Verb-second Effect on Quantifier Scope Interpretation

Asad Sayeed
Dept. of Philosophy, Linguistics,
and Theory of Science
University of Gothenburg, Sweden
`asad.sayeed@gu.se`

Matthias Lindemann[1], Vera Demberg[1,2]
[1]Dept. of Lang. Sci. and Tech.
[2]Dept. of Math. and CS
Saarland University, Germany
`malinux@t-online.de`
`vera@coli.uni-saarland.de`

Abstract

Sentences like *Every child climbed a tree* have at least two interpretations depending on the precedence order of the universal quantifier and the indefinite. Previous experimental work explores the role that different mechanisms such as semantic reanalysis and world knowledge may have in enabling each interpretation. This paper discusses a web-based task that uses the verb-second characteristic of German main clauses to estimate the influence of word order variation over world knowledge.

1 Introduction

In this paper, we present results from an online cloze (fill-in-the-blank) experiment in German designed to exploit the verb-second syntax of German main clauses in order to understand the relationship between factors involved in quantifier scope ambiguity resolution. An example of quantifier scope ambiguity in English is the sentence "Everybody loves somebody", which has at least two readings: either that (a) every person in the universe of discourse loves some person, but there are possibly different people being loved (linear scope), or (b) that there is exactly one person who is loved by everyone in the universe of discourse[1] (inverse scope).

Humans are both able to perceive scope ambiguities and resolve them in context with relative ease—such that reading (a) is quickly preferred, even though it is a seemingly "out-of-context" example. Potential ambiguities are actually relatively common (Koller et al., 2010), even if example sentences used to understand their formal properties seem rather contrived. Indeed, the

study of the formal properties of scope-ambiguous sentences has a long history in formal semantics (Ruys and Winter, 2011), especially as it pertains to the conditions under which ambiguous readings are available. However, psycholinguistic models of how scope ambiguities are *resolved* (as opposed to merely being made formally available) is a much less developed area of research, but whose development is necessary for, e.g., better computational models of language understanding.

The key issue is the integration of multiple factors: local context, background world-knowledge, "classical" syntax-semantics interface constraints, and so on. For example, in the two readings of *Every child climbed a tree*, common sense generally suggests that the children are climbing multiple trees, which may come from general overall experience of tree-climbing, the affordances (Gibson, 1977) of trees, and so on. Nevertheless, similar sentences have more easily available inverse scope readings, such as *Every jeweler appraised a diamond*. When, if at all, does the human processor "activate" the knowledge that both readings are available (from formal syntactic and semantic structure), and how does this formal knowledge interact with lexical-pragmatic knowledge about trees or diamonds?

2 Background

Early work (e.g., Fodor, 1982) suggested that quantifiers and word order interact, such that when an existential quantifier precedes a universal quantifier, the processor must act to reverse the linear order in order to get a plural reading for the existential quantifier, and *vice versa*. Various psycholinguistic techniques have been brought to bear on the question, especially using the English language as the experimental medium, among them judgement studies, eye-tracking, self-paced read-

---

[1]Taken in a narrowly logical way, reading (b) permits (a) to be true. Pragmatically speaking, however, reading (a) suggests that there are likely multiple/different people being loved. See Radó and Bott (2018) for an argument that the fine-grained logical implication is important from a processing point of view.

(1) Jeder Spion hat diesen/einen/diese Auftrag/Aufträge erhalten. [*Der/die Auftrag/Aufträge war(en)*]
Every spy_NOM has this/a/these order(s)_ACC received. [*The order(s) was/were*]
gefährlich und riskant.
dangerous and risky.

'Every spy received this/a/these order(s). [*The order(s) was/were*] dangerous and risky.'

(2) Diesen/Einen/Diese Auftrag/Aufträge hat jeder Spion erhalten. [*Der/die Auftrag/Aufträge war(en)*]
This/A/These order(s)_ACC has every spy_NOM received. [*The order(s) was/were*]
gefährlich und riskant.
dangerous and risky.

'Every spy received this/a/these order(s). [*The order(s) was/were*] dangerous and risky.'

Figure 1: Stimulus example.

ing, and event-related potentials (ERP). Kurtzman and MacDonald (1993) performed a foundational judgement study in English with a "context-continuation" experimental paradigm, where an ambiguously quantified (universal-existential or existential-universal) sentence (the "context") is followed by another sentence (the "continuation") that contains a singular or plural reference to the existentially quantified noun phrase in the context.

(3) *Context:* Every gardener watered **a plant**.
Continuation: **The plant** was tall and sturdy.

Example (3) illustrates the context-continuation paradigm. In this case, the singular subject of the continuation requires an inverse scope interpretation of the context.

Testing the acceptability of these sentence pairs is more sensitive to online processing than asking for explicit interpretation outright, and Kurtzman and MacDonald used judgement tasks of this nature and found no single principle under which a preference for particular interpretation could be identified. They tested swapping the linear order of quantifiers (e.g., "A gardener watered every plant"), and found no effect of linear order. One problem with such a study is that the quantifiers are swapped while keeping the semantic roles filled by the same nouns, so that it is not possible to separate the effect of linear order from common-sense interpretation.

Filik et al. (2004) performed judgement and eye-tracking studies on English sentences by using double object constructions where the direct and indirect objects had ambiguous universal and existential quantifier scopes. It is possible to swap these in English. This is under a theory where a grammatical hierarchy of constituents (Ioup, 1975) conflicts with linear order. They found that

the linear order effect interacts with the grammatical hierarchy effects in both experimental modalities. However, double object constructions involve either an optional argument or one of a subset of verbs representing events that obligatorily require indirect objects, restricting the ability of this type of study to explore the role of world knowledge[2].

Dwivedi (2013) performed a reading-time study based on stimuli similar to those of Kurtzman and MacDonald, with universal-existential order only for the context sentence and found no significant effects reflecting an online reanalysis process or a competition process in the manner of Kurtzman and MacDonald (1993) or Filik et al. (2004). On the contrary, in a question-answering task after each sentence pair, participants remembered and identified singular continuations at below chance level, being strongly guided by pragmatic intuitions. Dwivedi takes this to reflect a "heuristic first, algorithmic second" theory of scope processing—that the human processor may be *able* to reanalyze ambiguous quantifiers, but does not do so unless very strongly provoked to do so, and instead relies almost entirely on background and contextual knowledge.

Subject-verb-object (SVO) transitive constructions present a greater opportunity to explore the role of semantic events and associated knowledge than double object constructions, but use of the English language presents some limitations. A better exploration of the roles of algorithmic knowledge, linear order, and reanalysis processes in quantifier scope ambiguity processing could be performed in a language that allows for variation of order of the quantifiers without a fundamental change to the semantics of the situation be-

---

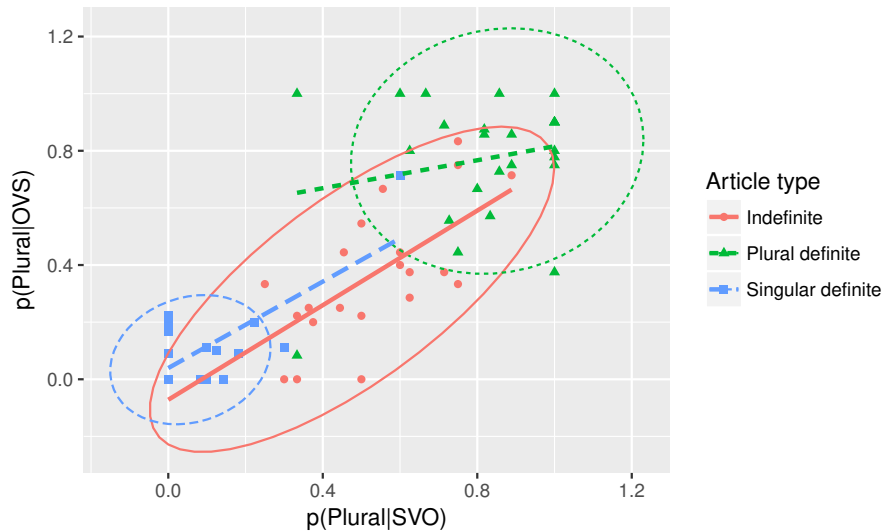[2]See Dwivedi (2013) for further discussion of Filik et al.'s stimuli.

Figure 2: Correlation plot of the probability of plural interpretations given SVO order vs. OVS order, for indefinite, definite singular, and definite plural conditions. Data ellipses illustrate means and standard deviations for each condition in two dimensions (Friendly et al., 2013).

ing represented. Simply swapping object and subject in simple declarative clauses is unacceptable in English, but much less so in verb-second, case-marking languages like German.

Radó and Bott (2018) investigate the role of quantifier distributivity relative to linear order using a series of context-continuation judgement tasks in German. Part of their manipulation is main-clause SVO order vs OSV given German's verb-second characteristic; in German main clauses, quantifier scope ambiguities are possible. However, they use visual aids to explicitly disambiguate the intended interpretation of the sentence, so that the judgement task is about matching the appropriateness of the sentence to the image. Thus, world knowledge effects are factored out of their experiment.

In this work, we instead exploit the SVO/OVS flexibility of German main clauses to better characterise how linear order, formal semantic structure, and world knowledge work together to produce interpretations. We hypothesize that Dwivedi's manipulation yields her result for English because English-speakers rarely, if ever, expect to see linear order variation. The low acceptability of an OVS order in declarative sentences in English means that English users have a much lower expectation of having to reverse the observed linear scope order in order to align their pragmatically-driven interpretation with the observed sentence. Correspondingly, English-

speakers only infrequently invoke "algorithmic" processing mechanisms. German-speakers find OVS order much more acceptable than English-speakers. This means that German-speakers are more likely to confront variation in word order, leading to mismatches between world knowledge and linear order. Since German-speakers expect to confront mismatches more often, syntactic and semantic structure, including word order, will have a more visible effect on number interpretation, providing more visible resistance to the influence of background knowledge. In order to test this, we make use of a web-based cloze task under the context-continuation paradigm, giving participants a free choice of interpretation.

3 Methods and Materials

We translated the English stimuli from Dwivedi (2013) to German, making adjustments or replacements where the syntax or semantics did not work with normal German usage. For the German context sentences, we varied the word order from SVO to OVS order[3] and we had three conditions for the object article: definite singular, indefinite, and definite plural, for six total conditions. The

---

[3]A reviewer points out that our experiment could work, in theory, for any language where the object can optionally precede the subject in a declarative main clause, not merely for verb-second languages. We consider this highly plausible, but we chose not to commit to a more ambitious hypothesis without further analysis of the associated linguistic phenomena beyond Germanic verb-second.

	b	Std. Error	z	$Pr(>|z|)$	
Intercept	-1.1176	0.3732	-2.994	0.00275	**
Linearity(SVO)	0.9260	0.2844	3.256	0.00113	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 1: Fitted mixed-effects model of word order on plural interpretation for objects with the indefinite article.

continuation sentences were supplied with a blank space for the subject and verb, with a final complement describing the targeted subject.

Examples (1) and (2) represent the stimuli we used. The italicized portions are the intended cloze fillers: participants are expected to fill in the subject and the verb, which would thereby reveal their intuitions about number.

The stimuli were randomized into six lists, and 24 fillers were created and added to each list. The stimuli were presented using Lingoturk[4], a workbench for developing and hosting web-based experiments (Pusse et al., 2016), one stimulus per screen, after two training exercises. A mix of participants were recruited, both native-speaker undergraduate students at a university in a German-speaking country and online via Prolific[5], which permits selection by native language.

The collected data (filled blanks) were then coded by a native speaker for whether they represented an unambiguously singular or plural judgement, were ambiguous (e.g., the participant substituted a subject-verb combination that did not refer to the object of the context sentence), or were uninterpretable.

4 Results and Discussion

There were a total of 66 participants, of which 31 were students and 35 were Prolific participants. The student participants were collected first, upon which we discovered that two items had errors. We discarded the data for those items, corrected the errors, and ran the full set with the Prolific participants. This yielded 1546 responses, 1236 of which were unambiguous references to the object of the context sentence with interpretable number[6].

We expected from their semantic interpretations that singular and plural definite articles would produce strong singular and plural interpretations re-

spectively, regardless of any underlying pragmatic bias of the item. We plotted the probability of plural interpretation per item given SVO order vs. OVS order (figure 2). It illustrated that singular and plural interpretations are, as expected, little affected by order when the determiner is definite: they tend to agree with the number of the determiner, and items in the plot cluster near zero probability of plural interpretation for the singular definite condition and high probability of plural interpretation in the plural definite condition. With the indefinite article, however, we obtain nearly the full range of plural biases, with some items having low plural probability with either word order and some high.

We fitted a binomial logit regression model with mixed effects for the indefinite article condition under the hypothesis $Plurality \sim Linearity + (1|Item) + (1 + Linearity|Participant)$; that is, the plural response has a fixed effect of linearity with a random intercept per item (given variation in the overall semantic plural bias of the item) and a correlated random intercept and slope for linearity relative to participant, reflecting overall individual tendencies to give plural interpretations and individual effects of linear order. This was performed in R under the `lme4` package (Bates et al., 2014). The fixed-effects model output is presented in table 1. This model produces a significant effect on plural interpretation in indefinite order sentences, with log-odds ratio $b = 0.926$ in favour of a plural interpretation when the context sentence is in SVO order.

While the plurality of each item is highly preserved in either order, OVS items are indeed less likely to be interpreted as plural, resulting in a slope of the fitted line for the indefinite condition in figure 2 that is visibly less than 1 and reflecting the outcome of the binomial logit regression.

Dwivedi's self-paced reading results in English present a picture of a system that employs reanalysis on ambiguous scopes mostly *in extremis*, but otherwise relies principally on lexical-pragmatic associations in interpretation. Our results do not

---

[4] https://github.com/FlorianPusse/Lingoturk
[5] https://www.prolific.ac
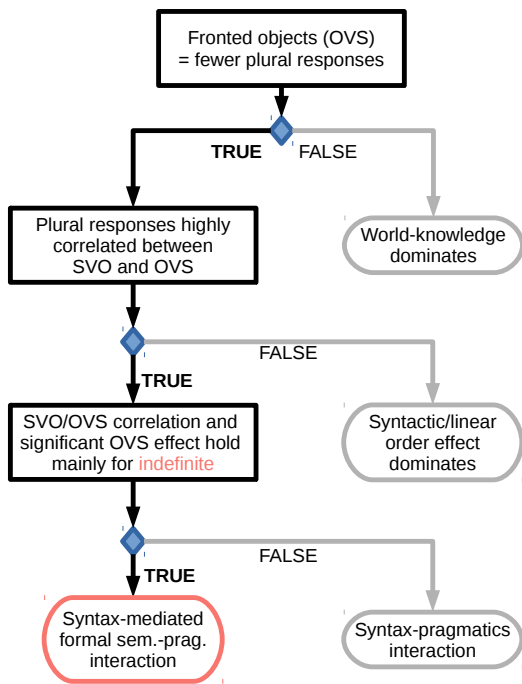[6] The data are available upon request to first author.

Figure 3: Possible outcomes of the experiment relative to hypothesis over dominant mechanisms of ambiguous quantifier scope processing.

challenge the centrality of background world and lexical knowledge, but instead, reflect a somewhat stronger role for the "algorithmic" component of scope processing. Specifically, German speakers' syntactic expectations admit OVS orders more naturally than English speakers, so that when an OVS order is encountered, it creates a "headwind", triggered by an initial indefinite article, that the powerful force of world knowledge must overcome. This mechanism is specifically invoked by the appearance of an indefinite object noun phrase in the initial position, while singular definite articles totally override world knowledge regardless of order.

Dwivedi measured plural bias for her original 2013 stimuli using a forced choice singular/plural interpretation task. We instead opted for a cloze task, which we believe has higher ecological validity—the task avoids forcing participants to choose among options when they may not feel that any of them are suitable or natural. A potential disadvantage of free completions is that participants sometimes filled the subject/verb field in a manner that would not let us determine whether they intended a singular or a plural reference to the object phrase or with a nonsense response; these responses had to be discarded before analysis. In our case, this happened in 20% of the data, still

leaving us with a sufficient amount of interpretable data. The advantage of this approach in the investigation of quantifier scope processing is that it does not explicitly call on participants to think about and judge plurality as a concept, but rather to come up with a phrase that makes sense to them in context.

5 Conclusions and Future Work

Figure 3 provides a visual description of the hypothesis space we considered for this experiment. To the proposition that world knowledge is overly dominant, we found that our result pushes the needle back in the direction of competition-based hypotheses. The specific role of indefinite articles triggering these ambiguities over word order points once again to a complex interaction at the interface between syntax, semantics, and pragmatics. We also demonstrated the experimental value of free(r) word order languages with case marking.

Our cloze task illustrates an aspect of the role of syntactic variation in quantifier interpretation. In order to investigate its specific time course, the most immediate future work is to take our stimuli to other experimental modalities, including self-paced reading and eyetracking.

In the longer run, we plan to use corpus (Sayeed, 2017) and machine learning investigations to determine the empirical basis of plurality judgements in ambiguously scoped events.

Acknowledgements

References

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2014. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Veena D Dwivedi. 2013. Interpreting quantifier scope ambiguity: Evidence of heuristic first, algorithmic second processing. *PloS one*, 8(11):e81461.

Ruth Filik, Kevin B. Paterson, and Simon P. Liversedge. 2004. Processing doubly quantified sentences: Evidence from eye movements. *Psychonomic Bulletin & Review*, 11(5):953–959.

Janet Dean Fodor. 1982. The mental representation of quantifiers. In *Processes, beliefs, and questions*, pages 129–164. Springer.

Michael Friendly, Georges Monette, and John Fox. 2013. Elliptical insights: Understanding statistical methods through elliptical geometry. *Statist. Sci.*, 28(1):1–39.

James J. Gibson. 1977. The theory of affordances. In John Bransford Robert E Shaw, editor, *Perceiving, acting, and knowing: toward an ecological psychology*, pages pp.67–82. Hillsdale, N.J. : Lawrence Erlbaum Associates.

Georgette Ioup. 1975. Some universals for quantifier scope. *Syntax and semantics*, 4:37–58.

Alexander Koller, Stefan Thater, and Manfred Pinkal. 2010. Scope underspecification with tree descriptions: Theory and practice. In Matthew W. Crocker and Jrg Siekmann, editors, *Resource-Adaptive Cognitive Processes*, Cognitive Technologies, pages 337–364. Springer.

Howard S Kurtzman and Maryellen C MacDonald. 1993. Resolution of quantifier scope ambiguities. *Cognition*, 48(3):243–279.

Florian Pusse, Asad Sayeed, and Vera Demberg. 2016. Lingoturk: managing crowdsourced tasks for psycholinguistics. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*, pages 57–61, San Diego, California. Association for Computational Linguistics.

Janina Radó and Oliver Bott. 2018. What do speaker judgments tell us about theories of quantifier scope in German? *Glossa: a journal of general linguistics*, 3(1).

EG Ruys and Yoad Winter. 2011. Quantifier scope in formal linguistics. In *Handbook of philosophical logic*, pages 159–225. Springer.

Asad Sayeed. 2017. Towards an annotation framework for incremental scope specification update. In *Proceedings of the conference on logic and machine learning in natural language (LaML 2017)*, Gothenburg.