN.ORD and GRAM.FUN interact when they affect wide-scope preferences? Finally, (***iii***) does the lexical realization (LEX.REAL) of a quantifier, e.g., *each* vs. *every* vs. *all* have a significant contribution to preference for wide scope independently of LIN.ORD and GRAM.FUN? These questions are impossible to answer by introspection and impractical to answer via psycholinguistic methods unless the investigation is restricted to particular contrasts, e.g. focusing solely on *every* vs. *each* or on *a* vs. *some*.

II. The Corpus To address these questions, we assembled a corpus of sentences from LSAT logic puzzles and tagged it for quantifier scope. Logic puzzles are ideal for such an investigation because sentences with two or more quantifiers are frequent in this register. In addition, ambiguity must be minimal because test takers are expected to select a single correct answer. Finally, the LSAT explicitly states assumptions that might be left to world knowledge or shared discourse context in ordinary conversation. This aspect of the corpus allows us to control for the role of world knowledge to a great extent.

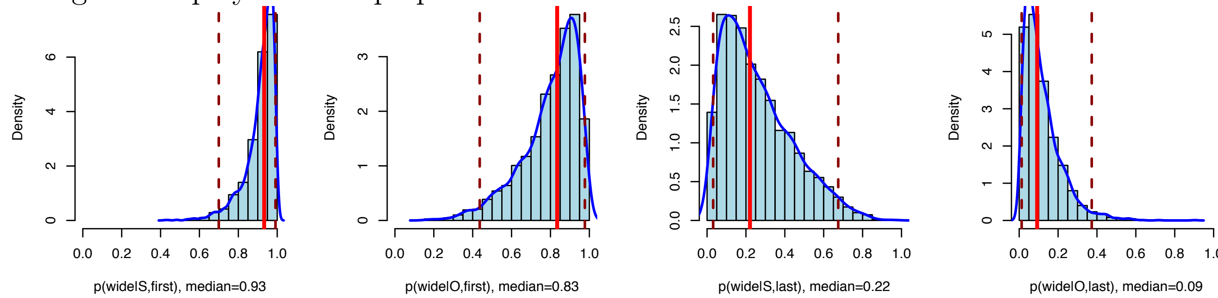
As shown in (1)-(2) above, each quantifier in a multiple-quantifier sentence was tagged for: (***i***) SCOPE – $1, 2, \ldots$; in cases where no truth-conditional difference was clear, the felicity of "such that" paraphrases provided the ultimate criterion; (***ii***) GRAM.FUN – we distinguished Subject, Object, Pivot and Adjunct, as well as individual prepositions; (***iii***) LEX.REAL – we tagged the entire complex determiner in cases like *more.than.two* or *a.different*. Though not explicitly tagged, LIN.ORD was recoverable from relative tag order.

III. Modeling and Results We focus on sentences where: there are exactly two quantifiers, the quantifiers interact scopally (i.e., no cumulative readings), and at least one is S or O. Given that some sentences have both an S and an O quantifier and the scope of one completely determines the scope of the other, we randomly sample one quantifier from each of these sentences to avoid double counting. Final dataset: 348 quantifiers / observations.

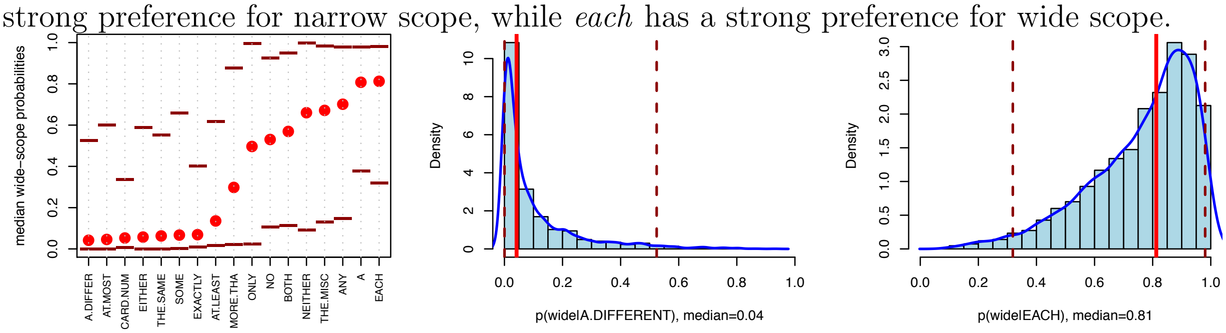
The paper compares several models, but we focus here on one (fitted with the lme4 package in R and in WinBUGS with vague priors): a mixed-effects logistic regression with (***i***) SCOPE (2 levels: *narrow*, *wide*; "success" level: *wide*) as the response variable, (***ii***) two fixed effects LIN.ORD (2 levels: *first*, *last*; reference level: *first*) and GRAM.FUN (2 levels: *S*, *O*; reference level: *S*) and (***iii***) one random effect LEX.REAL (17 levels: *a*, *a.different*, ...). Lexical realizations are modeled as random effects because our sample does not exhaust the population of English quantifiers and, in addition, various lexicalizations are closely related to one another, e.g., *each*, *every* and *all* or modified numerals like *exactly 2* and *at most 2*.

The frequentist estimates for the fixed effects (formula for the *glmer*() function: SCOPE $\sim$ LIN.ORD + GRAM.FUN + 1|LEX.REAL) are: INTERCEPT (i.e., *first* and S) $1.45 (p = 0.0108)$, LIN.ORDLAST $-3.75 (p = 2.23e-08)$ and GRAM.FUNO $-1.04 (p = 0.0502)$. As expected, being first and a Subject increases preference for wide scope, while being last or an Object decreases it. There was no significant interaction between LIN.ORD and GRAM.FUN.

We can more easily understand the estimates and their associated uncertainty in terms of wide-scope probabilities. The 4 figures below display the posterior distributions (obtained in WinBUGS) of the preference for, i.e., probability of, wide scope for the two grammatical functions S and O and the two linear-order positions *first* and *last*—together with the median probability (red) and the 0.025 and 0.975 quantiles (dark red); for presentational convenience, the figures display wide-scope preferences for the lexical item *both*.



The median preference for, i.e., probability of, wide scope for all 17 quantifiers instantiated in the corpus and the 0.025 and 0.975 quantiles are provided in the leftmost figure below (based on the posterior samples obtained in WinBUGS). We also provide the full probability distributions for the two quantifiers at the extremes of the spectrum, *a.different* and *each* (for presentational convenience, the figures below display wide-scope preferences when the lexical items occur first and as Objects): *a.different* (on its sentence-internal reading) has a strong preference for narrow scope, while *each* has a strong preference for wide scope.



The model with LEX.REAL random effects in addition to the LIN.ORD and GRAM.FUN fixed effects has a higher predictive adequacy than the model with fixed effects only (Somers' Dxy is 0.9 with random effects and only 0.68 without), indicating a much more important role for lexical semantics in scope disambiguation than previously assumed (e.g. by [**6**]).

References: [**1**] Anderson, C. (2004). *The Structure and Real-Time Comprehension of Quantifier Scope Ambiguity.* PhD Thesis, Northwestern. [**2**] Higgins, D. & J. Sadock (2003). A Machine Learning Approach to Modeling Scope Preferences. *Comp. Ling.* 29(1). [**3**] Ioup, G. (1975). *The Treatment of Quantifier Scope ...* PhD Thesis, CUNY. [**4**] Kurtzman, H. & M. MacDonald (1993). *Resolution of Quantifier Scope Ambiguities.* Cognition 48. [**5**] Tunstall, S. (1998). *The Interpretation of Quantifiers.* PhD Thesis, UMass. [**6**] Villalta, E. (2003). *The Role of Context in the Resolution of Quantifier Scope ...*, J Semantics 20(2).