Quantifying quantifier scope: a cross-methodological comparison

Oliver Bott and Janina Radó
SFB 441 – University of Tübingen, Germany
oliver.bott@uni-tuebingen.de

In recent years theoreticians have increasingly turned to experimental methods such as questionnaires to get a broader empirical base for their theories. While syntactic judgments can easily be obtained by having informants rate sentences, potential semantic readings are more difficult to access via direct querying: in order to rate paraphrases, for instance, one repeatedly has to compute, remember, and compare two meanings. It is nevertheless very important to collect data from naive speakers for semantic phenomena as well, especially when judgments are unclear or subtle. A case in point is quantifier scope, where the available readings reported in the literature are sometimes quite controversial. For example *Everyone loves someone* has been treated by some as the prototypical example of a scope-ambiguous sentence while for others it only has the wide scope universal ($\forall\exists$) reading. It thus seems desirable to obtain more solid evidence using experimental methods. First, however, we must insure that our methods satisfy at least two criteria: they are simple for the subjects to perform (i.e. no need to juggle two readings), yet sensitive enough to detect fine distinctions.

We report the results of testing three methods that intuitively fulfill these criteria. To guarantee that subjects need only consider one meaning at a time the scope-ambiguous sentences were always presented together with a form of disambiguation. We wanted to find out which method(s) of disambiguation prove to be useful for studying quantifier scope. The methods we tested were: 1) question–answer pairs (QA) in which the question determined the scope of the following target sentence, 2) a picture verification task (PV1) employing abstract set diagrams (e.g. Gillen 1991, Jackson and Lewis 2005) and 3) another picture verification task (PV2) using natural-looking scenarios. This way we contrasted linguistic context (QA) with visual disambiguation in concrete (PV2) vs. abstract (PV1) situations.

The scope-ambiguous sentences themselves were kept maximally similar across the experiments to allow cross-methodological comparison. We chose to manipulate two factors which have been repeatedly claimed to affect quantifier scope: *linear order of quantifiers* and *distributivity*. The latter factor was included since *jede* (and 'each' in English) has been claimed to take wide scope more easily than *alle* (and 'every') (e.g.

Beghelli and Stowell 2002, Pafel 2005, Tunstall 1998). A sample set of sentences is given in (1a-d).

(1) (a) *Genau einen dieser Professoren haben alle Studentinnen*
exactly one of these professors have all fem. students
angehimmelt.
adored.
All students adored exactly one professor.

(b) *Genau einen dieser Professoren hat jede Studentin*
exactly one of these professors has every fem. student
angehimmelt.
adored.
Every student adored exactly one professor.

(c) Alle Studentinnen haben genau einen dieser Professoren angehimmelt.
All students adored exactly one professor.

(d) Jede Studentin hat genau einen dieser Professoren angehimmelt.
Every student adored exactly one professor.

To assess scope preferences we asked subjects to judge how well a sentence fits with the disambiguation using the Magnitude Estimation method which is capable of detecting subtle distinctions in acceptability (Bard et al., 1996).

To make sure that all our methods are sensitive to scope we first pretested them on unambiguous quantified sentences (e.g. *Exactly one professor is such that he is adored by every student*). In the pretests QA, PV1 and PV2 all showed comparably good results.

Next we applied our methods to scope-ambiguous items like those in (1). In QA and PV1 both factors, *order* as well as *distributivity* showed an effect in the expected direction. The results obtained using the two methods were quite similar suggesting that both were able to measure scope preferences. Further, QA and PV1 validly conformed to judgements reported in the literature as well as to a corpus study we carried out using the Cosmas Corpus [1]. However, PV1 and QA differed in terms of inter-rater reliability: judgements were more consistent across participants in PV1 than in QA.

Surprisingly, PV2, which used highly natural scenarios, differed from the other two experiments by only exhibiting an across-the-board preference for the $\forall\exists$ reading. This global shift might be attributed to extralinguistic factors: To verify a sentence relative to a disambiguating model in PV2 subjects may have used a particular visual search strategy which resulted in a predominant $\exists\forall$ reading.

[1] Cosmas Corpus: `www.ids-mannheim.de/cosmas2/`

Our cross-methodological study has major implications for semantic theory and its empirical basis. First, we observed fine-grained distinctions in scope preferences across the different conditions we tested. This suggests that a binary distinction between ambiguous and unambiguous sentences is not sufficient to describe scope phenomena. Second, we were able to identify two experimental methods (QA and PV1) that allowed us to gather subtle semantic acceptability judgments. At the same time our results show that intuitively attractive methods like PV2 may turn out to be unsatisfactory on closer scrutiny. Therefore, cross-methodological comparison is very important to obtain valid data on semantic phenomena because the method might itself affect the results. Finally, our methods can easily be applied to other semantic phenomena as well.

Appendix

To simplify the presentation of the materials, the following examples only show the *–distributive* versions.

Sample materials for Experiment 1: "Question–answer pairs" (QA)

Note: In the experiment, each version of the question (A or A') was paired with each version of the answer (B or B').

A: *Kann man eigentlich von jeder Studentin sagen, dass sie genau einen Professor angehimmelt hat?*
can one actually of every fem. student say that she exactly one professor adored has
Can it be said of every student that she adored exactly one professor?

A': *Kann man eigentlich von genau einem Professor sagen, dass ihn jede Studentin angehimmelt hat?*
can one actually of exactly one professor say that him every fem. student adored has
Can it be said of exactly one professor that every student adored him?

B: *Ja, stimmt. Jede Studentin hat genau einen Professor angehimmelt.*
yes correct every fem. student has exactly one professor adored

B': *Ja, stimmt. Genau einen Professor hat jede Studentin angehimmelt.*
yes correct exactly one professor has every fem. student adored

Sample materials for Experiment 2: “Picture verification 1” (PV1)

Note: In the experiment each word order (SO vs. OS) was paired with each type of diagram.

SO	*Jede*	*Studentin*	*hat*	*genau*	*einen*	*Professor*	*angehimmelt.*
	every	fem. student	has	exactly	one	professor	adored
OS	*Genau*	*einen*	*Professor*	*hat*	*jede*	*Studentin*	*angehimmelt.*
	exactly	one	professor	has	every	fem. student	adored

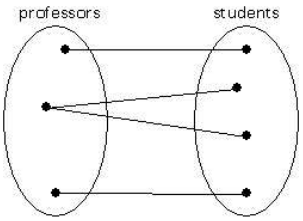

Figure 1: Diagram for ∀∃ reading

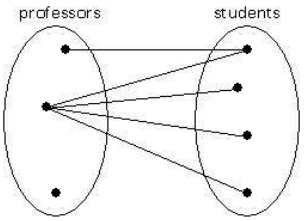

Figure 2: Diagram for ∃∀ reading

Sample materials for Experiment 3: “Picture verification 2” (PV2)

Note: In the experiment each word order (SO vs. OS) was paired with each type of diagram.

SO	*Jedes*	*Kind*	*hat*	*genau*	*ein*	*Dreieck*	*in*	*seiner*	*Spielecke.*
	every	child	has	exactly	one	triangle	in	his	corner
OS	*Genau*	*ein*	*Dreieck*	*hat*	*jedes*	*Kind*	*in*	*seiner*	*Spielecke.*
	exactly	one	corner	has	every	child	in	his	corner

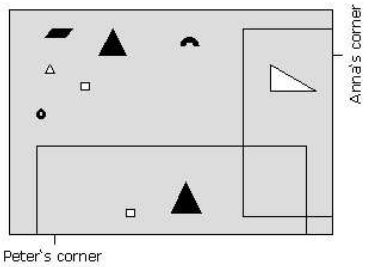

Figure 3: Diagram for ∀∃ reading

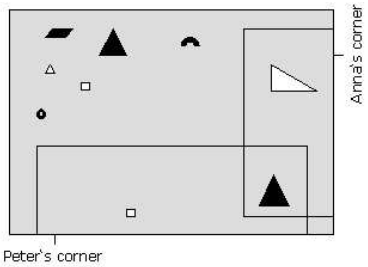

Figure 4: Diagram for ∃∀ reading

References

Bard, E., D. Robertson, and A. Sorace (1996). Magnitude estimation of linguistic acceptability. Language, **72**.

Beghelli, F. and T. Stowell (2002). Distributivity and negation: The syntax of *each* and *every*. In A. Szabolcsi, ed., Ways of Scope Taking, pp. 65–75.

Gillen, K. (1991). The Comprehension of Doubly Quantified Sentences. Ph.D. thesis, University of Durham.

Jackson, S. and W. Lewis (2005). The relation between prosody and logical scope varies by the operator. Poster presented at the CUNY Conference Tucson, Arizona.

Pafel, J. (2005). Quantifier Scope in German. An Investigation into the Relation between Syntax and Semantics. Johns Benjamins, Amsterdam, The Netherlands.

Tunstall, S. (1998). The Interpretation of Quantifiers: Semantics and Processing. Ph.D. thesis, University of Massachusetts, Amherst.