Glossa a journal of general linguistics

Radó, Janina 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): 91. 1–29, DOI: https://doi.org/10.5334/gjgl.309

RESEARCH

What do speaker judgments tell us about theories of quantifier scope in German?

Janina Radó[1] and Oliver Bott[2]

[1] Institut für Linguistik, Anglistik, Universität Stuttgart, DE

[2] SPP XPrag.de and SFB 833, Universität Tübingen, DE

Corresponding author: Janina Radó (janina.rado@ling.uni-stuttgart.de)

In this paper we use German data to evaluate configurational and multi-factor approaches to quantifier scope. Configurational theories derive scope relations syntactically at the level of Logical Form; semantic and pragmatic factors are either built into the syntactic representation or ignored, at least during the first derivational step. By contrast, multi-factor approaches consider syntactic, semantic and pragmatic properties of quantifiers as multiple constraints affecting quantifier scope. We examined predictions for quantifier scope in German of the configurational theory by Frey (1993) and of the multi-factor account by Pafel (2005). These fundamentally different approaches were tested in a series of picture verification experiments to assess scope preferences in doubly quantified German sentences. The results show that at least three factors affect the preferred scope. Our findings are neither fully consistent with Frey's configurational theory nor with Pafel's multi-factor approach; both theories made incorrect predictions for German doubly quantified sentences with a subject-before-object word order. For object-before-subject sentences, however, the experimental data by and large support the predictions of Pafel's (2005) multi-factor approach.

Keywords: quantifier scope; multi-factor approaches; configurational approaches

1 Introduction

Intuitive judgments are the main data source semanticists rely on to formulate their theories. Often, however, semantic judgments are not very clear and preferences are gradient rather than clear-cut and categorical. In particular, scope ambiguity of quantifiers is a phenomenon where intuitions tend to be shaky.

The aim of the present paper is to systematically investigate the influence of some factors that have been claimed to influence quantifier scope preferences (e.g. Ioup 1975; Kuno 1991; Beghelli & Stowell 1997; Pafel 2005):

(1) a. The linear order of quantifiers
b. Distributivity
c. Discourse anaphoricity

These three factors have played an important role in various theories of quantifier interaction. In particular linear order (e.g. Reinhart 1983; Kuno 1991; Frey 1993; Pafel 2005) and distributivity of quantifiers (e.g. Kuno 1991; Beghelli & Stowell 1997; Szabolcsi 1997; Pafel 2005) have commonly been claimed to affect scope. The same holds for the discourse status of quantifying expressions. Quantifiers involving a partitive construction such as *each of these* make explicit reference to a contextually given restrictor set. D-linking, which is related to discourse anaphoricity in that it also presupposes a contextually given

restrictor set, has been discussed in the generative literature on wh-operators (e.g. Pesetsky 1987; Cinque 1990), where it has been argued that d-linked phrases can be interpreted in situ whereas non-d-linked phrases have to undergo movement at LF. In a similar vein, Kuno (1991) and Pafel (2005) both predict that quantifiers with a discourse anaphoric restriction should take wide scope more easily than quantifiers without.

The factors have been captured in different ways in different types of theoretical approaches. In this paper we contrast syntactic approaches to quantifier scope with multi-factor theories. The latter take into account syntactic as well as non-syntactic factors without drawing any qualitative distinctions between them. We will completely ignore semantic and pragmatic approaches to the resolution of quantifier scope ambiguity (see among others Hendriks 1993; Barker 2002). Even with this limitation in mind we cannot do justice to all existing scope theories but will focus on a very limited set of proposals dealing with quantifier scope in German (for general reviews see Kiss 2006; Ruys & Winter 2010; Szabolcsi 2010; and Steedman 2011).

1.1 Configurational accounts

Configurational accounts treat scope ambiguities as a kind ofsyntactic ambiguity. Montague (1973) was the first to propose a syntactic rule *quantifying in* which was specifically designed to generate inverse scope readings of multiply quantified sentences. Given different syntactically disambiguated structures, semantic interpretation can proceed straightforwardly in a compositional fashion. Similarly, rather than taking the surface structure as input for semantics, May (1977) argues that syntax continues the derivation until *Logical Form* (LF), the level at which interpretation takes place. Quantifiers move covertly to a position from which they c-command their scope. This mechanism is called *Quantifier Raising* (QR) and has been adopted in several modified versions (May 1985; Hornstein & Weinberg 1990; Reinhart 1995; Heim & Kratzer 1998; Chierchia & McConnell-Ginet 2000; Fox 2003). In unconstrained versions of QR the prediction is that a sentence containing *n* quantifiers should have *n*! readings. Minimalist versions of configurational theories differ in terms of the underlying grammatical architecture but share the assumption that semantic interpretation can target unpronounced constituents of syntactic representations with multiple occurrences of quantifiers in feature-driven derivations (Fox 2003).

Besides QR, another covert syntactic transformation has played an important role in configurational accounts: the reconstruction of quantifiers into their base positions. A proposal along these lines for German is Höhle (1991). For German, it is assumed that scope reconstruction rather than QR is the mechanism underlying inverse scope phenomena (see Sauerland 2003). Using this mechanism scope can be directly read off from surface structure. In this paper, we will focus on the theory proposed in Frey (1993) to account for operator scope in German. Consider the following German examples with verum focus (Höhle 2002), indicated by capitalization of the verb.

(2) a. Fast jeder Lehrer LOBTE mindestens einen Schüler.
'Almost every teacher praised at least one student.'

b. Mindestens einen Schüler LOBTE fast jeder Lehrer.
at least one student.ACC praised almost every teacher.NOM
'Almost every teacher praised at least one student.'

(3) a. [Fast jeder Lehrer$_1$ [lobte$_2$ [t$_1$ mindestens einen Schüler t$_2$]]]

b. [Mindestens einen Schüler$_1$ [lobte$_2$ [fast jeder Lehrer t$_1$ t$_2$]]]

The corresponding syntactic representations involve two movement operations. The verb is moved to the V2 position and one of the quantificational DPs is moved to the prefield:

the subject in (2a), and the object in (2b). There is an intuitive difference between (2a) and (2b): While (2b) is ambiguous, (2a) seems to only allow for a linear scope construal. Frey proposes a scope principle to account for this difference; here is a simplified version, which will do for our purposes (cf. Frey 1993).

(4) Scope Principle (simplified, informal version): A quantifier α can take scope over another quantifier β if α and β are in the same syntactic domain and α syntactically commands β or its trace.

In the syntactic representation of (2a) the subject DP *almost every teacher* c-commands the object quantifier *at least one student*, but not vice versa, thus only one scope reading is possible. This is different in (2b), where the fronted object quantifier takes command over the subject quantifier, but the subject quantifier c-commands the trace t_1 of the object quantifier. This way, the configuration in which the two quantifiers appear poses constraints on the available scope readings.

The theory comes with two important restrictions. First, it only accounts for the interpretation of truly quantificational phrases, i.e. DPs that are not ambiguous between quantificational and other interpretations such as indefinites (Fodor & Sag 1982), which are known to allow for specific, referential interpretations with exceptional wide scope. This is why in the above examples we have German equivalents of *at least one* instead of indefinites such as *a* or *some*. The same goes for group denoting DPs such as *all* or even *every*, which – in one way or another– make reference to plural objects. For this reason the above examples used German equivalents of *almost every* instead of plain *every*. Second, Frey observes that prosody may make interpretations available that are impossible otherwise (see also Krifka 1998). He controls for intonation by always using verum focus to avoid stress on any of the quantificational DPs. If these two restrictions are met, his configurational account predicts scope ambiguity to be absent in German sentences such as (2a) with subject-before-object word order (cf. research question 1 below).

Let us now turn to sentences with object-before-subject word order. As shown above, this syntactic configuration opens up the possibility of inverse scope. This does not mean, however, that all sentences of this type should be fully ambiguous. Other, non-syntactic factors may still influence scope preferences. We can think of them as filters on the set of possible readings that are grammatically licensed in a given configuration. In principle, it may even be the case that one of the syntactically licensed readings is completely ruled out by some other factor or a combination of factors.

Indeed, it has been observed that quantifying expressions are not uniform with respect to their tendency to take wide scope (Szabolcsi 2010), compare e.g. *each* and *every* vs. *all*. Some configurational accounts have directly built influences of prima facie non-syntactic factors into their representations. For instance, Beghelli & Stowell (1997) have developed a minimalist version of QR in terms of feature checking. Quantifiers take scope from different positions in a split-CP structure. In their theory distributive phrases like *each boy* have to check their features at LF at a projection DistP high up in the tree, while non-distributive phrases like *all boys* may remain low. This accounts for the influence of distributivity on a quantifier's tendency to take wide scope. What at first sight seem to be lexical semantic distinctions between different determiners are thus claimed to be configurational differences between logical forms. In principle this strategy can be applied to many factors that affect the scopal behavior of quantifying expressions.

Beghelli & Stowell's theory aims to account for quantifier scope in English and was not intended to apply to scope in German. As there is no worked-out proposal for German of this sort, we will focus our discussion on the configurational account of Frey (1993). Obviously, the two proposals are very different and may in fact be incompatible with each

other. We do not suggest to combine them, nor do we want to test Beghelli & Stowell's theory in a language it was not designed for. It may seem tempting, though, to also try to account for other, prima facie non-syntactic scope factors in configurational terms. However, incorporating different scope factors directly into the hierarchical representations has important consequences: Scope factors should interact in an asymmetrical fashion. Consider the following LF with dedicated positions SPECXP and SPECYP for quantifiers with the two scope-relevant properties *X* and *Y*.

(5) $[_{XP}$ SPECXP $[_{YP}$ SPECYP …]]

If the quantifier feature *X* has to be checked in a high position at LF it should not matter whether the other quantifier needs feature checking at some lower hierarchical level (here, feature *Y*). Manipulating factor *Y* should only affect quantifier scope as long as the other quantifier does not have to be checked with respect to feature *X*. To put it differently, a consequence of this type of configurational account is that independent manipulations of two scope factors are predicted to take the form of (under-additive) statistical interactions. We will come back to this point when discussing the purely additive effects observed for manipulations of the different scope factors in Experiments 2 and 3 in section 4 (cf. research question 3 below).

An anonymous reviewer pointed out that the just given characterization of configurational accounts does not do justice to a great number of configurational theories that have been proposed. Often syntactic interpretation is thought of as generating all combinatorially possible scope readings within a given syntactic domain. Some of these readings may subsequently be filtered out by other factors such as derivational economy in combination with logical entailments, plausibility, information structure, and so forth. On closer scrutiny configurational accounts should therefore be also thought of as multi-factor theories. The experimental results reported below are admittedly fully consistent with such a view. However, the point just made may apply to such conceptions of relative scope as well depending on how the filtering mechanism is spelled out. If, for instance, lexical properties such as the distributivity of a quantifier filter out certain readings before discourse properties are considered, we would still expect to find asymmetrical dependencies between scope factors. Thus, an important question is whether constraining factors are applied in parallel or in a strictly serial fashion. The results reported below indicate that the factors tested in the present paper work in parallel because they symmetrically contribute to the available readings. Our data therefore still have important implications for the architecture of the grammar.

1.2 Multi-factor approaches

Multi-factor approachess assume a rather indirect relationship between syntax and semantics. Syntactic constellation is just one of a number of factors which determine the relative scope of quantifiers. Unlike configurational accounts, multi-factor accounts aim to model the precise distribution of readings observed for multiply quantified sentences. For that reason, they first and foremost strive for descriptive adequacy. Multi-factor approaches comprise very early theories like those of Kroch (1974); Ioup (1975) and VanLehn (1978) but also the somewhat more recent proposals by Kuno (1991) and Pafel (2005). According to the latter two accounts a genuinely syntactic factor (LINEAR ORDER) plus a series of other factors such as DISTRIBUTIVITY and DISCOURSE ANAPHORICITY of quantifiers influence relative scope. In Pafel's account the relative scope of German multiply quantified sentences is predicted via a linear additive model which uses weighted factors to determine the scope value of a given quantifier and rank the quantifiers with respect to their

values. Multi-factor theories claim that the scope determining factors are independent from each other and that they add up in a purely cumulative fashion.

To illustrate how scope preferences are derived under a multi-factor approach, we come back to the examples (2a) and (2b) from above. The following predictions are derived from Pafel's (2005) account of relative scope in German with one slight modification. Pafel uses threshold values, i.e. difference scores between quantifiers' "scope potentials", to distinguish between ambiguous and unambiguous sentences. We do not employ thresholds here because existing experimental work suggests that their use leads to a loss of explained variance relative to a model without thresholds (Bott & Radó 2007). The only relevant factors in Pafel's model with respect to (2a) and (2b) are LINEAR ORDER, GRAMMATICAL FUNCTION, and DISTRIBUTIVITY. Other factors such as DISCOURSE ANAPHORICITY ("discourse binding"), THEMATIC PROPERTIES of arguments related to psych verbs, INTONATION and so forth do not play a role here. Each quantifier is evaluated with respect to these dimensions by computing its respective scope value (s-val) according to a linear additive model using factor weights. If a quantifier in the prefield precedes another one (1ST, for short) a value of 1.5 is added, if it functions as subject (SUBJ) a value of 1 is added, and if it is distributive (DIST) its scope value is increased by 1. Thus, the following scope values are derived:

(6) a. s-val (*almost every teacher*) = 1ST & SUBJ & DIST = 1.5 + 1 + 1 = 3.5 (ex. (2a))
s-val (*at least one student*) = 0 (ex. (2a))

b. s-val (*at least one student*) = 1ST = 1.5 (ex. (2b))
s-val (*almost every teacher*) = SUBJ & DIST = 1 + 1 = 2 (ex. (2b))

Comparing the respective scope values shows that the predicted readings are largely the same as those predicted by Frey (1993).[1] For (2a), the scope value of *almost every teacher* is much higher as the one for *at least one student*, which corresponds to a very strong prefence for linear scope, whereas the much closer values for (2) indicate a more or less ambiguous sentence with possibly a slight preference for inverse scope. However, if we exchange the quantifiers, the predictions of the two theories differ considerably (DISCOURSE ANAPHORICITY (D-ANA) has a factor weight of 1.)

(7) a. Genau ein Lehrer LOBTE fast jeden dieser Schüler.
exactly one teacher.NOM praised almost every of these students.ACC

b. s-val (*exactly one teacher*) = 1ST & SUBJ = 1.5 + 1 = 2.5
s-val (*almost every of these students*) = DIST & D-ANA = 1 + 1 = 2

Thus in Pafel's multi-factor account a small change of quantifiers should result in a rather large change in scope distributions. (7) with subject-before-object word order is predicted to be scope ambiguous with a slight preference for a linear scope construal. For doubly quantified sentences with an object-before-subject word order the prediction is again that factors other than the syntactic configuation should have a strong influence on scope preferences. Consider (8).

(8) a. Fast jeden dieser Schüler LOBTE genau ein Lehrer.
almost every of these students.ACC praised exactly one teacher.NOM
'Exactly one teacher praised almost every one of these students.'

[1] And even more so in Pafel's original model with thresholds where (2a) is predicted to unambiguously exhibit surface scope.

b. s-val (*almost every one of these students*) = 3.5
s-val (*exactly one teacher*) = 1

c. Genau einen Schüler LOBTE fast jeder dieser Lehrer.
Exactly one student.ACC praised almost every of these teachers.NOM
'Almost every one of these teachers praised exactly one student.'

d. s-val (*exactly one student*) = 1.5
s-val (*almost every one of these teachers*) = 3

While (8a) is predicted to show a very strong preference for linear scope, in (8c) inverse scope should actually be preferred over the linear reading (cf. research question 2 below).

The discussion of these examples illustrates that the aim of multi-factor accounts is rather different from configurational theories: They try to account for subtle differences in scope preferences at a rather descriptive level. Unlike configurational accounts multi-factor theories are far less concerned with identifying a single mechanism that explains why particular readings are available or unavailable. Most notably, their predictions lend themselves to psycholinguistic studies because they can be directly translated into quantitative predictions for experiments.

Based on the predictions of the theories outlined above we conducted a study of quantifier scope in German. In particular, we addressed the following questions:

i. Do doubly quantified sentences with subject-before-object word order have unambiguous linear scope as predicted by Frey (1993), or are they potentially ambiguous as predicted by Pafel (2005), cf. (2a) vs. (7)?
ii. Do we find evidence for the scope distributions predicted by Pafel's multi-factor account for object-before-subject sentences of the types illustrated in (8)?
iii. If the scope factors claimed relevant for quantifier scope do in fact show the predicted influences, how do they interact? Do we find evidence for purely additive effects as essentially predicted by additive linear models of the sort just outlined? Or, do we rather find under-additive effects as may be expected under a configurational account that encodes non-syntactic factors at dedicated positions in the LF (Beghelli & Stowell 1997)?

The first two questions will be addressed in our first experiment (section 3), which is an explicit comparison of Frey's and Pafel's theory. The third question concerns configurational and multi-factor theories in general and will be taken up in Experiments 2 and 3 (section 4). Before going into the experimental study, we will first relate our research questions to existing psycholinguistic literature on quantifier scope ambiguity, then in section 2 we will discuss methodological considerations important for assessing quantifier scope preferences.

1.3 Psycholinguistic studies on quantifier scope

Multi-factor models have been quite prominent in the psycholinguistic literature on the processing of quantifier scope. We cannot go into detail here but refer the reader to the overview in Tunstall (1998: Chapter 2), and the more recent studies by Paterson et al. (2008) and Brasovaeanu & Dotlacil (2015). It has been repeatedly claimed that the language processor simultaneously relies on different sources of information to generate the possible readings of multiply quantified sentences (e.g. Kurtzman & MacDonald 1993; Filik et al. 2004; Paterson et al. 2008). Unfortunately, the existing psycholinguistic studies cannot easily be related to the scope theories reviewed above.

One reason is that many of the existing studies used suboptimal disambiguations. In Kurtzman & MacDonald (1993); Tunstall (1998) and Dwivedi (2013) as well as in the Filik et al. (2004) and Paterson et al. (2008) studies a potentially scope ambiguous sentence with a universal and an existential quantifier like *every kid climbed some tree* was combined either with a singular or with a plural continuation (e.g. *this tree/these trees was/were full of apples*) which was intended to disambiguate the scope. However, it has often been noted both in the semantic (e.g. Kempson & Cormack 1981) and in the processing literature (e.g. Tunstall 1998) that the ∃∀-reading entails the ∀∃-reading. Therefore, the singular continuation is compatible with both interpretations and does not achieve real disambiguation. This can be illustrated by combining an unambiguously ∀∃-sentence with a singular continuation in (9).

(9) Every child is such that it climbed some tree. This tree was full of apples.

If the singular continuation were only compatible with a wide scope existential reading then the discourse in (9) should be incoherent, as this reading is excluded in the first sentence. The discourse is coherent, though, showing that the singular continuation does not fulfill its function of providing disambiguation. This could result in an overestimation of the ∃∀ readings. More generally, the method used to assess whether a particular reading is available must not introduce a distorting bias. In the experiments presented here, we will therefore use a picture verification task that has been shown to yield reliable and valid results (Bott & Radó 2007).

Second, to evaluate the theoretical approaches outlined above we need to systematically investigate the interplay of the various factors that have received attention in the semantic literature on quantifier scope. As outlined above, a crucial test of multi-factor theories is examining their prediction of purely additive scope influences. However, none of the existing psycholinguistic studies has manipulated multiple factors in tandem using the same participants, items and experimental procedure.

At the same time, there is a body of experimental work examining these factors separately. The potential influence of LINEAR ORDER has been investigated but the findings are mixed. While VanLehn (1978); Fodor (1982); Gillen (1991); Kurtzman & MacDonald (1993); Tunstall (1998); Anderson (2004) and Paterson et al. (2008) reported effects of linear order and/or c-command relations, Catlin & Micham (1975); Ioup (1975); and Micham et al. (1980) found no effects of surface order but only of grammatical function. The effect of DISTRIBUTIVITY has been tested experimentally by Tunstall (1998); Bott & Radó (2007; 2008); Radó & Bott (2012); and Brasovaeanu & Dotlacil (2015). Although the findings are not unanimous,[2] distributive quantifiers do in fact seem to have a stronger tendency to take wide scope than non-distributive quantifiers. To our knowledge, Syrett et al. (2016) is the only study that has investigated the influence of DISCOURSE ANAPHORICITY on quantifier interpretation in adults. In one of the four experiments reported in the paper, they manipulated whether or not quantifiers included a partitive construction. Their findings suggest that discourse anaphoric quantifiers take inverse scope over another quantifier more easily than their non-discourse anaphoric counterparts. In addition, experimental work on language acquisition by Miller & Schmitt (2004) and Musolino & Gualmini (2004) suggests that discourse anaphoric partitive quantifiers can be more easily interpreted with inverse scope than non-partitive quantifier phrases. To summarize, the three factors manipulated in the present study can be expected to influence quantifier

[2] See Filik et al. (2004) and Paterson et al. (2008), who did not find stronger effects for *each* than for *every*. In their studies, however, DISTRIBUTIVITY was only manipulated across experiments.

scope and are thus good candidates for studying research question 3, viz. the nature of the interaction between scope factors.

Finally, almost all of these studies investigated quantifier scope in English. Because English has strict SVO word order, in an active declarative sentence the subject always precedes the object. Thus the effect of linear order/c-command relations and of the grammatical function of the quantifiers cannot be readily distinguished. In other languages word order is much less constrained and these factors can be teased apart more easily. We therefore chose German, where the direct object can precede the subject.[3] This is illustrated in (10a) and (10b).

(10) a. Jeder Lehrer lobte genau einen Schüler.
every teacher.NOM praised exactly one student.ACC
'Every teacher praised exactly one student.'

b. Genau einen Schüler lobte jeder Lehrer.
exactly one student.ACC praised every teacher.NOM
'Every teacher praised exactly one student.'

However, there is one line of study that is closely related to the questions addressed in the present paper. Bott & Schlotterbeck (2012) examined whether German object-before-subject and subject-before-object sentences both allow inverse scope readings. They were particularly interested in the readings that become available during the online interpretation of doubly quantified sentences. Their study provided evidence that subject-before-object sentences differ from object-before-subject sentences in that only the latter gave rise to online effects of scope ambiguity. The final interpretations as measured in an offline task, however, indicated ambiguity in both types of constructions. Furthermore, Bott & Schlotterbeck (2012) only manipulated the configuration that the quantifiers appeared in, they did not investigate the interactions with other, non-syntactic scope factors.

To sum up, we think that the existing studies are not sufficient to evaluate the scope accounts discussed above. Nonetheless, they provide initial support that scope is influenced by different factors. In the following we will present three picture verification experiments that were designed to test the two scope theories we focus on: Frey's (1993) configurational account and the multi-factor approach by Pafel (2005).

2 Methodological considerations

To determine scope preferences, we have to relate a potentially scope ambiguous construction to each of its potential meanings and measure how well they fit. To this end, we used a picture verification task and had participants judge how well doubly quantified sentences match their disambiguations. All sentences contained a universal quantifier (*jeder* ('every/each')[4] or *alle* ('all'); ∀) and *genau ein* ('exactly one') ('$\exists^{=1}$'). The disambiguations were set diagrams only consistent with one scope reading, one model disambiguating towards $\exists^{=1}\forall$, the other disambiguating towards $\forall\exists^{=1}$. Figure 1 shows two disambiguating models for sentence (11). Set diagrams of the sort used here have been successfully employed in a number of experimental studies on scope (see Bott & Radó 2007 for a cross-methodological study investigating the reliability and validity of the

[3] Although object-subject word order has been shown to be more complex to process than subject-object order (see e.g. Hemforth & Konieczny 1999 and the papers therein.) Note that we are not making any claims about processing but are only interested in the final interpretation.

[4] In the following, we will translate *jed-* sometimes by *every* and sometimes by *each* following our intuitions for English.

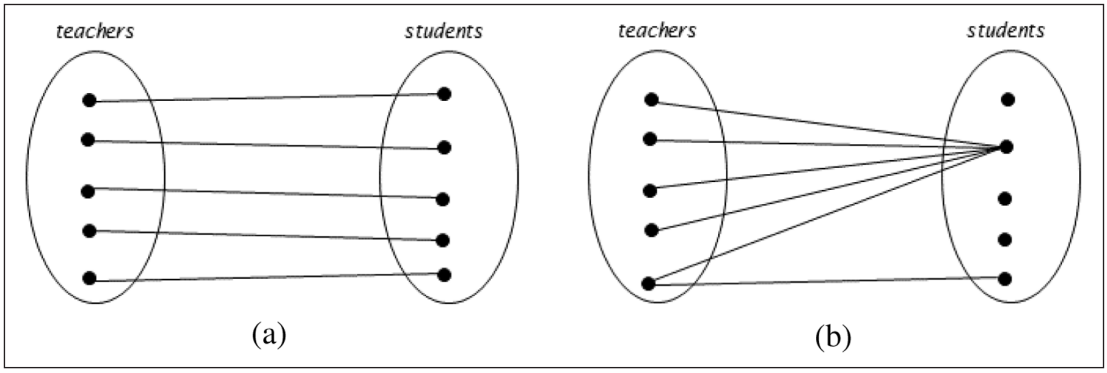

Figure 1: Disambiguations used in the picture verification experiments: **(a)** Diagram for $\forall\exists^{=1}$-reading; **(b)** Diagram for $\exists^{=1}\forall$-reading.

method, and Gillen 1991; Jackson & Lewis 2005; Bott & Schlotterbeck 2012; Robaldo et al. 2014; and Bott & Schlotterbeck 2015 for applications).

(11) Exactly one student was praised by every teacher.

The model on the right is true on the $\exists^{=1}\forall$ reading since there is exactly one student who was praised by all the teachers. It is false on the $\forall\exists^{=1}$ reading because one student was praised by more than one teacher. By contrast, the model on the left is true on the $\forall\exists^{=1}$ reading because for each teacher there is exactly one student whom he praised. It is false on the $\exists^{=1}\forall$ reading because there is no student who was praised by all the teachers. The $\exists^{=1}\forall$ diagrams were always of the kind in Figure 1(b), that is, all of them contained at least one additional line, which made the $\forall\exists^{=1}$ reading false.

We are now able to operationalize what should count as scope preference (**criterion to determine scope preference**): Reading A is preferred over reading B if the A-disambiguation is rated better than the B-disambiguation; if the ratings are equal, the sentence is fully ambiguous.[5]

In our experiments we used non-monotone *exactly one* in order to provide proper disambiguations for the sentences (see e.g. the methodological discussion in Ruys & Winter 2010). The choice of quantifier was also crucial for testing Frey's (1993) theory, which only applies to a small number of quantificational expressions, namely those that are not ambiguous between specific/referential and quantificational expressions (i.e. "indefinites") or between collective and quantificational expressions (i.e. "plurals"). The quantifiers *genau ein* ('exactly one') and *jeder* ('every') fulfill these criteria, they can thus be used to test Frey's configurational account. Let's first consider the possibility of referential interpretations of *exactly one*.

(12) a. Peter glaubt, dass ihn (genau) eine Vegetarierin liebt.
Peter believes that him_{OBJ} (exactly) one/a vegetarian.FEM.NOM loves
'Peter believes that he is loved by exactly one/a vegetarian.'

b. Falls (genau) einer meiner Onkels dieses Jahr stirbt, erbe ich ein
if (exactly) one/a of my uncles this year dies, inherit I a
Vermögen.
fortune
'If (exactly one)/some of my uncles dies this year, I'll inherit a fortune.'

[5] Unfortunately, the term *ambiguous* is itself ambiguous. In the following we will distinguish between merely *ambiguous*, i.e. more than one interpretation is available regardless of whether one is preferred over the other, and *fully* ambiguous, i.e. the available interpretations are roughly equally preferred.

Example (12a) employs a test used by Frey (1993). The sentence without *exactly*, i.e. indefinite *eine* ('a') allows for a de re reading where Peter believes that he is loved by some woman without knowing that she is a vegetarian. This reading is not possible with *exactly one*, mirroring Frey's observation for the monotone quantifier *mindestens ein-* ('at least one'). In other words, *genau ein* cannot take exceptional wide scope over *believe*. (12b) (cf. Fodor & Sag 1982) shows that *exactly one* cannot outscope the if-clause. The sentence only allows for the pragmatically marked interpretation that there is a strange inheritance rule according to which I inherit a fortune only if one but not more than one of my uncles dies (no matter which one). The specific interpretation, easily available without *genau*, becomes impossible once we add *exactly*.[6]

The same goes for the unavailability of collective interpretations of *jeder* ('every'). *Jeder* cannot co-occur with collective predicates such as *surround*, *form a line*, or *run apart* resulting in odd interpretations in (13a)–(13c). Moreover, as mentioned above Frey dismisses unmodified *jeder* and only uses *fast jeder* ('almost every'). The reason is that in his view unmodified *jeder* makes reference to a given restrictor set, as evidenced by strong readings of unmodified examples (Dalrymple et al. 1998; Schlotterbeck & Bott 2013). Pafel (2005), however, observes that basically the same readings are available for *jeder* and *fast jeder* in (13d). The most natural interpretation for the sentence is that for (almost) each pair of individuals from the domain the two individuals differ in their smell. This interpretation is available with or without *almost*, contrary to Frey's proposal. In order to avoid problems due to the inherent vagueness of *almost* we therefore used unmodified *jeder* in the experiments reported below.

(13) a. Alle Ritter haben/#Jeder Ritter hat die Burg umzingelt.
all knights have/#every knight has the castle surrounded
'All knights have/#Every knight has surrounded the castle.'

b. Alle Tänzer haben/#Jeder Tänzer hat eine Linie gebildet.
all dancers have/#every dancer has a line formed
'All dancers have/#Every dancer has formed a line.'

c. Alle Kinder sind/#Jedes Kind ist auseinander gelaufen.
all kids are/#every kid is apart run
'All kids have/#has run apart.'

d. Fast jeder Mensch/Jeder Mensch hat einen anderen Geruch.
almost every human/every human has a different smell
'Almost every human/Every human has a different smell.'

Thus the quantifiers *jeder* and *genau ein*, which we used to test Frey's theory in Experiment 1 (section 3), satisfy Frey's own requirements for being unambiguously quantificational. In Experiments 2–3 (section 4) we will also compare *every (one of these)* with *all (of these)*. The sole purpose of this comparison between *jeder* ('every') and *alle* ('all') is to examine distributivity effects; the results for conditions with *alle* will not be used to evaluate Frey's theory but will only become relevant when testing the additivity of scope factors.

[6] An anonymous reviewer suggested that a referential interpretation may be possible for *exactly one* in the following example with *will* ('want') instead of *believe*:

(i) Peter will genau eine Vegetarierin heiraten.
Peter wants exactly one vegetarian.FEM marry
'Peter wants to marry exactly one vegetarian.'

Our intuitions are the same as for *believe* in (12a), but the availability of a specific interpretation for *exactly one* should be tested experimentally. We have to leave this to future research.

Finally, in order to address the first research question stated above, it is crucial to define a criterion that allows us to decide whether doubly quantified sentences with consistently low ratings for the inverse interpretation are still in principle ambiguous or only have the linear reading. We will therefore compare potentially scope ambiguous sentences such as (7), (8a), and (8c) with scope disambiguated baseline control conditions in which quantifiers are separated by a clause boundary (see also Bott & Schlotterbeck 2012 for the same rationale). This is illustrated in the following English translations of the three experimental conditions to be tested in Experiment 1.

(14) a. Exactly one teacher is such that he praised each of these students.
b. Each of these students is such that he was praised by exactly one teacher.
c. Exactly one student is such that he was praised by each of these teachers.

The comparison with unambiguous constructions allows us to operationalize what we mean by saying that a sentence is, or rather is not compatible with some scope reading (**criterion to determine scope ambiguity**): Sentence S allows for scope reading A, iff the picture disambiguating for A is judged more acceptable when paired with S than when paired with the appropriate disambiguated baseline control S'. If there is no difference between S and S' with respect to their distributions of acceptability scores for the potential readings, S is unambiguous.

3 Frey (1993) vs. Pafel (2005) (Exp. 1)

In section 1 we introduced Frey's (1993) and Pafel's (2005) account using three doubly quantified constructions for which the two theories make rather different predictions. These three constructions were tested in the first experiment.

3.1 Methods

3.1.1 Materials

We constructed 36 doubly quantified sentences in three sentence conditions each (see Appendix A for the complete list of items). Here is a sample item.

(15) a. Genau einen Schüler lobte jeder Lehrer voller Wohlwollen.
exactly one pupil.ACC praised each teacher.NOM full-of goodwill
'Each teacher praised exactly one pupil full of goodwill.'
b. Jeden dieser Schüler lobte genau ein Lehrer…
each these pupils.ACC praised exactly one teacher.NOM…
'Exactly one teacher praised each of these pupils…'
c. Genau ein Lehrer lobte jeden dieser Schüler…
exactly one teacher.NOM praised each these pupils.ACC…
'Exactly one teacher praised each of these pupils…'

Conditions (15a) and (15b) are object-before-subject sentences. They differ, however, with respect to the relative factor weights of the quantifiers. While in example (15b) based on Pafel's (2005) account we expect full ambiguity with a slight tendency of the distributive subject quantifier *each teacher* to take scope over the topicalized object *exactly one student*, in (15b) the cumulative factor weights of linear order, distributivity and discourse anaphoricity of *jeden dieser Schüler* work against the subject quantifier *exactly one teacher* and we predict the sentence to be strongly biased towards linear scope. According to Pafel's (2005) model changing the order of the quantifiers in (15c) should yield an ambiguous sentence, this time slightly biased towards the linear interpretation. This is because the sentence-initial quantifier is the subject whereas the second quantifier is

distributive and discourse anaphoric. The magnitude of the preference should be exactly the same as in the first object-before-subject condition (15a), but go in the opposite direction. The factor weights and predicted scope values based on Pafel (2005) for each of the three constructions in (15) are summarized in (16).

(16) a. Genau einen Schüler$_{\exists^{=1}}$ lobte jeder Lehrer$_{\forall}$ voller Wohlwollen.
$\exists^{=1}$: 1.5 (1ST) = 1.5
$\forall$: 1 (SUBJ) + 1 (DIST) = 2
⊳ slight preference for inverse scope

b. Jeden dieser Schüler$_{\forall}$ lobte genau ein Lehrer$_{\exists^{=1}}$ voller Wohlwollen.
$\exists^{=1}$: 1 (subject) = 1
$\forall$: 1.5 (1ST) + 1 (DIST) + 1 (D-ANA) = 3.5
⊳ strong preference for linear scope

c. Genau ein Lehrer$_{\exists^{=1}}$ lobte jeden dieser Schüler$_{\forall}$ voller Wohlwollen.
$\exists^{=1}$: 1.5 (1ST) + 1 (SUBJ) = 2.5
$\forall$: 1 (DIST) + 1 (D-ANA) = 2
⊳ slight preference for linear scope

However, if scope ambiguity is only possible in object-before-subject sentences, as Frey predicts, then the subject-before-object sentence (15c) should have unambiguously linear scope.[7] To test this prediction we added three scope unambiguous controls with quantifiers appearing in clause bounded position. If the potentially ambiguous subject-before-object construction (15c) should turn out to be indistinguishable from a corresponding scope unambiguous control condition we can legitimately take it to be unambiguous. Consider the sample item in the three control conditions in (17).

(17) a. Für genau einen Schüler gilt: Ihn lobte jeder Lehrer.
for exactly one pupil holds: Him.ACC praised each teacher.NOM
'Exactly one pupil is such that he was praised by each teacher.'

b. Für jeden dieser Schüler gilt: Ihn lobte genau ein Lehrer.
for each these pupils holds: Him.ACC praised exactly one teacher.NOM
'Each of these pupils is such that he was praised by exactly one teacher.'

c. Für genau einen Lehrer gilt: Er lobte jeden dieser Schüler.
for exactly one teacher holds: He.NOM praised each these pupils.ACC
'Exactly one teacher is such that he praised each of these pupils.'

Each of the six sentence conditions was paired with a picture that was only compatible with the $\forall\exists^{=1}$ or the $\exists^{=1}\forall$ reading. This yielded a total of twelve sentence-picture pair

[7] An anonymous reviewer raised the possibility that the SVO sentence in (15c) does not represent the basic structure, rather, the object has scrambled over the adverb *voller Wohlwollen*, which marks the left edge of the VP. In this structure the moved object would c-command the trace of the subject inside the VP, thus in principle allowing for inverse scope. We believe, however, that this is not the case. If (15c) involves scrambling of the object then the basic structure before movement should be like (i):

(i) Ich glaube, dass voller Wohlwollen genau ein Lehrer jeden Schüler gelobt hat.
I think that full-of goodwill exactly one teacher.NOM each pupil.ACC praised has
'I think that exactly one teacher praised each pupil full of goodwill.'

However, (i) does not seem acceptable, at least not without a special intonation suggesting that the adverb has been fronted. It is much more likely that the underlying structure is as in (ii), i.e. with the adverb inside the VP.

(ii) Ich glaube, dass genau ein Lehrer jeden Schüler voller Wohlwollen gelobt hat.

Thus the direct object is presumably in its basic position in (15c) and only linear scope should be possible according to Frey's theory.

conditions according to a 3 (CONSTRUCTION: OS1 vs. OS2 vs. SO) × 2 (AMBIGUITY: AMB. vs. DISAMB.) × 2 (DISAMBIGUATION: LINEAR vs. INVERSE) factorial design. Sample pictures for all three construction types are given in Figure 2. The $\forall\exists^{=1}$-picture was always the same in the OS1, the OS2 and the SO sentences. The $\exists^{=1}\forall$ pictures in the OS1 order were just the mirror image of those used in the OS2 and SO orders. We added 61 filler sentence-picture pairs (15 false) and constructed 12 lists using a latin square design.

3.1.2 Procedure and participants

The experiment was programmed using WebExp 2 (Mayo et al. 2006). After reading written instructions, participants first completed a short practice session consisting of ten trials. That was followed by the experiment in one block. Sentence-picture pairs were presented in an individually randomized order. Judgments were provided on a seven-point scale. We measured both judgments and judgment times.

48 students at Tübingen University (mean age 24.7 years, range 20–33 years; 32 female) participated in the study for a payment of 5€. Participants were tested individually in a quiet computer pool at the Department of Modern Languages. An experimental session took about 30 minutes.

3.2 Results and discussion

The ratings were normalized by computing z-scores for each participant. Mean judgments are shown in Figure 3. We will report separate 2 (AMBIGUITY) × 2 (DISAMBIGUATION) repeated measures ANOVAs for the three constructions.

The SO conditions provide evidence for the possibility of inverse scope in this construction type. Even though the linear interpretation was strongly preferred over the inverse reading (mean z-scores: 0.41 vs. –0.89) this preference was significantly weaker than the corresponding difference in the unambiguous conditions (mean z-scores: 0.63 vs. –1.16). ANOVAs revealed a main effect (marginal by items) of AMBIGUITY ($F_1(1,47) = 4.14$, $p < .05$; $F_2(1,35) = 3.90$, $p = .06$), a significant main effect of DISAMBIGUATION ($F_1(1,47) = 474.45$, $p < .01$; $F_2(1,35) = 956.23$, $p < .01$) and a significant interaction between

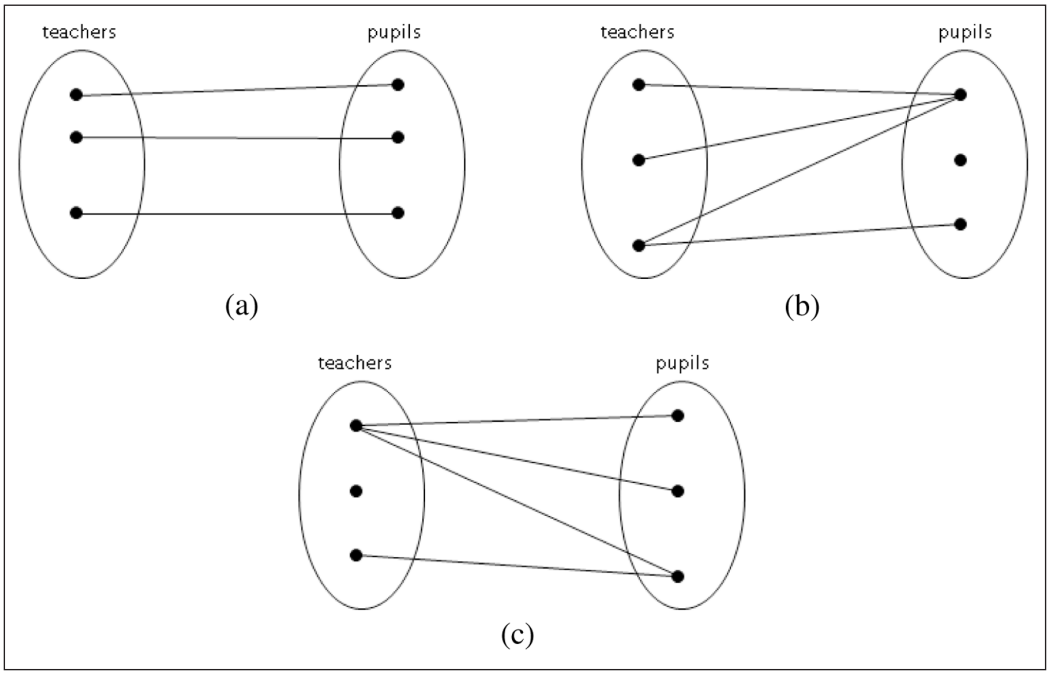

Figure 2: Diagrams used in Exp. 1: **(a)** Diagram for $\forall\exists^{=1}$-reading; **(b)** Diagram for $\exists^{=1}\forall$-reading in the OS1 conditions; **(c)** Diagram for $\exists^{=1}\forall$-reading in the OS2 and SO conditions.

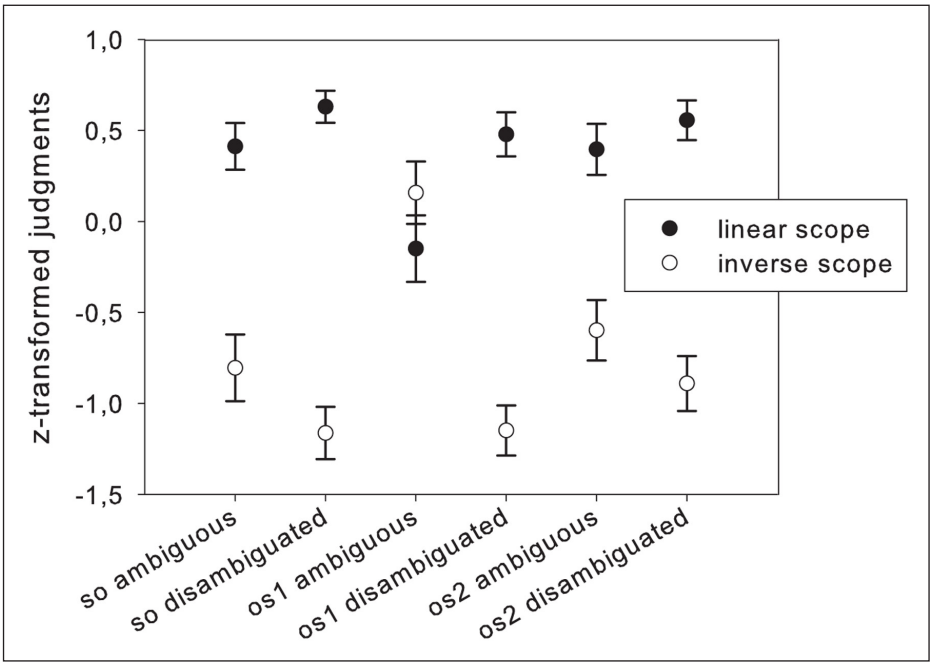

Figure 3: Judgments in Experiment 1 (plus 95% confidence intervals). so = *exactly one*$_{SUBJ}$ > *each of these*$_{OBJ}$; os1 = *exactly one*$_{OBJ}$ > *each*$_{SUBJ}$; os2 = *each of these*$_{OBJ}$ > *exactly one*$_{SUBJ}$.

AMBIGUITY and DISAMBIGUATION ($F_1(1,47) = 44.03, p < .01$; $F_2(1,35) = 33.16, p < .01$). This interaction was due to ambiguity in the ambiguous conditions as compared to the unambiguous controls. Finding ambiguity in this configuration of quantifiers provides prima facie evidence against Frey's (1993) configurational account. Inverse scope readings seem to exist even in constructions where the second quantifier does not c-command a trace of the first quantifier. Two paired t-tests were computed as follow-up analyses in order to break down the interaction. Interestingly, these analyses revealed that the ambiguity affected both readings. It was not only the case that set diagrams disambiguating towards inverse scope were judged better in the ambiguous construction than in the disambiguated control condition ($t_1(47) = 5.29, p < .01$; $t_2(35) = 4.65, p < .01$), but also the reverse was true: Set diagrams only consistent with the strongly preferred linear interpretation were judged significantly worse in the ambiguous condition than in the appropriate control condition ($t_1(47) = –3.41, p < .01$; $t_2(35) = –3.43, p < .01$).

In the OS1 conditions the object quantifier *genau ein* ('exactly one') preceded the distributive subject quantifier *jeder* ('every'). The potentially ambiguous construction was indeed compatible with both scope readings, but the inverse reading was somewhat preferred over the linear interpretation (mean z-scores: 0.16 vs. –0.15). The scope unambiguous OS1 control conditions patterned with the true and false fillers. They were accepted when combined with a picture only compatible with the linear interpretation and rejected when combined with an inverse picture. ANOVAs revealed significant main effects of AMBIGUITY ($F_1(1,47) = 35.0, p < .01$; $F_2(1,35) = 52.1, p < .01$) and of DISAMBIGUATION ($F_1(1,47) = 74.2, p < .01$; $F_2(1,35) = 98.9, p < .01$) and a significant interaction between AMBIGUITY and DISAMBIGUATION ($F_1(1,47) = 109.4, p < .01$; $F_2(1,35) = 197.2, p < .01$). Paired t-tests between the linear and the inverse ambiguous conditions revealed that the difference was marginal by participants and significant by items ($t_1(47) = 1.88, p = .07$; $t_2(35) = 2.58, p < .05$). Thus, fully in line with Pafel's (2005) prediction (cf. (16)) in this particular configuration of quantifiers inverse scope was slightly preferred over the linear reading.

The results of the OS2 conditions are also fully in line with Pafel's model. In the ambiguous construction linear scope was strongly preferred over inverse scope (mean z-scores: 0.40 vs. –0.60) but the difference between the two readings was less pronounced than in the disambiguated sentences (mean z-scores: 0.56 (linear) vs. –0.89 (inverse)). Thus, although strongly dispreferred, the inverse reading is still possible in the ambiguous OS2 construction. ANOVAs revealed a significant main effect of DISAMBIGUATION ($F_1(1,47) = 165.57$, $p < .01$; $F_2(1,35) = 345.19$, $p < .01$) and a significant interaction between AMBIGUITY and DISAMBIGUATION ($F_1(1,47) = 10.24$, $p < .01$; $F_2(1,35) = 8.15$, $p < .01$) but the main effect of AMBIGUITY was not significant ($F_{1/2} < 1$). The interaction is due to the fact that the difference between the judgments for the two scope readings was larger in the unambiguous cases than in the ambiguous ones. In particular, the mean z-score of –0.60 for the inverse scope in the ambiguous construction was still significantly higher than the z-score for the unambiguous control, as shown by paired t-tests: $t_1(47) = 2.48$, $p < .05$ and $t_2(35) = 2.26$, $p < .05$.

Comparing the relative differences between the linear and the inverse disambiguation in the SO and the OS1 conditions we see that Pafel's predictions are not met in the SO conditions. Under his theory we would have expected differences of the same size in both construction types. This was clearly not the case. While the linear and the inverse interpretation were more or less equally available in the ambiguous OS1 conditions, the SO conditions showed a strong preference for linear scope. So, could Frey's account be on the right track after all? As outlined in the introduction, Frey acknowledges the influence of non-syntactic factors such as lexical properties of quantifiers and intonation. We have shown in section 2 that the tested quantifiers are quantificational expressions according to Frey's criteria. However, participants may sometimes have implicitly chosen an intonation that made inverse scope available (see, e.g. Fodor 2002 for empirical evidence for implicit prosody during silent reading). More specifically, Krifka (1998), building on work by Büring (1997), has proposed that a rise-fall intonation contour might add the inverse reading, which is impossible otherwise. If this is correct, the ambiguity observed in the SO conditions may be due to the fact that participants sometimes chose a rise-fall intonation contour in order to license inverse scope. Note that this could also explain the difference between the findings of the present experiment and Bott & Schlotterbeck's (2012) results who did not find any evidence for inverse scope in the same construction when the sentences were presented in a self-paced fashion. It is plausible that in that kind of task readers are not able to construe the sentences with a rather marked rise-fall intonation contour. In an offline task as the one used here, however, choosing the appropriate implicit prosody is probably much easier. We therefore ran a pilot study explicitly addressing the role of intonation for scope inversion in *subject-before-object* sentences. The experiment used picture verification with auditory presentation of stimuli controlled for intonation.

3.3 The potential influence of intonation (Pilot study)

We constructed 20 items in four conditions according to a 2 × 2 (INTONATION × DISAMBIGUATION) within design. Target sentences with the universal distributive subject quantifier *jeder* preceding the object quantifier *genau ein* ('exactly one') were embedded in short dialogues:[8]

[8] An anonymous reviewer remarked that the inverse scope answer in (18a) may be excluded on pragmatic grounds: It may simply be irrelevant how many pieces the children share. To us, the question itself is ambiguous allowing for an interpretation *is it many black pieces that belong to both Andrea and Wolfgang?* licensing an inverse scope construal of the answer. When designing the experiment, we decided against

(18) a. A: Stimmt es, dass Andrea und Wolfgang viele schwarze Spielplättchen haben?
B: Nein, JE/des Kind hat genau EIN\schwarzes Spielplättchen in seiner Spielecke.
(A: 'Is it correct that Andrea and Wolfgang have many black pieces?'
B: 'No, every child has exactly one black piece in its corner.')

b. A: Wieviele schwarze Spielplättchen haben Andrea und Wolfgang?
B: Jedes Kind hat genau ein schwarzes Spielplättchen in seiner Spielecke.
(A: 'How many black pieces do Andrea and Wolfgang have?'
B: 'Every child has exactly one black piece.')

Each dialogue was paired with two disambiguating pictures such as those in Figure 4. The pictures always showed a table with partly overlapping play areas ("corners") of two children playing with geometrical pieces of various colours (sample item in black and white for printing reasons). Participants' task was to decide whether B's answer corresponds to a true description of the scenario. Speaker B's answer in the dialogue (18a) is an informative answer to the question under both scope readings. In its linear scope construal the answer is that neither of the two children has many black pieces because each has only one (cf. Figure 4(a)). Under the inverse scope construal, the answer says that there is only one black piece that Andrea and Wolfgang have in their joint area (cf. Figure 4(b)), which again shouldn't count as an instance of *many*. The dialogue (18b) served as baseline control. Here, only the linear reading should be possible.

Dialogues were spoken and recorded by a male (speaker A) and a female speaker (speaker B). Speaker B was instructed to produce the answer in the rise-fall dialogues (18a) with a rise-fall intonation. For the control condition (18b) speaker B was asked to use what seemed the most appropriate intonation for the answer given the question. The produced intonation contours of speaker B were analyzed with respect to their fundamental frequency using the PRAAT software package (Boersma 2001). The average intonation contours in the two dialogue conditions are shown in Figure 5. In the control condition about half of the items were realized with two foci on the two quantifiers and the other half with verum focus on the auxiliary. While recording the stimuli, speaker B changed

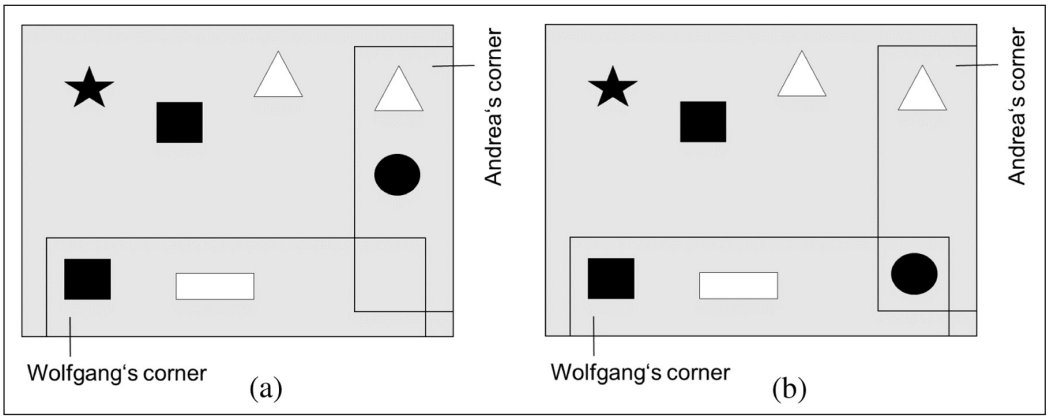

Figure 4: Diagrams used in the Pilot Study: **(a)** Diagram for $\forall\exists^{=1}$-reading; **(b)** Diagram for $\exists^{=1}\forall$-reading.

presenting decontextualised statements without a preceding question because naive informants might not be able to infer the underlying information structure. We agree that more careful testing is required. This is why we were hesitant to call this pilot study a "real experiment". Future work is required to confirm the assumptions underlying the pilot study.

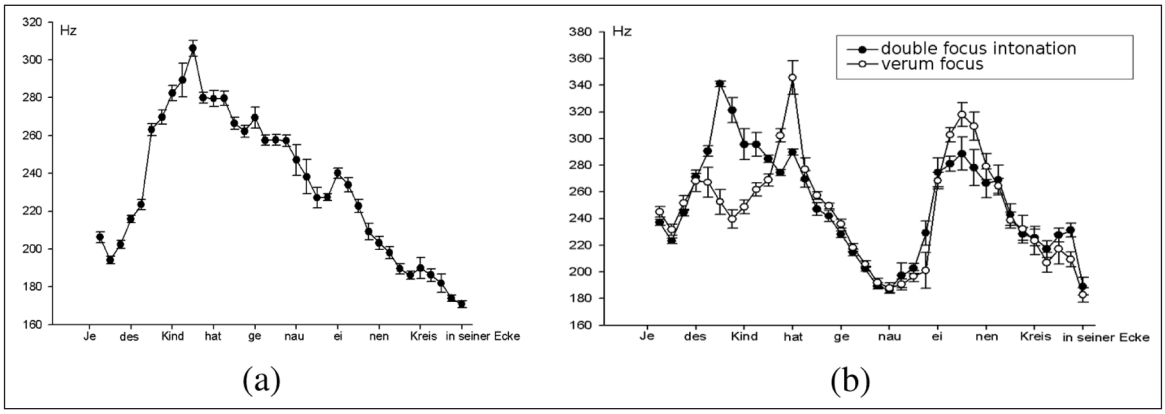

Figure 5: Intonation contours of the target sentences in the pilot study: **(a)** Mean F0-contour rise-fall used for dialogues of type (18a); **(b)** Mean F0-contour rise-fall used for dialogues of type (18b). Plots show mean F0-values split by sentence regions and 95% confidence intervals.

between these intonation contours and stated afterwards that to her both were possible answers to the *how many* question in (18b).

20 native German participants judged the 20 items plus 60 filler dialogues in a latin square design making sure that they saw each item in only one condition. After inspecting the picture and listening to the auditorily presented dialogue they had to judge whether the answer of speaker B was a true description of the picture by providing a "yes, true" or "no, false" answer.

Acceptance rates were as follows. In the control condition the linear disambiguation was accepted 98% of the time, whereas pictures only consistent with the inverse interpretation were accepted 17% of the time. Acceptance rates were almost identical in the rise-fall conditions: Linear scope disambiguations were accepted 96% of the time, whereas inverse scope disambiguations were accepted 20% of the time. A loglinear analysis on *yes, acceptable* versus *no, unacceptable* responses including the factors CONTOUR and DISAMBIGUATION revealed that the factor CONTOUR did not contribute to model fit, neither as a main effect nor in the form of an interaction with the other factors ($\chi^2(4) = .51$; $p = .97$). Thus, acceptance rates did not differ between conditions with a rise-fall contour and the control conditions.

This provides preliminary evidence against the view that doubly quantified German sentences with a subject-before-object word order can receive an inverse scope interpretation, or receive it more easily, if realized with a rise-fall intonation. Coming back to the results of Experiment 1 we thus conclude that it is unlikely that participants made inverse scope available by implicitly imposing a rise-fall intonation. We would like to emphasize, though, that the influence of intonation on scope interpretation needs to be investigated more carefully. We must leave this for future research.

4 The interplay of scope factors (Exp. 2 and 3)

Experiment 1 addressed the first two research questions. The observed scope ambiguity in the subject-before-object condition provided evidence against Frey's (1993) theory, but the pattern of results is not fully consistent with Pafel's (2005) multi-factor account. The following two picture verification task experiments investigated the combined effects of LINEAR ORDER, DISTRIBUTIVITY and DISCOURSE ANAPHORICITY in doubly quantified sentences. Their aim was to answer the third research question from section 1.2: *Do we find evidence for strictly additive influences of the different scope factors as predicted by multi-factor theories?* Since answering this question may make it necessary to measure rather subtle differences between sentence conditions, we employed the

Magnitude Estimation technique (Bard et al. 1996) which has been widely applied in experimental syntax, e.g. Keller (2000); Featherston (2005) but see Weskott & Fanselow (2011); Sprouse (2011).

4.1 Experimental materials

The constructions tested in Experiment 2 and 3 are presented in (19) and (20). The complete list of items is provided in Appendix B.

(19) a. Genau einen dieser Aufsätze hat jeder Student gelesen.
exactly one of-these papers.ACC has each student.NOM read
'Each student read exactly one of these papers.'

b. Genau einen dieser Aufsätze haben alle Studenten gelesen.
exactly one of-these papers.ACC have all students.NOM read
'All students read exactly one of these papers.'

c. Jeder Student hat genau einen dieser Aufsätze gelesen.
each student.NOM has exactly one of-these papers.ACC read
'Each student read exactly one of these papers.'

e. Alle Studenten haben genau einen dieser Aufsätze gelesen.
all students.NOM have exactly one of-these papers.ACC read
'All students read exactly one of these papers.'

In (19) the factors LINEAR ORDER and DISTRIBUTIVITY are manipulated. In (19a) and (19b) the $\exists^{=1}$-quantifier precedes $\forall$ whereas in (19c) and (19d) the order of quantifiers is reversed. The universal quantifier *jeder* ('every') used in (19a) and (19c) is distributive whereas *alle* ('all') in (19b) and (19d) is not. The factor DISCOURSE ANAPHORICITY is kept constant across conditions.

(20) a. Genau einen Aufsatz haben alle diese Studenten gelesen.
exactly one paper.ACC have all these students.NOM read
'All (of) these students read exactly one paper.'

b. Genau einen dieser Aufsätze haben alle Studenten gelesen.
exactly one of-these papers.ACC have all students.NOM read
'All students read exactly one of these papers.'

c. Genau einen Aufsatz hat jeder dieser Studenten gelesen.
exactly one paper.ACC has each these students.NOM read
'Each of these students read exactly one paper.'

d. Genau einen dieser Aufsätze hat jeder Student gelesen.
exactly one of-these papers.ACC has each student.NOM read
'Each student read exactly one of these papers.'

In (20) the factors DISTRIBUTIVITY and DISCOURSE ANAPHORICITY are manipulated. In (20b) and (20d) the first quantifier is related to a contextually salient referent, whereas in (20a) and (20c) it is the second quantifier that is linked to preceding discourse. In Experiment 3 LINEAR ORDER was kept constant across conditions: The sentences always had object-before-subject word order.

4.2 Manipulating linear order and distributivity (Exp. 2)

In Experiment 2, we manipulated LINEAR ORDER ($\forall$ before $\exists$ vs. $\exists$ before $\forall$) and DISTRIBUTIVITY (*jeder* vs. *alle*). A potentially scope ambiguous sentence like (19a) to (19d) was combined with a set diagram either disambiguating towards $\exists^{=1}\forall$ or towards $\forall\exists^{=1}$.

4.2.1 Methods

Participants: 56 native German speakers from Tübingen University (26 female, mean age 25.6 years) participated in the experiment for 5€. 5 additional participants were excluded from the analysis because of poor performance on the fillers.

Materials and design: We used the 24 doubly quantified transitive sentences in the four variants in (19). The universally quantified phrase was the subject and the existentially quantified and discourse anaphoric *exactly one of these*-quantifier was the direct object. Each item was paired with two disambiguating set diagrams like the ones in Figure 1 resulting in a $2 \times 2 \times 2$ factorial design with the within factors LINEAR ORDER, DISTRIBUTIVITY and DISAMBIGUATION. Additionally, we prepared 36 distractor sentences using different quantifiers, negation and definite descriptions. 16 fillers were true and 20 were false. We constructed eight lists according to a latin square design.

Procedure: Judgments were gathered using the Magnitude Estimation method Bard et al. (1996). Participants judged sentence-picture pairs relative to a reference item to which they had assigned an arbitrary value first. High values indicated that the sentence fits the picture, low values indicated a mismatch.

The experiment was a paper-and-pencil questionnaire. After reading written instructions, participants first completed a short practice session consisting of five trials. In the following experimental session, sentence-picture pairs were presented in an individually randomized order. Each sentence-picture pair was presented on a separate page in a small booklet with the reference item on top of each page (*Many hunters have shot two rabbits* in a situation in which 4 out of 10 hunters shot two and the rest shot none). Participants were explicitly instructed not to go back to earlier items and to complete the experiment in a page-by-page fashion.

4.2.2 Results and discussion

Each participant's judgments were normalized by transforming them into z-scores. The true fillers were rated with a mean z-score of 0.76 (standard deviation 0.80) and the false fillers with a mean z-score of –0.74 (standard deviation 0.74). Thus, participants paid attention to the semantic properties of the sentences.

The mean judgments are shown in Figure 6. LINEAR ORDER and DISTRIBUTIVITY both showed effects in the expected direction. A quantifier that preceded another tended to

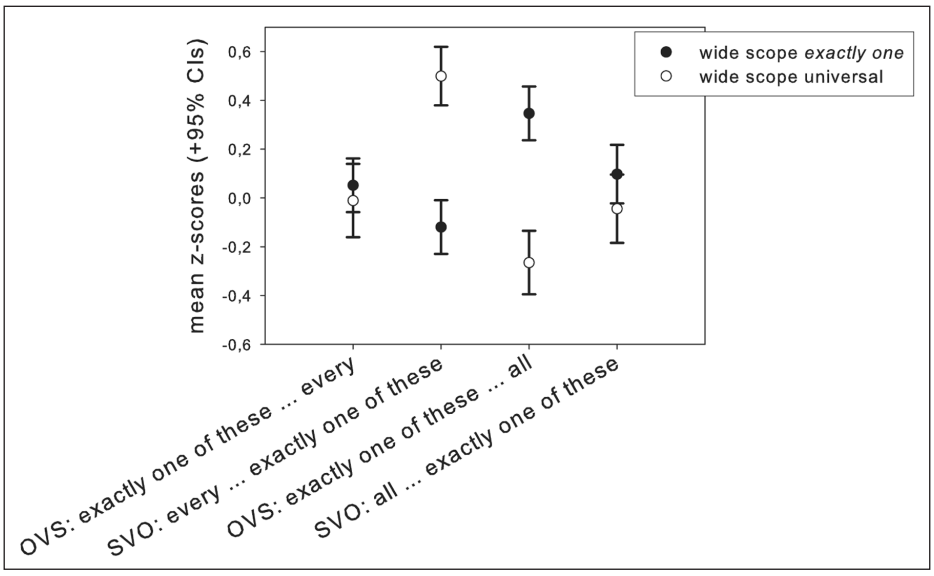

Figure 6: Judgments in Experiment 2 (plus 95% confidence intervals).

take wide scope. Similarly, distributive quantifiers with the determiner *jeder* ('every') took wide scope more easily than non-distributive *alle* ('all'). We computed repeated measures ANOVAs on the z-transformed judgments with the within factors DISTRIBUTIVITY (*jeder* vs. *alle*), ORDER ($\exists^{=1}$ BEFORE $\forall$ vs. $\forall$ BEFORE $\exists^{=1}$) and DISAMBIGUATION ($\forall\exists^{=1}$ vs. $\exists^{=1}\forall$). These revealed significant two-way interactions of DISTRIBUTIVITY and DISAMBIGUATION ($F_1(1,55) = 80.06$, $p < .01$; $F_2(1,23) = 54.15$, $p < .01$) and ORDER and DISAMBIGUATION ($F_1(1,55) = 20.71$, $p < .01$; $F_2(1,23) = 41.73$, $p < .01$). Besides these interactions only the interaction between DISTRIBUTIVITY and ORDER was significant ($F_1(1,55) = 5.40$, $p < .05$; $F_2(1, 23) = 4.72$, $p < .05$). This was due to the fact that sentences with *jeder* before *genau ein* received higher ratings than the other constructions irrespective of the disambiguation. Crucially, the three-way interaction between ORDER, DISTRIBUTIVITY and DISAMBIGUATION was not significant ($F_1(1,55) = 1.29$; $p = .26$; $F_2(1,23) = 1.37$; $p = .26$). This shows that the prediction of purely additive effects of combined scope factors is borne out.

Planned comparisons (with a Bonferroni corrected α of .0125) revealed significant differences between the $\exists^{=1}\forall$- and the $\forall\exists^{=1}$-disambiguation in the *jeder-before-ein*-construction ($t_1(55) = 5.36$, $p < .0125$; $t_2(23) = 6.12$, $p < .0125$) and in the *ein-before-alle*-construction ($t_1(55) = -5.21$, $p < .0125$; $t_2(23) = -6.59$, $p < .0125$) whereas the other two constructions showed no difference (all ts < 1.6). Sentences with distributive *jeder* were fully compatible with both readings when *genau ein* preceded *jeder* but were preferred $\forall\exists^{=1}$ when *jeder* was the first quantifier; the pattern was exactly the opposite for sentences containing the non-distributive determiner *alle*: *Alle* before *ein* was ambiguous but *ein* before *alle* favored a surface scope interpretation.

Figure 7 shows the observed difference scores between the mean judgments for the $\forall\exists^{=1}$ disambiguation minus the mean judgments for the $\exists^{=1}\forall$ disambiguation.[9] The observed difference scores are plotted against the difference scores predicted by Pafel's (2005) model. As we can see, apart from minor exceptions there is a rather close correspondence between the actual scope judgments and the values predicted by the theory.

The results of Experiment 2 show that both LINEAR ORDER and DISTRIBUTIVITY have an effect on the scope of doubly quantified sentences. Quantifiers containing the distributive determiner *jeder* ('every') take wide scope more easily than quantifiers with the non-distributive determiner *alle* ('all'). Similarly for LINEAR ORDER: linear scope is preferred over inverse readings. Interestingly, LINEAR ORDER and DISTRIBUTIVITY showed purely additive effects. The three-way interaction of ORDER, DISTRIBUTIVITY and DISAMBIGUATION was absent. This pattern fits the predictions of multi-factor theories like Kuno (1991) and Pafel (2005) but without thresholds. It is incompatible with the predicted unter-additive effects of scope theories incorporating scope factors into the syntactic configuration.

[9] Before the actual comparison the observed and the predicted difference scores were z-transformed for better comparability. Without transformation, the two scores are admittedly rather different from each other. One of the effects of the transformation is that it corrects for any main effects of DISAMBIGUATION. We consider this a welcome result. To see why, the reader is invited to consider the different ratings for the sentence conditions *exactly one of the … every* and *exactly one of the … all*, both with an OVS order in Exp. 2 and 3. These conditions were the same across experiments, however the results show a general bias towards a linear interpretation in Exp. 3 relative to Exp. 2. The reason for this difference could be that all the experimental sentences in Exp. 3 were OVS sentences whereas in Exp. 2 there were as many SVO as OVS sentences. This may result in what could be called an effect of EXPERIMENTAL CONTEXT. The influence of the factor SYNTACTIC FUNCTION is probably weakened once participants get used to a high proportion of OVS sentences. Consistent with the multi-factor framework this should result in a stronger preference for linear scope. Another case in point is LF priming as observed in experiments by Raffray & Pickering (2010). Their experiments show that the availability of a scope reading increases if the same scope was assigned in the trial before. Effects such as these make it difficult to directly relate linguistic theory to experimental data. Our solution in the present paper is to only interpret relative differences between the conditions tested in an experiment; and this is exactly what z-transformation achieves.

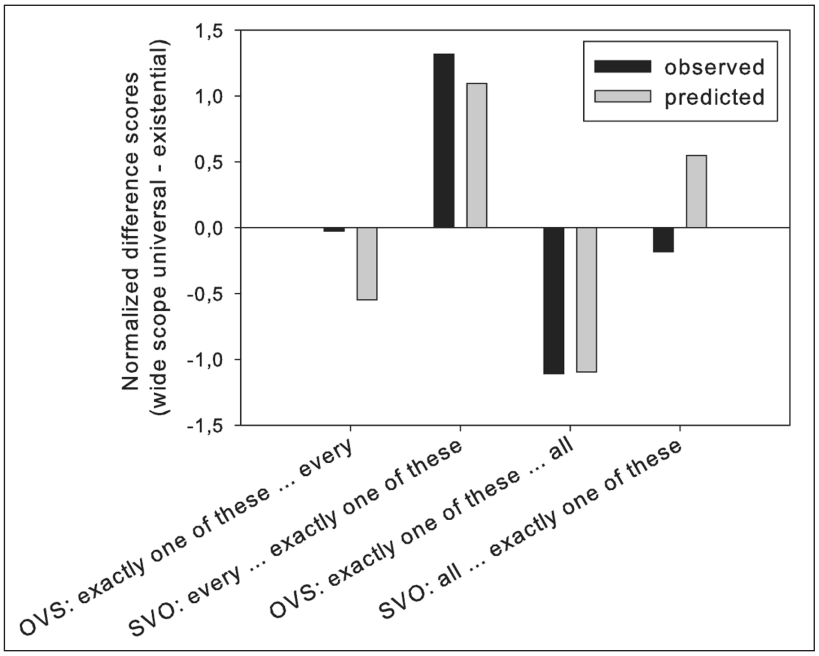

Figure 7: Observed and predicted difference scores in Experiment 2. Both observed and predicted mean difference scores in the four sentence conditions were normalized via z-transformations for better comparability.

4.3 Manipulating distributivity and discourse anaphoricity (Exp. 3)

Discourse anaphoricity of the restrictor set has not received very much attention in the literature on quantifier scope. An exception is Musolino & Gualmini (2004), who investigated inverse scope construals in children's understanding of sentences including negation and quantifiers such as (21).

(21) a. The smurf didn't catch two (of the) birds.
b. The troll didn't find some of the jewels.

They found that discourse anaphoric quantifiers significantly increased children's ability to compute non-isomorphic, inverse interpretations, which were largely absent in sentences without the partitive. Here, we employed this factor to test whether and how discourse related scope factors interact with lexical factors such as distributivity. Thus in Experiment 3 we manipulated DISTRIBUTIVITY and DISCOURSE ANAPHORICITY. Consider the sample item (22) repeated from (20).

(22) a. Genau einen Aufsatz haben alle diese Studenten gelesen.
exactly one paper.ACC have all these students.NOM read
'All (of) these students read exactly one paper.'

b. Genau einen dieser Aufsätze haben alle Studenten gelesen.
exactly one of-these papers.ACC have all students.NOM read
'All students read exactly one of these papers.'

c. Genau einen Aufsatz hat jeder dieser Studenten gelesen.
exactly one paper.ACC has each these students.NOM read
'Each of these students read exactly one paper.'

d. Genau einen dieser Aufsätze hat jeder Student gelesen.
exactly one of-these papers.ACC has each student.NOM read
'Each student read exactly one of these papers.'

The universal quantifier *alle* in (22a) and (22b) is non-distributive; *jeder* in (22c) and (22d) is distributive. The discourse-anaphoric phrase *dieser* links one of the quantifiers to the context: In (22a) and (22c) it is the universal quantifier and in (22b) and (22d) it is the existential quantifier that is discourse anaphoric.

Multi-factor theories predict that these factors should show purely additive effects. If the factors are explicitly encoded in the hierarchical configuration, however, we expect asymmetrical dependencies. Manipulating the factor encoded in a higher position of the tree should block effects due to a manipulation of a factor encoded in some lower projection in the tree. Experiment 3 tested these predictions.

4.3.1 Methods

The sentence materials used in Experiment 2 were adjusted to the design in (22). We manipulated DISTRIBUTIVITY and DISCOURSE ANAPHORICITY but kept word order constant with *genau ein* ('exactly one') always preceding the universal quantifier. The sentences were paired with the same disambiguating pictures as in Experiment 2 yielding a $2 \times 2 \times 2$ (DISTRIBUTIVITY $\times$ DISCOURSE ANAPHORICITY $\times$ DISAMBIGUATION) within design. The same filler sentences were used as in the previous experiment. The procedure was the same as in Experiment 2.

Participants: 24 native German speakers from Tübingen University (10 female, mean age 26.2 years) took part in the experiment. Each participant received 5€. Four additional participants were excluded from the analysis due to poor performance on the fillers.

4.3.2 Results and discussion

The true fillers were rated with a mean z score of 0.89 (standard deviation 0.68) and the false distractors with a mean z score of –0.70 (standard deviation 0.80).

The distribution of scope readings in Experiment 3 is shown in Figure 8. We computed repeated measures ANOVAs on z-transformed judgments with the within factors DISTRIBUTIVITY (*jeder* vs. *alle*), DISCOURSE ANAPHORICITY (DISCOURSE ANAPHORIC $\exists^{=1}$ vs. DISCOURSE ANAPHORIC $\forall$) and DISAMBIGUATION ($\exists^{=1}\forall$ vs. $\forall\exists^{=1}$). Both factors showed an influence on the distribution of scope readings. A discourse anaphoric quantifier tended to take wide scope more easily than a non-discourse anaphoric quantifier resulting in a significant interaction between DISCOURSE ANAPHORICITY and DISAMBIGUATION ($F_1(1,23)$

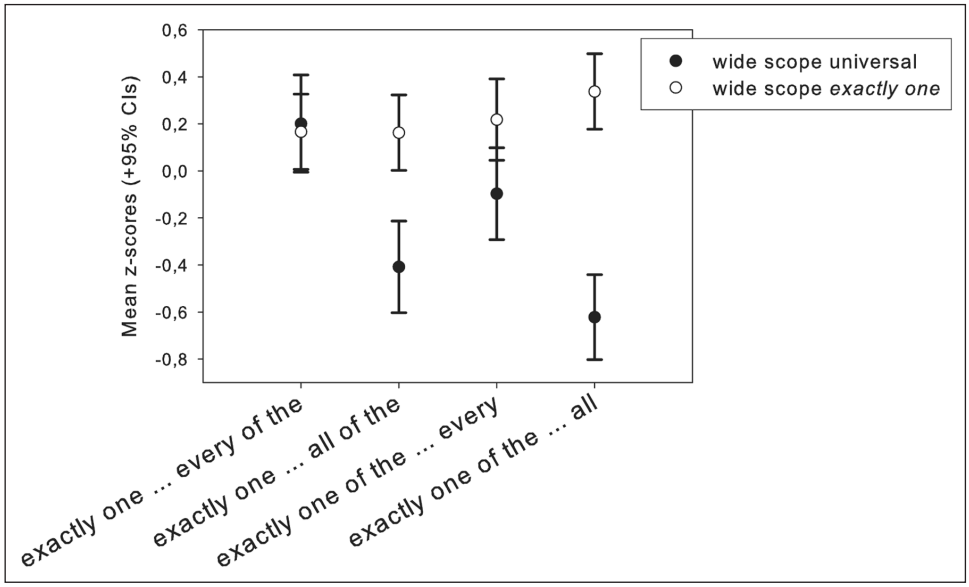

Figure 8: Judgments in Experiment 3 (plus 95% confidence intervals).

= 7.33, $p < .05$; $F_2(1,23) = 7.99$, $p < .05$). Furthermore, distributive *jeder* ('every') had a stronger tendency for wide scope than non-distributive *alle* ('all') leading to a significant interaction between DISTRIBUTIVITY and DISAMBIGUATION ($F_1(1,23) = 17.00$, $p < .01$; $F_2(1,23) = 19.47$, $p < .01$). Beside these effects there was a main effect of DISAMBIGUATION ($F_1(1,23) = 13.92$, $p < .01$; $F_2(1,23) = 32.55$, $p < .01$) which reflects a general preference for the $\exists^{=1}\forall$-interpretation. This preference is probably due to constant word order in Experiment 3 with the existential quantifier always preceding the universal quantifier. The main effect of DISTRIBUTIVITY was also significant ($F_1(1,23) = 13.48$, $p < .01$; $F_2(1,23) = 13.05$, $p < .01$). This is due to the fact that the wide-scope universal disambiguations were judged better for *jeder* than for *alle* ($t_1(23) = 4.78$, $p < .01$; $t_2(23) = 4.88$, $p < .01$). Crucially, there was no three-way interaction between DISCOURSE ANAPHORICITY, DISTRIBUTIVITY and DISAMBIGUATION ($F_1(1,23) = .04$; $p = .85$; $F_2(1,23) = .05$; $p = .83$). Thus, DISCOURSE ANAPHORICITY and DISTRIBUTIVITY showed purely additive effects.

Figure 9 shows the observed difference scores between the mean judgments for the linear disambiguation minus the mean judgments for the inverse disambiguation. The observed difference scores are plotted against the difference scores predicted by Pafel's (2005) theory. As in Experiment 2, there is a rather close correspondence between the actual scope judgments and the values predicted by the theory. Again, this is only true for observed and predicted values after z-transformation (cf. fn. 9). We hypothesize that the stronger-than-expected tendency for linear interpretations in the present experiment compared to the previous one is due to a higher proportion of OVS sentences in Exp. 3. Note that a decrease in the factor weight for SUBJECTHOOD leads to a stronger preference for linear scope in OVS sentences. We consider it an interesting question for future research to investigate how contextual factors such as the broader context of the experiment probabilistically contribute to the relative scope of quantifiers.

This experiment provides evidence that DISCOURSE ANAPHORICITY also affects scope preferences: Discourse anaphoric quantifiers take wide scope more easily than non-anaphoric quantifiers. Again, the effects of the scope factors were purely additive, a result

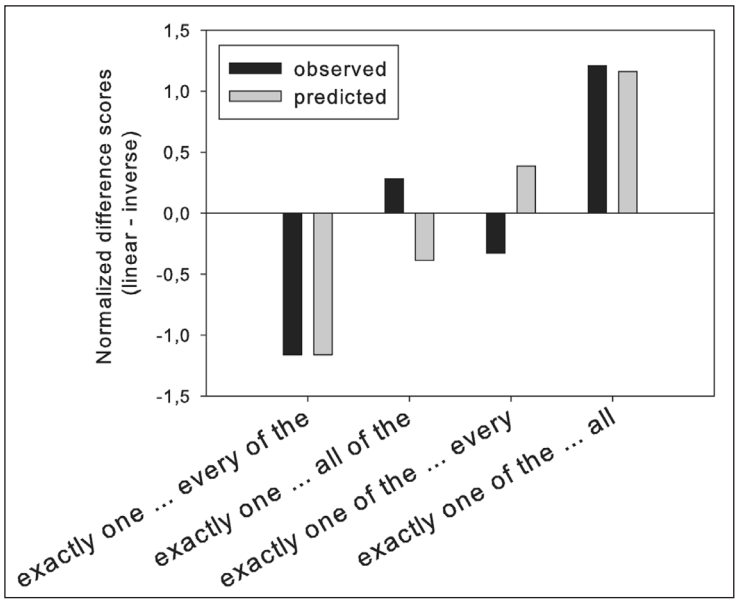

Figure 9: Observed and predicted difference scores in Experiment 3. Both observed and predicted mean difference scores in the four sentence conditions were normalized via z-transformations for better comparability.

that corroborates the findings from Experiment 2. This pattern of effects is exactly what is predicted by multi-factor theories of quantifier scope.

5 General discussion

The results of the reported picture verification task experiments showed that the scope relations in doubly quantified sentences depend on a number of factors: the linear order of the quantifiers, whether they are distributive and whether they are explicitly restricted by context. All three factors showed cumulative effects, that is, they added up rather than interacting with each other.

In section 1 we gave an overview of theories on quantifier scope. Simplifying matters a great deal we focussed on two types of theories on scope interaction: configurational and multi-factor theories. Furthermore, as the presented experiments measured scope preferences in German, we chose the configurational theory by Frey (1993) and contrasted it with the multi-factor account proposed in Pafel (2005). We now discuss the implications of the experimental findings for each of them in turn.

The configurational theories we considered here disambiguate the relative scope of quantifying expressions in the syntax. In Frey's account inverse scope in German can only arise in configurations in which the second quantifier has syntactic command over a trace left in the base position of a quantifier in a fronted position. The results of Experiment 1 provide evidence against Frey's scope principle. Using quantificational expressions that satisfy Frey's constraints, inverse scope was more acceptable in the tested *subject-before-object* construction than in the unambiguous baseline control condition. According to our criterion for determining scope ambiguity, we therefore conclude that the first research question *are German doubly quantified subject-before-object sentences scope unambiguous* must be answered in the negative. This claim is further supported by our pilot study suggesting that inverse scope in this construction type cannot be explained by proposed prosodic influences on quantifier scope in German. We would also like to point out that we found strong differences in scope preferences for sentences with fronted object quantifiers. As it stands, Frey's theory does not make any specific predictions for these except for the general claim that they should in principle allow for both linear and inverse scope construals. For these cases, the empirical coverage of the theory is thus rather limited and something needs to be added to explain the observed differences.

In the introduction we mentioned configurational proposals that encode different scope factors in dedicated positions in the LF. In its original form (May 1977) the movement operation was fully unconstrained but later developments like Beghelli & Stowell (1997) make use of dedicated scope positions to account for individual properties of quantifiers. The purely additive nature of effects is, however, problematic for this type of theory because manipulating a factor in a higher scope position should preclude the influence of another factor in a lower position. In Experiments 2 and 3 we did not find any evidence of asymmetrical dependencies (under-additive interactions) between factors. This provides prima facie evidence against an account trying to accommodate the investigated factors in a hierarchical structure.

Unconstrained versions of configurational accounts would predict all combinatorially possible scope readings to be available, at least in principle. This is in line with the present findings. What needs to be done is to work out an explanative model of scope preference incorporating the interplay of the different factors here shown to be relevant for the relative scope of quantifiers. Furthermore, we have argued in the introduction that the presented data pose some constraints on such a model, namely that the factors that filter out

particular scope readings have to contribute in parallel. We hope that the data presented here will serve as a valuable data source for working out such a model.

The findings of our experiments lend some support to multi-factor theories of quantifier scope. First, relative scope is influenced by a number of factors such as LINEAR ORDER, DISTRIBUTIVITY and DISCOURSE ANAPHORICITY. Second, scope preferences were graded rather than categorical and, at least for object-before-subject sentences, largely fit the predictions derived from the tested linear model of scope preference. Third, the factors interact in a purely additive fashion. Contrary to Pafel's (2005) model, our findings do not support the assumption of thresholds distinguishing ambiguous from unambiguous sentences. All the constructions we tested in our experiments were scope ambiguous. Experiment 1 makes the strongest case for this claim. Although one scope reading was extremely preferred over the other, the dispreferred interpretation was still judged better than a clearly unavailable one in scope disambiguated controls. Moreover, the ratings of canonical subject-before-object sentences in Experiment 1 indicated that Pafel's (2005) account makes the wrong predictions for this kind of sentences. Here, the theory clearly has to be adjusted to fit the data. We have also seen that the broader experimental context has an influence on scope judgments. Identical conditions tested in Exp. 2 and 3 received consistently different judgments. This shows that we need additional linking assumptions to relate linguistic theory to experimental work on scope preferences.

Besides the theoretical implications, our study adds to the psycholinguistic work on quantifier interaction. Previous results were not sufficient to decide whether LINEAR ORDER has an influence above and beyond GRAMMATICAL FUNCTION. While VanLehn (1978), Fodor (1982), Gillen (1991), and Kurtzman & MacDonald (1993) provided support in favor of this claim, Ioup (1975), Catlin & Micham (1975), and Micham et al. (1980) argued that the grammatical function of the quantifiers is more important than linear order. The constructions we used made it possible to tease apart the two factors. We found a strong effect of LINEAR ORDER indicating that the order of quantifiers influences scope independent of grammatical function. Furthermore, our study lends support to the findings of Tunstall (1998), Bott & Radó (2009), Radó & Bott (2012), and Brasovaeanu & Dotlacil (2015) in showing effects of DISTRIBUTIVITY. Finally, it adds DISCOURSE ANAPHORICITY to the list of scope factors, which has not been investigated before with respect to scope preferences in adults (Syrett et al. 2016) but has been demonstrated to yield rather strong effects in children's interpretation of scope (Miller & Schmitt 2004; Musolino & Gualmini 2004).

What is more, our study is one of the first experimental investigations on the combined influence of multiple scope factors. We have demonstrated that the factors claimed to be relevant for scope interpretation show purely additive effects as predicted by multi-factor models that take into account the linear combinations of multiple constraints. In spite of being descriptively adequate, multi-factor theories are not very popular in the linguistic literature. There are reasons for this. So far, this type of theory just states that a factor is relevant for quantifier scope without offering an explanation why this should be the case. To also strive for explanatory adequacy, multi-factor theories need to address this question.

Abbreviations

1ST = linearily first quantifier phrase, SUBJ = subject quantifier phrase, DIST = distributive quantifier phrase, D-ANA = discourse anaphoric quantifier phrase, NOM = nominative, ACC = accusative, FEM = feminine

Additional Files

The additional files for this article can be found as follows:

- **Appendix A.** Items tested in Exp. 1. DOI: https://doi.org/10.5334/gjgl.309.s1
- **Appendix B.** Items tested in Exp. 2/3. DOI: https://doi.org/10.5334/gjgl.309.s1

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft within the SFB 441 "Linguistische Datenstrukturen" and SFB 833 "Bedeutungskonstitution", as well as the project "Composition in Context" in the DFG Priority Program *XPrag.de* "New Pragmatic Theories Based on Experimental Evidence". Thanks to Jürgen Pafel and Wolfgang Sternefeld for helpful discussions. We are also very grateful to three anonymous reviewers for insightful comments that have significantly improved this paper. All remaining errors are our own.

Competing Interests

The authors have no competing interests to declare.

References

Anderson, Catherine. 2004. *The structure and real-time comprehension of quantifier-scope ambiguity*. Chicago, IL: Northwestern University dissertation.

Bard, Ellen, Dan Robertson & Antonella Sorace. 1996. Magnitude estimation of linguistic acceptability. *Language* 72(1). 32–68. DOI: https://doi.org/10.2307/416793

Barker, Chris. 2002. Continuations and the nature of quantification. *Natural Language Semantics* 10(3). 211–242. DOI: https://doi.org/10.1023/A:1022183511876

Beghelli, Filippo & Tim Stowell. 1997. Distributivity and negation: The syntax of *each* and *every*. In Anna Szabolcsi (ed.), *Ways of scope taking*, 71–107. Dordrecht: Kluwer Academic Publishers.

Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5(9/10). 341–345.

Bott, Oliver & Fabian Schlotterbeck. 2012. Incremental truth value judgments. In Britta Stolterfoht & Sam Featherston (eds.), *Empirical approaches to linguistic theory*, 3–28. Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9781614510888.3

Bott, Oliver & Fabian Schlotterbeck. 2015. The processing domain of scope interaction. *Journal of Semantics* 32(1). 39–92. DOI: https://doi.org/10.1093/jos/fft015

Bott, Oliver & Janina Radó. 2007. Quantifying quantifier scope. In Sam Featherston & Wolfgang Sternefeld (eds.), *Roots. Linguistics in search of its evidential base*. Walther de Gruyter.

Bott, Oliver & Janina Radó. 2009. How to provide exactly one interpretation for every sentence, or what eye movements reveal about quantifier scope. In Susanne Winkler & Sam Featherston (eds.), *The fruits of empirical linguistics* 1. Process, Berlin: Walther de Gruyter. DOI: https://doi.org/10.1515/9783110216141.25

Brasovaeanu, Adrian & Jakub Dotlacil. 2015. Strategies for scope taking. *Natural Language Semantics* 23(1). 1–19. DOI: https://doi.org/10.1007/s11050-014-9109-1

Büring, Daniel. 1997. The great scope inversion conspiracy. *Linguistics and Philosophy* 20. 175–194. DOI: https://doi.org/10.1023/A:1005397026866

Catlin, Jack & Dennis L. Micham. 1975. Semantic representations as procedures for verification. *Journal of Psycholinguistic Research* 4(3). 209–225. DOI: https://doi.org/10.1007/BF01066927

Chierchia, Gennaro & Sally McConnell-Ginet. 2000. *Meaning and grammar*. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1990. *Types of A'-dependencies*. Cambridge, MA: MIT Press.
Dalrymple, Mary, Makoto Kanazawa, Yookyung Kim, Sam McHombo & Stanley Peters. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy* 21(2). 159–210. DOI: https://doi.org/10.1023/A:1005330227480
Dwivedi, Veena. 2013. Interpreting quantifier scope ambiguity: Evidence of heuristic first, algorithmic second processing. *PLoS ONE* 8(11). e81461. DOI: https://doi.org/10.1371/journal.pone.0081461
Featherston, Sam. 2005. That-trace in German. *Lingua* 115(9). 1277–1302. DOI: https://doi.org/10.1016/j.lingua.2004.04.001
Filik, Ruth, Kevin B. Paterson & Simon P. Liversedge. 2004. Processing doubly quantified sentences: Evidence from eye movements. *Psychonomic Bulletin & Review* 11. 953–959. DOI: https://doi.org/10.3758/BF03196727
Fodor, Janet D. 1982. The mental representation of quantifiers. In Stanley Peters & Esa Saarinen (eds.), *Processes, beliefs, and questions*, 129–164. Dordrecht: D. Reidel. DOI: https://doi.org/10.1007/978-94-015-7668-0_5
Fodor, Janet D. 2002. Prosodic disambiguation in silent reading. In Masako Hirotani (ed.), *Proceedings of the North East Linguistic Society (NELS)* 32. 111–132. Amherst, MA: GSLA.
Fodor, Janet D. & Ivan A. Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5(3). 355–398. DOI: https://doi.org/10.1007/BF00351459
Fox, Danny. 2003. On Logical Form. In Randall Hendrick (ed.), *Minimalist syntax*, 82–123. Boston: Blackwell. DOI: https://doi.org/10.1002/9780470758342.ch2
Frey, Werner. 1993. *Syntaktische Bedingungen für die semantische Interpretation: Über Bindung, implizite Argumente und Skopus*. Berlin: Akademie Verlag.
Gillen, Kathryn. 1991. *The comprehension of doubly quantified sentences*. Durham: University of Durham dissertation.
Heim, Irene & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Cambridge, MA: Blackwell.
Hemforth, Barbara & Lars Konieczny. 1999. *German sentence processing* (Studies in Theoretical Psycholinguistics). University of Michigan, MI, Kluwer Academic Publishers.
Hendriks, Herman. 1993. *Studied flexibility: Categories and types in syntax and semantics*. Amsterdam: ILLC Dissertation Series.
Höhle, Tilman N. 1991. On reconstruction and coordination. In Hubert Haider & Klaus Netter (eds.), *Representation and derivation in the theory of grammar*, 139–197. Dordrecht: Kluwer. DOI: https://doi.org/10.1007/978-94-011-3446-0_5
Höhle, Tilman N. 2002. Über Verum Fokus im Deutschen. *Linguistische Berichte* 4. 112–141.
Hornstein, Norbert & Amy Weinberg. 1990. The necessity of LF. *The Linguistic Review* 7. 129–167. DOI: https://doi.org/10.1515/tlir.1990.7.2.129
Ioup, Georgette. 1975. Some universals for quantifier scope. In John Kimball (ed.), *Syntax and Semantics*, 37–58. New York: Academic Press 4th edn.
Jackson, Scott & Will Lewis. 2005. The relation between prosody and logical scope varies by the operator. *Poster presented at the CUNY Conference Tucson*. Arizona.
Keller, Frank. 2000. *Gradience in grammar: Experimental and computational aspects of degrees of grammaticality*. Edinburgh: University of Edinburgh dissertation.
Kempson, Ruth & Annabel Cormack. 1981. Ambiguity and quantification. *Linguistics and Philosophy* 4(2). 259–309. DOI: https://doi.org/10.1007/BF00350141
Kiss, Katalin É. 2006. Quantifier scope ambiguities. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax* 4. 1–34. Malden, MA: Blackwell. DOI: https://doi.org/10.1002/9780470996591.ch53

Krifka, Manfred. 1998. Scope inversion in German under the rise-fall pattern. *Linguistic Inquiry* 29(1). 75–112. DOI: https://doi.org/10.1162/002438998553662

Kroch, Anthony S. 1974. *The semantics of scope in English*. Cambridge, MA: MIT dissertation.

Kuno, Susumu. 1991. Remarks on quantifier scope. In Heizo Nakajima (ed.), *Current English linguistics in Japan*, 261–287. Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110854213.261

Kurtzman, Howard S. & Maryellen C. MacDonald. 1993. Resolution of quantifier scope ambiguities. *Cognition* 48. 243–279. DOI: https://doi.org/10.1016/0010-0277(93)90042-T

May, Robert. 1977. *The grammar of quantification*. Cambridge, MA: MIT dissertation.

May, Robert. 1985. *Logical Form*. Cambridge, MA: MIT Press.

Mayo, Neill, Martin Corley & Frank Keller. 2006. Webexp2: Experimenter's manual. See the manual at: http://www.webexp.info/.

Micham, Dennis L., Jack Catlin, Nancy VanDerven & Kathleen Loveland. 1980. Lexical and structural cues in quantifier scope relations. *Journal of Psycholinguistic Research* 9(4). 367–377. DOI: https://doi.org/10.1007/BF01067449

Miller, Karen & Christina Schmitt. 2004. An experimental study on child comprehension of Spanish indefinites and bare singulars. *Language Acquisition* 12(3/4). 247–256. DOI: https://doi.org/10.1080/10489223.1995.9671745

Montague, Richard. 1973. On the proper treatment of quantification in ordinary English. In Richard H. Thomason (ed.), *Formal philosophy: Selected papers of Richard Montague*, 247–270. New Haven, CT: Yale University Press.

Musolino, Julien & Andrea Gualmini. 2004. The role of partitivity in child language. *Language Acquisition* 12. 97–107. DOI: https://doi.org/10.1207/s15327817la1201_5

Pafel, Jürgen. 2005. *Quantifier scope in German*. (Linguistics Today 84). Amsterdam: John Benjamins.

Paterson, Kevin B., Ruth Filik & Simon P. Liversedge. 2008. Competition during the processing of quantifier scope ambiguities: Evidence from eye movements during reading. *The Quarterly Journal of Experimental Psychology* 61(3). 459–473. DOI: https://doi.org/10.1080/17470210701255317

Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In Eric Reuland & Alice ter Meulen (eds.), *The representation of (in)definiteness*, 98–129. Cambridge, MA: MIT Press.

Radó, Janina & Oliver Bott. 2012. Underspecified representations of scope ambiguity? In Maria Aloni, Vadim Kimmelmann, Floris Roelofsen, Galit W. Sassoon, Kathrin Schulz & Matthijs Westera (eds.), *Logic, language and meaning: 18th Amsterdam Colloquium* (Lecture Notes in Computer Science 7218), 180–189. Amsterdam: Springer. DOI: https://doi.org/10.1007/978-3-642-31482-7_19

Raffray, Claudine N. & Martin J. Pickering. 2010. How do people construct Logical Form during language comprehension? *Psychological Science* 21(8). 1090–1097. DOI: https://doi.org/10.1177/0956797610375446

Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. London/Sydney: Croom Helm.

Reinhart, Tanya. 1995. Interface strategies (OTS working papers in linguistics). Ms. Universiteit Utrecht.

Robaldo, Livio, Jakub Szymanik & Ben Meijering. 2014. On the identification of quantifiers' witness sets: A study of multi-quantifier sentences. *Journal of Logic, Language and Information* 23. 53–81. DOI: https://doi.org/10.1007/s10849-014-9197-9

Ruys, Eddy G. & Yoad Winter. 2010. Quantifier scope in formal linguistics. In M. Dov Gabbay & Franz Guenthner (eds.), *Handbook of philosophical logic* 16. 159–225. Dordrecht: Springer.

Sauerland, Uli. 2003. On Quantifier Raising in German. Unpublished manuscript, University of Tübingen.

Schlotterbeck, Fabian & Oliver Bott. 2013. Easy solutions for a hard problem? The computational complexity of reciprocals with quantificational antecedents. *Journal of Logic, Language and Information* 22(4). 363–390. DOI: https://doi.org/10.1007/s10849-013-9181-9

Sprouse, Jon. 2011. A test of the cognitive assumptions of magnitude estimation: Commutativity does not hold for acceptability judgments. *Language* 87. 274–288. DOI: https://doi.org/10.1353/lan.2011.0028

Steedman, Mark. 2011. *Taking scope – the natural semantics of quantifiers*. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262017077.001.0001

Syrett, Kristen, Adrian Brasoveanu & Atsushi Oho. 2016. An experimental investigation of the scope of object comparative quantifier phrases. Ms. Rutgers University.

Szabolcsi, Anna. 1997. Strategies for taking scope. In Anna Szabolcsi (ed.), *Ways of scope taking*, 109–155. Dordrecht: Kluwer Academic Publishers. DOI: https://doi.org/10.1007/978-94-011-5814-5_4

Szabolcsi, Anna. 2010. *Quantification*. Cambridge, MA: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511781681

Tunstall, Susanne L. 1998. *The interpretation of quantifiers: Semantics and processing*. Amherst, MA: University of Massachusetts dissertation.

VanLehn, Kurt A. 1978. Determing the scope of English quantifiers. *Technical Report AI-TR-483 Artificial Intelligence Laboratory*. Cambridge, MA: Massachusetts Institute of Technology.

Weskott, Thomas & Gisbert Fanselow. 2011. On the informativity of different measures of linguistic acceptability. *Language* 87. 249–273. DOI: https://doi.org/10.1353/lan.2011.0041

How to cite this article: Radó, Janina 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): 91. 1–29, DOI: https://doi.org/10.5334/gjgl.309

Submitted: 05 December 2016 **Accepted:** 18 October 2017 **Published:** 20 August 2017

]u[*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press. **OPEN ACCESS**