GLOSSA
a journal of general linguistics

Open Library of Humanities

Assessing the availability of inverse scope in German in the covered box paradigm

Gisbert Fanselow, University of Potsdam, DE, fanselow@uni-potsdam.de

Malte Zimmermann, University of Potsdam, DE, mazimmer@uni-potsdam.de

Mareike Philipp, University of Potsdam, DE, maphilip@uni-potsdam.de

This paper presents the results of a novel experimental approach to relative quantifier scope in German that elicits data in an indirect manner. Applying the covered-box method (Huang et al. 2013) to scope phenomena, we show that inverse scope is available to some extent in the free constituent order language German, thereby validating earlier findings on other syntactic configurations in German (Radó & Bott 2018) and empirical claims on other free constituent order languages (Japanese, Russian, Hindi), as well as recent corpus findings in Webelhuth (2020). Moreover, the results of the indirect covered-box experiment replicate findings from an earlier direct-query experiment with comparable target items, in which participants were asked directly about the availability of surface scope and inverse scope readings. The configuration of interest consisted of canonical transitive clauses with deaccented existential subject and universal object QPs, in which the restriction of the universal QP was controlled for by the context.

1. Inverse scope in canonical transitive sentences: context and free word order

The relative scope of quantifiers and operators is a key phenomenon at the syntax-semantics interface. The first proposal that the syntactic relation of c-command[1] plays a crucial role in identifying the scope of operators and quantifiers can be found in Klima (1964).[2] The idea that the scope of a quantifier is identical to its c-command domain became somewhat standard with the contributions of Reinhart (1976), May (1977), Chomsky (1981) and Hornstein (1984; 1995), although more flexible models have since been formulated as well (May 1985).

The strict c-command model must be reconciled with the fact that sentences such as (1) exhibit scope-ambiguity in English. In addition to the preferred interpretation (1a), in which syntactic c-command and semantic scope are in harmony for the two quantificational expressions, there is the (often dispreferred) option of reading the sentences as in (1b), i.e. with inverse scope: the universal object QP takes semantic wide scope although the existential object QP c-commands it at surface structure.

(1)   *Some student* admires *every professor*.
   a. $\exists x \, (student'(x) \; \& \; \forall y(professor'(y) \rightarrow admire(x,y)))$
   b. $\forall y(professor'(y) \rightarrow \exists x \, (student'(x) \; \& \; admire(x,y)))$

If scope is indeed linked to syntactic c-command, it is necessary to postulate an optional covert operation ("Quantifier Raising", QR) that rearranges the two quantified expressions *some student* and *every professor* in such a way that the latter comes to c-command the former at some inaudible level of representation, Logical Form (LF): (2) would be a simple way of representing such a Logical Form.

(2)   every professor$_i$ (some student admires t$_i$)

This two-component model postulates a strict condition for scope in interaction with a covert movement process. As a result, it potentially faces an overgeneration problem: scope ambiguities are predicted in many syntactic contexts in which they in fact do not arise. Double object constructions are a famous example from English, see (3) (Larson 1990).[3] In (3b), the inverse reading is unavailable, contrary to the prediction from the model sketched above.

(3)   a.   The teacher assigned one problem to every student.   ✓∃∀   ✓∀∃
       b.   The teacher assigned one student every problem.   ✓∃∀   *∀∃

(Larson 1990: 604)

---

[1] A node a c-commands a node b iff the first (branching) node above a also dominates b and b is not part of a.

[2] Klima (1964) used the term "in construction with" for c-command.

[3] See also Beghelli & Stowell (1997) for the frequent absence of scope ambiguities in English, depending on quantifier type, whereas Bruening (2019) presents experimental evidence for the view that English double object construction are scope-ambiguous after all.

More generally, fixed word order SVO languages from East and Southeast Asia seem not to behave as expected: canonical active SVO sentences in Mandarin Chinese (Aoun & Li 1993) and Vietnamese (Thuan Tran, p.c.) come only with the surface scope interpretation; see also Scontras et al. (2017) for recent experimental evidence to this effect in Mandarin, and Larson & Wu (2018) for an account of such observations in terms of the syntactic representation of topicality.

The present paper is concerned with a different situation that has been claimed to involve scope rigidity. It focuses on the existence of scope ambiguities (or lack thereof) in canonical transitive sentences in languages with free constituent order (+scrambling). Some free constituent order languages such as Hungarian have a special clausal domain reserved for hosting quantifiers. In this structural domain, scope is rigidly connected to surface c-command (É. Kiss 1991; Szabolcsi 1997). For free word order languages lacking such a quantifier field, a recurrent claim is that transitive sentences with canonical order (S > O) are scope-rigid as well. As we shall see, such claims require qualification in a quite general way.

Ionin (2001) characterizes Russian (free constituent order, canonical SVO order) as exhibiting only surface scope in SVO sentences. By contrast, Antonyuk (2015) sketches a more nuanced picture: most of her informants accept inverse scope as a marked option for SVO order. Inverse scope is quite easily available for (4a) and (4b), taken from Antonyuk (2015: 26–27).

(4) a. Dva studenta   pročitali   každuju knigu.   [*Russian*]
         Two students(NOM) read(PST.PL) every   book(ACC)
         'Two students read every book'

    b. Neskol'ko xirurgov operirovali každogo pacienta.
         Several   surgeons operated   every   patient
         'Several surgeons operated on every patient'

Interestingly, the situations that Antonyuk specifies as capable of licensing inverse scope are those in which there is a contextually given set of entities quantified over by the direct object (e.g., the set of books from a reading list of a seminar, or the patients that need to undergo surgery in a certain clinic on a certain day). In subsequent work, however, Antonyuk (2019) adopts an even more general perspective on scope in Russian. She provides evidence that Quantifier Raising, and consequently inverse readings in Russian, are in fact available under the very same conditions as in English, despite scrambling being an option in Russian but not in English.

Mahajan (2018) discusses scope in the free constituent order and SOV language Hindi. He observes a certain scope rigidity for sentences with unmarked SOV order. Scope seems confined to surface c-command, but Mahajan is careful to note that some speakers accept inverse scope for unscrambled SOV order, too. Guha, Moitra & Marty (2021) present the results of a truth value judgment test which shows that canonical SOV sentences also allow for scope ambiguities in Bangla. Early studies of Japanese (also SOV and +scrambling) argued for a restriction of normal

order SOV sentences to surface scope (Kuroda 1969/1970), but later studies (e.g., Kitagawa 1990) have contested this view. Hayashishita (2013) and, in particular, Ueyama & Hayashishita (2020) show that inverse scope in SOV sentences is possible if there is exactly one group of objects in the context that can form the extension of the object – in line with the context characterization in Antonyuk (2015) for Russian.

The experiment reported on in the present paper is concerned with German, a language that also fits the characterization of a basic SOV order language with scrambling. Much of the work in generative syntax assumes that sentences with canonical S > O order disallow inverse scope (e.g., Bobaljik & Wurmbrand 2012, but see Beck & Gergel 2014 for an opposing view), and typically cites the detailed introspection-based investigation of Frey (1993) as key evidence. Notice, though, that Frey confines his claims to a very particular syntactic configuration, namely to clauses with verum focus intonation and two modified quantifiers *mindestens ein* 'at least one' and *fast jeder* 'almost everyone', in order to avoid referential readings. Under different prosodic conditions, by contrast, inverse scope may be possible even with canonical SOV orders (Krifka 1998). Drawing on the results of behavioral experiments, Radó & Bott (2018) establish that German main clauses with SVO order, in fact, allow for scope ambiguities, as well. Their findings are in line with recent corpus analyses showing that inverse scope indeed also occurs with SOV order in attested production data (Webelhuth 2020).

Webelhuth (2020) observes that these examples with SOV order and inverse scope interpretation all share the property that the set quantified over by the object is discourse-old, or easily inferable. In that respect, his findings are comparable to the facts observed for Japanese or Russian. The experiment reported below attempts to establish the availability of inverse scope in German SOV sentences in the kind of context described by Ueyama & Hayashishita (2020) and Webelhuth (2020): the set of entities quantified over (or referred to) by the objects is given in the preceding context. We employ a novel method that, we hope, is particularly suitable for dealing with ambiguity, viz. the covered box paradigm (Huang, Spelke & Snedeker 2013). The experiment builds on and extends earlier studies carried out in the context of our SFB (CRC) research projects, viz. Winkel (2018) and Blum (2020), and Philipp & Zimmermann (submitted).

2. A covered box experiment on quantifier scope

We designed our experiment with the aim of confirming prior experimental and corpus-based evidence to the effect that inverse scope is also possible in German sentences with quantified subjects and objects in canonical SO order. More specifically, in contexts where the set quantified over is discourse-given. The covered-box paradigm we adopted is a more indirect method for gauging acceptance of inverse readings for sequences of two sentences. The first of these sentences

introduces a set of entities about which the second adds a follow-up statement involving two quantifiers, one in subject and one in object position.

Szendrői (2022) discusses a problem with standard direct-query experiments on quantifier scope, which directly probe for truth-value judgments: adult speakers of English typically have no problems in accepting the critical inverse readings in truth-value judgment tasks, as long as the inverse scope reading is associated with the *yes*-answer. The high acceptability of inverse readings stands in contrast to the fact that these readings are dispreferred even in English (Anderson 2004; Reinhart 2006). This should, if anything, make them more difficult to detect. Szendrői (2022) attributes the unexpectedly high acceptability of inverse scope readings to the pragmatic *Principle of Charity* (Wilson 1959; Davidson 1984: ch.13), following Crain & Thornton (1998: 211f.), who call upon this principle in their studies on language acquisition. These authors observe that their child participants tend to respond *yes* in truth-value judgment tasks on potentially ambiguous sentences when the reading associated with the *yes*-answer is available to them – at least in the absence of salient alternatives: "This bias to say "Yes" is apparently enough to boost the dispreferred interpretation of an ambiguous sentence" (ibid.). While this phenomenon may be a problem for studies which have the goal of detecting which reading is the *preferred* one, it is in fact an advantage for studies which aim to detect whether a potentially existent but dispreferred reading is available at all. This holds even more so for languages such as German, in which the inverse reading is generally much less available than in languages like English. This means that it should be more difficult to detect whether a reading is just dispreferred or, in fact, completely absent. In our experiment, we take advantage of this effect by employing some of the factors Szendrői (2022) holds responsible for activating the *Principle of Charity*: participants judge the truth or falsity of a sentence with respect to a single context, and the *yes*-answer is associated with the inverse reading.

At the same time, however, we employ a more indirect task than the direct-query method: a covered-box paradigm. The main reasons for this are methodological concerns about testing potentially ambiguous sentences, as prominently voiced in Matthewson (2004). Matthewson was the first to point out that non-linguist speakers cannot give a meta-judgment as to whether a sentence is ambiguous or not, since this would require the speakers to reflect, rather than to give an intuitive judgment. In addition, non-linguist speakers show a tendency to associate every sentence with just a single meaning, which drives them towards disambiguation (typically, the preferred reading) when both readings are presented side-by-side. In the covered-box paradigm we avoid these problems by presenting participants with two options to choose from. However, one of them is left implicit (covered), such that the participant cannot directly compare the two potential readings. In addition, the visible (uncovered) alternative exhibits variability in the form of different types of statements and different ways of phrasing, so the participant is also less

likely to detect and consequently compare the two crucial readings for the target sentences (see section 2.1 below for details on the study design). The covered-box paradigm thus allows us to take advantage of the *Principle of Charity* to detect dispreferred readings, while at the same time avoiding the possibility that participants will directly compare the two relevant readings within or across items. In sum, the aim of the study is to gauge whether the covered-box experimental paradigm can successfully be applied to the investigation of inverse scope in German.

2.1 Experiment design

2.1.1 Materials

All items consisted of a sequence of context sentence, target sentence and follow-up statement. The follow-up statement presented a potential paraphrase of the target sentence, which participants judged for truth. There were three types of possible follow-up statements which differed as to whether they claimed that there was (i) a single X, or (ii) more than one X, or (iii) no X involved in the event described by the context and target sentence. Each participant saw only one of these three statements per item. In total, we had 18 target items and 48 filler items, which were divided up into 6 lists in a Latin Square design. That is, each participant saw six target items with paraphrasing statements of type (i), (ii), and (iii), respectively. The same was followed with the filler items. No participant saw any item twice.

   Target Items: The target items consisted of a doubly quantified sentence with an existential subject QP and a universal direct object QP, as illustrated in (5) below. The existential QP was headed by the reduced form *'n/'ne* of the indefinite article, and the universal QP featured the distributive quantifier *jede/jeder* 'every, each'. The reduced form of the indefinite was chosen to control for potentially confounding effects of prosody. Since the reduced form cannot carry stress, two confounding interpretations are excluded: (i) an interpretation as a specific (numeral) indefinite NP (Ebert & Hinterwimmer 2010), which takes obligatory wide (surface) scope (Ebert 2009); (ii) an interpretation as a contrastive topic under the rise-fall-contour (Krifka 1998), which would confound towards the inverse reading. The nouns of both QPs were introduced in the context sentence as part of an embedded clause under a future-directed attitude verb (*order, hope, promise*, etc…). This way, the QP-content had the same information structural status of *given*. Importantly, the antecedent of the critical existential QP was introduced in number neutral form, by using PPs headed by *per* 'via, by' and containing a bare singular NP, which are underspecified for semantic number in German. This was meant to ensure that participants would not be biased towards the *single X* or the *more than one X* reading merely because of shallow number matching with the context NP. A potential bias could still be triggered by the morphological form, but since the form is singular, it would only result in the generation of false negatives (= surface readings), but not false positives.

(5)   Der Agrarexperte   hatte empfohlen,   dass die Felder per Regenkanal bewässert
           the agriculture.expert had   recommended that the fields via open.canal irrigated
           werden sollten, und tatsächlich hat dann auch **'n Regenkanal jedes Feld**
           be   should, and in.fact   has then also a open.canal   every field
           bewässert.
           irrigated.
           'The agriculture expert had recommended that the fields should be irrigated through open canals, and then, in fact, an open canal irrigated every field.'

Each of the three follow-up statements illustrated in (6) came in one of two versions with the same meaning, but with slightly different forms. The different versions were created to avoid training effects. Otherwise, participants might have gotten used to the sentence structures and started to superficially match a certain sentence type to a certain response purely out of habit throughout the course of the experiment. In the target items, statement 1 ($n = 1$) corresponded to the surface reading and statement 2 ($n > 1$) to the inverse reading, while statement 3 ($n = 0$) was unambiguously false. The contexts and linguistic materials were chosen in such a way that both readings were plausible in terms of world knowledge. The experimental materials were constructed by modifying material from earlier studies (Philipp & Zimmermann, submitted).

(6)   i.   Evaluated Statement 1: ($n = 1$)
        a.   Version 1: Alle Felder wurden von demselben Regenkanal bewässert.
                    'All fields were irrigated by the same open canal.'
        b.   Version 2: Die Felder wurden insgesamt nur von einem einzigen Regenkanal
                    bewässert.
                    'Overall, the fields were irrigated by only a single open canal.'

    ii.   Evaluated Statement 2: ($n > 1$)
        c.   Version 1: Insgesamt gab es mehr als einen Regenkanal, der die Felder bewässert
                    hat.
                    'Overall, there was more than one open canal that irrigated the fields.'
        d.   Version 2: Insgesamt gab es mehrere Regenkanäle, die die Felder bewässert haben.
                    'Overall, there were several open canals that irrigated the fields.'

    iii. Evaluated Statement 3: ($n = 0$)
        e.   Version 1: Die Felder wurden von keinem einzigen Regenkanal bewässert.
                    'The fields were not irrigated by a single open canal.'
        f.   Version 2: Kein Feld wurde von einem Regenkanal bewässert.
                    'No field was irrigated by a open canal.'

Fillers: There were five types of fillers, examples of which are shown below. As in the targets, we varied the wording of the follow-up statements slightly to avoid training effects.

(F1) **Basic Filler Type (9 items):** The basic condition only involved a singular existential without a universal quantifier. Therefore, it was unambiguous, with only statement 1 being true.

(7) Die Mathestudenten hatten sich geeinigt, 'nen Film im Campus-Kino zu
the maths.students had themselves agreed a film at.the campus.cinema to
gucken, und tatsächlich haben sie dann auch **'nen Film** geguckt.
watch, and in.fact have they then also a film watched
'The maths students had agreed to watch a film at the campus cinema, and then, in fact, they did watch a film.'

    i. Statement 1: (n = 1)
        Es wurde insgesamt nur ein Film von den Mathestudenten geguckt.
        'Overall, the maths students watched only one film.'

    ii. Statement 2: (n > 1)
        Es wurde insgesamt mehr als ein Film von den Mathestudenten geguckt.
        'Overall, the maths students watched more than one film.'

    iii. Statement 3: (n = 0)
        Es wurde kein Film von den Mathestudenten geguckt.
        'No film was watched by the maths students.'

(F2) **No X-Fillers (15 items):** This filler type was included to ensure that statement 3 was not always false. This condition stated that the event happened for no X. Therefore, only statement 3 should be judged as true.

(8) Die Kunstauktionatoren hatten behauptet, später 'nen echten Dürer zu versteigern,
the art.auctioneers had claimed later a real Durer to auction
und tatsächlich haben sie dann aber **keinen echten Dürer** versteigert.
and in.fact have they then but no real Durer auctioned
'The art auctioneers had claimed they would auction off a real Durer later, but then, in fact, they did not auction off a real Durer.'

    i. Statement 1: (n = 1)
        Es wurde insgesamt genau ein echter Dürer von den Kunstauktionatoren versteigert.
        'Overall, exactly one real Durer was auctioned off by the art auctioneers.'

    ii. Statement 2: (n > 1)
        Es wurde insgesamt mehr als ein echter Dürer von den Kunstauktionatoren versteigert.
        'Overall, more than one real Durer was auctioned off by the art auctioneers.'

    iii. Statement 3: (n = 0)
        Es wurde kein echter Dürer von den Kunstauktionatoren versteigert.
        'No real Durer was auctioned off by the art auctioneers.'

(F3) **The Each-Filler Type (6 items):** This condition contained the strongly distributive word *jeweils* (= each) plus a reduced existential QP distributing over a plural antecedent (Zimmermann 2002). The interpretation of such distance-distributive sentences should strongly favour statement 2.

(9)   Die Büroangestellten hatten darauf gesetzt, beim Umzug 'nen höhenverstellbaren
the office.workers had onto bet at.the move a height-adjustable
Tisch zu ergattern, und tatsächlich haben sie dann auch **jeweils** 'nen
desk to get.hold.of and in.fact have they then also each a
höhenverstellbaren Tisch ergattert.
height-adjustable table got.hold.of
'The office workers had set their hopes on getting hold of a height-adjustable desk
during the move, and then, in fact, they each got hold of a height-adjustable desk.'

    i.   Statement 1: ($n = 1$)
Die Büroangestellten haben zusammengenommen genau einen höhenverstellbaren
Tisch ergattert.
'Altogether, the office workers got hold of exactly one height-adjustable desk.'

    ii.   Statement 2: ($n > 1$)
Die Büroangestellten haben zusammengenommen mehr als einen
höhenverstellbaren Tisch ergattert.
'Altogether, the office workers got hold of more than one height-adjustable desk.'

    iii.   Statement 3: ($n = 0$)
Die Büroangestellten haben zusammengenommen keinen einzigen
höhenverstellbaren Tisch ergattert.
'Altogether, the office workers didn't get hold of a single height-adjustable desk.'

(F4) **Universal Filler Type (6 items):** This condition consisted of doubly quantified sentences like the targets, but the order of quantifiers was reversed, with the universal QP in canonical subject position and the existential QP in canonical object position. This sentence is ambiguous or vague, but it should favour a distributive reading in the presence of the reduced indefinite article, and hence statement 2.

(10)   Der Gärtner hatte vorgeschlagen, dass die Bäume von Freiwilligen beschnitten
the gardener had suggested that the trees by volunteers pruned
werden sollen, und tatsächlich hat dann auch **jeder Freiwillige 'nen Baum**
be should and in.fact has then also every volunteer a tree
beschnitten.
pruned
'The gardener had suggested that the trees should be pruned by volunteers, and then, in fact, every volunteer pruned a tree.'

i. Statement 1: (n = 1)
    Insgesamt haben alle Freiwilligen nur einen einzigen Baum beschnitten.
    'Overall, all the volunteers pruned only a single tree.'

ii. Statement 2: (n > 1)
    Insgesamt gab es mehrere Freiwillige, die einen Baum beschnitten haben.
    'Overall, there were several volunteers who pruned a tree.'

iii. Statement 3: (n = 0)
     Insgesamt gab es keine Freiwilligen, die einen Baum beschnitten haben.
     'Overall, there were no volunteers who pruned a tree.'

(F5) **RC-ambiguity (12 items):** This condition contained a different kind of ambiguity, namely a structural ambiguity between high-attachment vs. low-attachment in restrictive relative clauses. In addition to being a distractor, this condition was added to establish whether participants have a tendency of rejecting readings that are available but dispreferred. In the case of RC-ambiguities, the high-attachment construal is generally considered more costly in processing (Hemforth et al. 2015), and hence dispreferred, just as inverse scope readings are. The logic is as follows: should the acceptance of inverse reading turn out to be very low, it would be unclear whether this is due to inverse readings being unavailable or whether participants simply reject those readings because they are dispreferred (though available). In the case of the latter, we should see the same rejection of dispreferred, but available readings with instances of RC-ambiguity. These are known to have two readings, with one of them being dispreferred (the high-attachment reading). In contrast, should the acceptance of high-attachment RC-readings be high and the acceptance of inverse readings very low, this would constitute evidence that the low acceptability of inverse readings is not just due to a dispreference. These items hence serve as controls for whether our method is suitable for measuring availability, rather than preference. In (F5), statement 1 paraphrases the high-attachment reading and statement 2 the low-attachment reading.

(11) Der Richter hatte gehofft, dass die heutige   Verhandlung im   Gruppenprozess
      the judge   had   hoped   that the today.ADJ hearing   in.the group.trial
      neue Aussagen   bringen würde, und tatsächlich hat dann auch **'n Angehöriger** von
      new testimonies bring   would and in.fact   has then also a relative   of
      einem Angeklagten, **der** bisher geschwiegen   hatte, 'ne Aussage   gemacht.
      a   defendant   who so.far remained.silent has   a   testimony made
      'The judge had hoped that today's hearing in the group trial would bring about new testimonies, and then, in fact, a relative of a defendant who so far had remained silent gave a testimony.'

    i.   Statement 1: (high attachment to NP1)

        Es gibt einen Angehörigen, der bisher geschwiegen hat.

        'There is a/one relative who so far has remained silent.'

   ii.   Statement 2: (low attachment to NP2)

        Es gibt einen Angeklagten, der bisher geschwiegen hat.

        'There is a/one defendant who so far has remained silent.'

  iii.   Statement 3: ($n = 0$)

        Es gibt keinen, der bisher geschwiegen hat.

        'There is no one who so far has remained silent.'

2.1.2 Participants

We tested 43 participants between 18 and 50 years of age, with a mean age of 29 years. 24 participants were female and 19 were male. The participants were recruited via the online platform Prolific. They received a reimbursement of 6.60£ for their participation. All of them stated that they were monolingual speakers of German. We excluded one participant from the analysis because they did not reach the accuracy threshold of 10 out of 12 in the control items (the twelve control items were the filler items F1-F4 in Condition Statement 3, which were unambiguously true in F2 and unambiguously false in F1/F3/F4).

2.1.3 Method & procedure

The study was conducted online with the free software OnExp. We employed a forced-choice truth-value-judgment task in the covered box paradigm (Huang et al. 2013). Participants were instructed that they would always see a sentence and two statements about this sentence, one being true and the other one false. However, only one of the statements was visible, while the other one was covered up. Participants had to choose the statement that they considered true. Since one of the two statements was covered up with no option of uncovering, participants should have only opted for the covered option if they were certain that the visible sentence was false. In this way, we hoped participants would opt for visible statements, expressing dispreferred scope options, given that the other statement (here: the covered one) must be false. Experimental items and statements were pooled in such a way that each of statements 1 to 3 was the visible one in one third of the items for each participant. This way, each participant saw six target items with the visible statement 1, another six with visible statement 2, and a final six with visible statement 3, distributed across lists. Participants gave informed consent about their participation and the use of their data. They were informed beforehand about data protection issues and that they were free to abort the experiment at any time, thereby complying with standard ethics requirements on linguistic experiments.

2.2 Predictions

Across participants, all target and filler sentences were followed by three different kinds of visible statements. For the target sentences, statement 1 probed for surface scope readings, statement 2 probed for inverse readings, whereas statement 3 was a control and unambiguously false. We predict that visible statement 3 should never be chosen for the target sentences. The predictions for the two scope readings are as follows: As visible statement 1 probes for surface scope, it should be easily accepted across items and participants. By contrast, if inverse scope readings are available, at least to some extent, visible statement 2 should be accepted to some degree that is clearly above the (non-)acceptance of control statement 3. Moreover, since inverse scope is generally thought to be more difficult to obtain, we expected the acceptance rate of statement 2 to be lower than that of statement 1 (though, again, clearly higher than statement 3).

Most filler conditions also allowed for clear predictions, as discussed in the *Materials* section above. Possible exceptions are Condition Statement 1 in F4 (Universal Filler), where we expected Statement 1 to be a dispreferred, but still a possible paraphrase of the filler item (marked by (✓)). For the structural ambiguity condition F5 (RC-ambiguity), we expected both statement 1 and 2 to be available, possibly with a higher acceptance of the Low Attachment-option Statement 2, which has been shown to be preferred in processing (Hemforth et al. 2015).

The predictions for the three statements for targets and filler items are summarised in **Table 1**.

		Statement 1 (n = 1/NP1)	Statement 2 (n > 1/NP2)	Statement 3 (n = 0)
Targets		✓	✓ (< statement 1)	*
Fillers	**Basic**	✓	*	*
	Kein	*	*	✓
	Each	*	✓	*
	Universal	(✓)	✓	*
	RC-ambiguity	✓	✓	*

Table 1: Predicted acceptance of the visible statement as the correct paraphrase for target items on different scope construals and for fillers.

2.3 Results

The descriptive results for targets and fillers are shown in **Table 2**. The table shows the proportion of cases for which the respective statement was judged as true when it was presented as the visible paraphrase.

		Statement 1 (n = 1/NP1)	Statement 2 (n > 1/NP2)	Statement 3 (n = 0)
Targets		0.71	0.34	0.02
Fillers	**F1 Basic**	0.87	0.06	0.02
	F2 Kein	0.00	0.00	0.97
	F3 Each	0.08	0.90	0.00
	F4 Universal	0.12	0.74	0.00
	F5 RC-ambiguity	0.77	0.71	0.01

Table 2: Proportion of cases when the respective condition was judged as true out of all the cases when this condition was the visible one.

For the scopally ambiguous target sentences, the visible follow-up statement was only accepted in 2% of all cases when it was the false statement 3, which is as predicted. Statement 2, which corresponds to the inverse reading, was judged as true in 34% of cases. Statement 1, which corresponds to the surface reading, was judged as true in 71% of the cases. Applying a binom-test, it can be shown that the p-values for surface scope (71%) and inverse scope (34%) are both <0.01, i.e., these results differ significantly from chance level.

The fillers, which also served as controls, show the predicted results. This is a very good indicator that participants understood the procedure, and that the experiment delivered reliable results. Statement 3, which was always clearly true or false, received values at, or very close to bottom or ceiling. The acceptance rate was at 0-2% when it was false (F1, F3 to F5) and at 97% when it was true (F2). As for statements 1 and 2, the false statement in the unambiguous conditions F1 *basic* (2) and F2 *kein* (1) was only accepted in 0–6% of all cases, whereas the true statement was accepted in 87–97% of all cases. In conditions F3 *each* and F4 *universal*, we see a similar tendency, with only 8–12% acceptance rate for the false or dispreferred statements as opposed to 74–90% acceptance rate for the true or preferred statements. In condition F5 *RC-ambiguity*, both readings with high- (1) and low-attachment (2) were accepted to a relatively high degree, with a slightly higher value for high-attachment (77%) than low-attachment (71%).[4]

---

[4] This result is surprising in that the acceptance rate for both dispreferred options (inverse scope and high attachment, respectively) is high enough to suggest that these readings are available. At the same time, though, the values for the two RC-conditions are similar, i.e., the difference in preference does not show at all in this testing method. The situation is different for scope: even though the inverse reading is accepted in one third of all cases, this is still much lower than the acceptance rate for the surface reading, i.e., the difference in preference still shows. The different patterns could follow from differences in the linguistic structures involved, which may be connected to differences in processing. Alternatively, the difference in preference in the RC case is simply much smaller than in the scope case. Our method of testing for availability would then make a small difference in preference completely disappear in the acceptability rates, whereas a larger difference in the scope case would still show.

In addition to these overall means, we also observed a high degree of variability between speakers. Looking at the acceptance rates of the visible statement for individual participants in **Figure 1**, we observe a wide range of response patterns: some participants never accept the inverse reading, whereas others accept the inverse reading for all 6 items. However, the majority of participants (88%) accepted the inverse reading in 0 to 3 cases out of a total of 6.
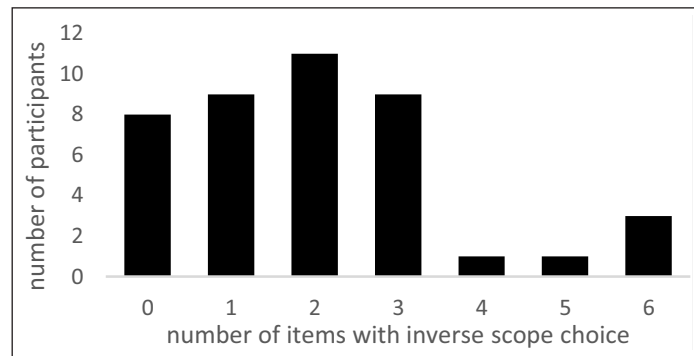
Figure 1: Number of items with inverse scope choice out of 6 items in total (x-axis) and number of participants who accepted the inverse reading that many times (y-axis).

Looking next at the individual items, we also find some variability. This variability is presumably triggered by differences in the plausibility of the individual surface scope and inverse scope scenarios for each item. Crucially, the numbers are not driven by a few outliers, since 94% are in the range of 0–60%. This is shown in **Figure 2**.
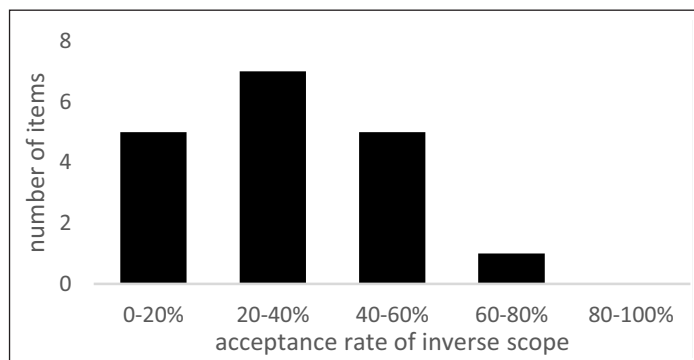
Figure 2: Acceptance rate of inverse scope (x-axis) and number of items for which the inverse reading was accepted to that rate (y-axis).

3. Discussion

In this section, we first sum up the main results in section 3.1. Section 3.2 compares the results to findings from a previous experimental study on otherwise fully comparable material, which

employed a more standard direct-query method for sentences with potential scope ambiguities. This section also discusses some open empirical questions. Finally, section 3.3 presents a short discussion of how our findings may be integrated into existing theoretical approaches to relative quantifier scope from a cross-linguistic perspective.

3.1 Main results

There are two main results. First, **Table 2** above shows the expected outcome for all the control items. This makes us confident that the experiment is ecologically valid: participants were able to, and did in fact, comply with our instructions in the intended manner. As for our main empirical question, **Table 2** also shows that the experimental participants accepted an inverse scope reading for sentences with existential subject QP and universal object QP in canonical word order in roughly a third of all cases. Given that this acceptance rate is way above the rejection threshold of 0–2% for the clearly false control sentences, we can therefore conclude that inverse scope readings for canonical SOV sentences are available to some degree, also in German. We will return to the question of how this result fits with theoretical approaches to relative quantifier scope in free constituent order languages in section 3.3.

3.2 Comparison to a previous experiment and some open questions

The experimental results are by-and-large comparable to the results of an earlier experimental study. This study elicited scope data by means of a forced choice task that directly probed for the availability of a given scope reading (Philipp & Zimmermann, submitted). As in the present experiment, participants were presented with future-oriented contextual scenarios introducing the NP-contents of existential and universal QPs, thereby making them given and deaccented. As in the experiment here, the context sentence was followed up by the target sentence describing what actually happened, which was in turn followed up by a direct polar content question probing for surface scope (Q-SR) or inverse scope (Q-IR), respectively. The forced choice task consisted of participants having to decide whether or not to answer the polar question in the affirmative. An affirmative answer was taken to indicate that the questioned scope reading was available; see Philipp & Zimmermann (submitted) for details and more discussion. (12) shows a full target item with Q-SR and Q-IR (in English translation), respectively:

(12)   Context: Der Polizeibeamte hatte vermutet, dass **die Einbrecher** von neu angebrachten Überwachungskameras aufgenommen worden sein könnten, und tatsächlich…
'The police officer hoped that the burglars might be recorded by newly installed surveillance cameras, and then, in fact…'

a. Target … hat dann **['ne neu angebrachte Überwachungskamera] [jeden**
has then a newly installed surveillance.camera every
Einbrecher] aufgenommen.
burglar recorded
'… a newly installed surveillance camera recorded every burglar.'

b. Can this sentence be understood to mean that, overall, …

i. Q-SR … **only a single** newly installed surveillance camera recorded the
burglars? *yes/no*

ii. Q-IR … **more than one** newly installed surveillance camera recorded the
burglars? *yes/no*

As shown in (12), the target items in this experiment matched the targets from the present experiment in the crucial context condition: the restriction of the universal QP (here *Einbrecher* 'burglar') in the target clause is explicitly introduced by a definite plural DP (*die Einbrecher* 'the burglars') in the context. The only difference in design concerned the number marking on the antecedent of the existential QP, which carries plural marking in (12). This raised the risk of inducing a bias for falsely answering Q-IR (probing for inverse scope) in the affirmative, on the basis of a shallow number matching effect. However, participants' responses in control conditions showed no evidence for such morphological number matching whatsoever; see Philipp & Zimmermann (submitted) for discussion. Recall that this potentially biasing factor for IR-responses was avoided in the present experiment by using a bare singular NP as the antecedent of the existential QP. Finally, the direct-query experiment also used three of the filler conditions from the covered-box experiment, namely F1 (Basic), F3 (Each), and F4 (Universal). The results for target sentences and fillers are shown in **Table 3**, together with the comparable results of the covered-box experiment from **Table 2** (in brackets).

		Q-SR	Q-IR
Targets		0.82 (0.71)	0.39 (0.34)
Fillers	**F1 Basic**	0.94 (0.87)	0.10 (0.06)
	F3 Each	0.08 (0.08)	0.89 (0.90)
	F4 Universal	0.12 (0.12)	0.95 (0.74)

Table 3: Proportion of affirmative responses to Q-SR and Q-IR for target items and fillers in forced-choice experiment. In brackets, proportion of acceptance of visible statement 2 of the covered-box experiment from Table 2 (repeated).

The results of the two experiments are comparable overall, and in some conditions even identical, thereby validating both as viable experimental methods for the investigation of relative quantifier scope. In particular, the direct-query task also gave clear results in the three control

conditions, providing solid evidence for the existence of surface scope (0.82) and inverse scope reading (0.39), the latter as a dispreferred option. Moreover, proportions of affirmative responses in the direct-query method were overall slightly higher across conditions. We suspect that the slightly higher acceptability in the direct-query method may be an effect of task and a difference in instructions: in the direct-query method, participants were explicitly instructed to give a *yes*-response whenever a reading is *possible,* even if they can also think of another interpretation. In the covered-box paradigm, participants were *forced to choose* between a visible and a covered option and they might sometimes suspect that the covered option is the better one. Moreover, there is also a pronounced difference in the results of the covered-box method when compared to a classical forced-choice procedure in which participants must actively choose between the two simultaneously displayed scope readings, see Anderson (2004). Such experimental settings allow for participants to always opt for the preferred (surface) scope reading, thereby, in the worst case, making the dispreferred (inverse) reading largely invisible. Taken together, our findings show that the covered-box method works as predicted and is, in fact, superior to classical forced-choice methods for detecting dispreferred readings. It thus seems likely that participants applied the *Principle of Charity* in both the direct-query and the covered-box method, but to a higher degree in the former than the latter. Furthermore, we can observe that both conditions probing for the availability of surface scope, the default option, exhibit a somewhat unexpected non-ceiling effect with only 71% acceptance rate in covered-box and 82% *yes*-responses in direct-query. One possible way of accounting for such non-ceiling effects would be to relate them to processing considerations. It appears plausible that, when confronted with a potentially ambiguous clause, participants will opt for the first of the two readings that comes to mind. Depending on world knowledge and subjective experience, it is quite likely that some participants may first come up with a situation verifying the inverse scope reading, and then subsequently stick to this construal: rejecting visible statement 1 or answering Q-SR with *no*, respectively (see, e.g., Sanford & Garrod (1998) for an account of processing-driven, subpropositional derivation of meanings, and Scontras et al. (2017) and Ionin (2014), for comparable non-ceiling effects with surface scope readings).

Turning back to the covered-box experiment, at least two questions remain open. The first is how to account for the relatively high inter-speaker variability shown in **Figure 1**. Is this an effect of different answer strategies adopted by the individual participants? Or are we dealing with dialectal or idiolectal differences concerning scope-taking? For English, the literature offers repeated informal observations of such differences between surface scopers, inverse scopers, and flexible scopers, beginning with early work by Carden (1970); see Brasoveanu & Dotlačil (2015:10) for extended discussion and references. In a forced-choice experiment on relative quantifier scope in English, Brasoveanu & Dotlačil (2015) observe parallel sub-group formation, with 21.2% surface scope takers, 9.4% inverse scope takers, and 69.4% of all participants

showing a variable behaviour. Comparing this to our results for German (elicited with a different methodology), we find almost identical percentages, with 19% consistent surface scopers, 7% inverse scopers, and again the largest subgroup of 74% showing a variable behaviour. This cross-linguistic parallel is especially interesting in view of the traditional view that inverse scope is generally easier to obtain in English than in German. The inter-speaker variability results do not really support this view, while leaving open the possibility that the observed variability is due to non-linguistic factors such as idiosyncratic response strategies.

Finally, we are left with the question of why participants accepted Statement 1 (High-Attachment) and Statement 2 (Low-Attachment) in F5 (RC-Ambiguity) to similar degrees, and even slightly more so for Statement 1 (High Attachment), the dispreferred option according to the literature. We can only offer speculations, but it seems possible that there is a qualitative difference between the resolution of structural attachment ambiguities, on the one hand, and the resolution of scope ambiguities, on the other. In particular, it is necessary to make a choice on attachment height in order to be able to derive truth-conditions for clauses with structural ambiguities. In contrast, with doubly quantified sentences it is possible to derive a consistent set of truth-conditions on the semantic surface scope construal without making a (sub-) conscious choice. In total, then, it seems that both attachment options in RC-ambiguity are salient enough to serve as an acceptable parse for the majority of speakers.

3.3 Integrating the results into the theoretical landscape

The main result of our covered-box experiment is that inverse scope is acceptable for German sentences with normal S > O word order in contexts where a preceding sentence introduces a set of objects, that, in turn, serves as the restriction for the universal distributive quantifier in the target clause. This finding is in line with the corpus evidence in Webelhuth (2020) and it replicates our own previous experimental findings in a novel experimental design. This experiment is also in line with recent claims on other free constituent order languages, which were cited above. We therefore tentatively conclude that free constituent order languages are, in fact, not as rigid with respect to scope as is often assumed. How could this be captured in a model of the syntax-semantics interface? We believe that an answer can be found in a combination of proposals made by Müller (1999; 2001) and Bobaljik &Wurmbrand (2012).

Müller (2001) has provided an optimality-theoretic account of various *order preservation* effects in natural language syntax in terms of the violable constraint PARMOV (parallel movement). PARMOV requires that c-command relations between A and B at some level L in the derivation be preserved at the next level. Müller (2001) observes that PARMOV implies scope rigidity for a free constituent order language such as German if (as he assumes), the PARMOV constraint compares the surface representation of a sentence with its LF when it comes to quantifier scope. The model of Bobaljik &Wurmbrand (2012) is very similar: they postulate the Scope Transparency Principle

ScoT (13) that demands that the precedence, or rather, c-command relations at LF encoding scope should be mirrored in the surface PF representation of a sentence.

(13)   Scope Transparency (ScoT)
        If the order of two elements at LF is A»B, the order at PF is A»B.

(Bobaljik & Wurmbrand 2012: 373)

When a language allows free constituent order, ScoT implies that LF-scope relations be expressed by surface structure c-command in free constituent order languages, just as PARMOV does.[5]

However, this conclusion is only warranted to the extent that (a) scrambling can apply freely (so that the surface structure constellations mirroring scope can be generated with no additional cost), and that (b) there are no further constraints on surface structures in free constituent order languages that may require a serialization that is in conflict with the quantifier scope relations. Closer scrutiny shows, in fact, that PARMOV is ranked *below* the constraints that Müller (1999) postulates in his model for the principles governing free constituent order in German. Consequently, surface c-command *can* be in conflict with LF-scope relations.

Consider, for instance, the ordering principle GIVEN > NEW, which is operative in many free word order languages, and which can trigger the derivation of a clause with O > S order in German. Assuming that the scope relations to be expressed at LF are part of the input of the optimality theoretic evaluation (as is compatible with the proposal of Bobaljik & Wurmbrand 2012), an LF in which the subject outscopes the object may nevertheless be expressed by an OSV sentence if the surface order is conditioned by the GIVEN > NEW principle, which is ranked higher than PARMOV in Müller (2001).

Since all surface ordering principles form a cluster in Müller (1999), including those that encode normal word order (such as nominative > accusative), the low rank of PARMOV relative to this cluster implies that all serialization alternatives should allow scope ambiguities – in line with the corpus data presented in Webelhuth (2020) for a wide range of constructions.

Putting it differently, the idea that free word order languages should be scope-rigid because there are no strict ordering laws that would prevent the surface realization of scope (as the rigid laws of English do) is due to a misconception of the nature of free constituent order. Constituent order is not arbitrary in German or Japanese. It is governed by a set of constraints from various domains of language (pragmatics, prosody, case order, etc.) that are equally able to block the transparent surface representation of quantifier scope. There is no difference to fixed order languages in this respect. Rather, the difference resides in the fact that the ordering

---

[5] This is so in the absence of further ordering constraints on LF that might be in conflict with quantifier scope; see Bobaljik & Wurmbrand (2012) for a discussion of the role of contrastive topicality or focality in the determination of relative quantifier scope.

constraints in German or Japanese are of a more varied nature, that they may be in conflict with each other, and that conflict resolution allows for much more flexibility in the actual choice of a given constituent order (cf. Müller 1999).

Notice that this argument relies on Müller's assumption that PARMOV is ranked below the cluster of ordering constraints. In an OT-account, one would expect to find languages in which the ranking is reversed. In such a language, scope should indeed correspond to surface c-command. Presumably, Hungarian is a language of this type, at least according to the existing generalizations in the literature.

Finally, our argument makes the prediction that the availability of scope ambiguities in sentences with normal word order does not depend on the particular information structure constellation that was tested in our experiments. Whether that is true or not, i.e., whether we can identify scope ambiguities for normal word order sentences with other information structures, will have to be tested in further experiments.

4. Conclusion

We have presented the results of a covered-box experiment on the availability of inverse scope in German. While this indirect experimental method has shown good results for other semantic phenomena (Huang et al. 2013), this is the first covered-box investigation of inverse scope, to the best of our knowledge. Our target sentences involved syntactic configurations that have not been investigated in previous experimental work on German; see Radó & Bott (2018) and references therein. The configuration of interest consisted of doubly quantified transitive clauses with the reduced existential quantifier *'n/'ne* in subject position and with a distributive universal *jeder*-NP in object position, in canonical S > O word order, and with verum intonation. The target items were presented in contexts that made the NP-content of both quantified expressions given and de-accented. The context sentences also provided the semantic restriction for the universal quantifier in the target clause, in the form of a plural DP. The experimental results show that inverse scope is available to some extent in the free constituent order language German, thereby validating earlier findings for other syntactic configurations in German (Radó & Bott 2018) and empirical claims for other free constituent languages (Japanese, Russian, Hindi), as well as recent corpus findings for German in Webelhuth (2020). Furthermore, the results of the covered-box experiment replicate findings from an earlier forced choice experiment on parallel target items, in which participants were asked directly about the availability of surface and inverse scope readings. Finally, we discussed ways of modifying theoretical models of scope rigidity, such as Bobaljik & Wurmbrand (2012) and Müller (1999a), so that they can account for our results.

Abbreviations

ACC = accusative, ADJ = adjective, MASC = masculine, NEUT = neuter, NOM = nominative, PL = plural, PRT = particle, PST = past

Supplementary files

Supplementary files are available under the following link: https://doi.org/10.17605/OSF.IO/NHK56

The material includes:

- **Supplementary file 1:** Experimental_stimuli. Full set of the stimuli used in the experiment.
- **Supplementary file 2:** Experimental_results. Data set of the experimental results.

Ethics and consent

The experiment was conducted within project C02 of the Collaborative Research Council 1287: The limits of variability in language. Within the overall proposal, there was no need to obtain individual clearing of sub-experiments by the ethics committee and the project was wholesale approved without ethics statement by the DFG as proposed, including the listed experiments. In carrying out the experiment, we adhered to standard practices regarding data protection (in compliance with UP rules on data protection) and experimental ethics.

Funding information

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.

Antonyuk, Svitlana. 2015. *Quantifier Scope and Scope Freezing in Russian*. Stony Brook, NY: Stony Brook University dissertation.

Antonyuk, Svitlana. 2019. Quantifier Scope in Russian. *Glossa: a journal of general linguistics* 4(1). 54. DOI: https://doi.org/10.5334/gjgl.562

Aoun, Joseph & Li, Yen-Hui Audrey. 1993. *Syntax of scope*. Cambridge, MA: MIT Press.

Beck, Sigrid. & Gergel, Remus. 2014. *Contrasting English and German grammar*. Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110346190

Beghelli, Filippo & Stowell, Tim. 1997. Distributivity and negation: the syntax of *each* and *every*. In: Anna Szabolcsi (ed.), *Ways of Scope Taking,* vol. 65, 71–98. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-94-011-5814-5_3

Blum, Lia. 2020. *Scrambling und inverser Skopus im Luxemburgischen: Eine empirische Untersuchung* [Scrambling and inverse scope in Luxembourgish: An empirical study]. Potsdam: University of Potsdam BSc thesis.

Bobaljik, Jonathan D. & Wurmbrand, Susi. 2012. Word order and scope: Transparent interfaces and the 3/4 signature. *Linguistic Inquiry* 43. 371–421. DOI: https://doi.org/10.1162/LING_a_00094

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

Bruening, Benjamin. 2019. Experiments reveal that scope is not frozen in English double object constructions. Ms. University of Delaware.

Carden, Guy. 1970. A note on conflicting idiolects. *Linguistic Inquiry* 1. 281–290.

Chomsky, Noam. 1981. *Lectures on Government and Binding.* Dordrecht: Foris.

Crain, Stephen & Thornton, Rosalind. 1998. *Investigations in Universal Grammar.* Cambridge, MA: MIT Press.

Davidson, Donald. 1984. *Inquiries into Truth and Interpretation.* Oxford: Clarendon Press. DOI: https://doi.org/10.1093/0199246297.001.0001

É. Kiss, Katalin. 1991. Logical structure in syntactic structure. In Huang, C. T. James & May, Robert (eds.), *Logical structure and linguistic structure: cross-linguistic perspectives,* 111–147. Dordrecht: Kluwer. DOI: https://doi.org/10.1007/978-94-011-3472-9_5

Ebert, Cornelia. 2009. *Quantificational Topics – A Scopal Treatment of Exceptional Wide Scope Phenomena.* Dordrecht: Springer.

Ebert, Cornelia & Hinterwimmer, Stefan. 2010. Quantificational variability effects with plural definites: Quantification over individuals or situations? *Journal of Semantics* 27(2). 139–176. DOI: https://doi.org/10.1093/jos/ffq003

Frey, Werner. 1993. *Syntaktische Bedingungen für die semantische Interpretation: über Bindung, implizite Argumente und Skopus.* [Syntactic conditions for semantic interpretation: about binding, implicit arguments and scope]. Berlin: Akademie Verlag.

Guha, Ishani & Moitra, Swarnendu & Marty, Paul. 2021. Inverse Scope in Scrambling Languages: the case of Bangla. Paper, presented at WCCFL 2021, Tucson, Arizona.

Hayashishita, Junichi. 2013. On the nature of inverse scope readings. *Gengo Kenkyu* 143. 29–68.

Hemforth, Barbara & Fernandez, Susana & Clifton Jr, Charles & Frazier, Lyn & Konieczny, Lars & Walter, Michael. 2015. Relative clause attachment in German, English, Spanish and French: Effects of position and length. *Lingua* 166. 43–64. DOI: https://doi.org/10.1016/j.lingua.2015.08.010

Hornstein, Norbert. 1984. *Logic as Grammar.* Cambridge, MA: MIT Press.

Hornstein, Norbert. 1995. *Logical Form: From GB to Minimalism.* Oxford: Blackwell.

Huang, Yi Ting & Spelke, Elizabeth & Snedeker, Jesse. 2013. What Exactly do Numbers Mean? *Language Learning and Development* 9(2). 105–129. DOI: https://doi.org/10.1080/15475441.2012.658731

Ionin, Tania. 2001. The one girl who was kissed by every boy: Scope, scrambling and discourse function in Russian. In van Koppen, Marjo & Sio, Joanna & de Vos, Mark (eds.), *Proceedings of ConSole X*, 65–80.

Ionin, Tania. 2014. Some indefinites scope out of islands: an experimental investigation. In Aguilar-Guevara, Ana & Zwarts, Joost & Le Bruyn, Bert (eds.), *Weak Referentiality*, 45–72. Amsterdam: John Benjamins.

Kitagawa, Yoshihisa. 1990. Anti-scrambling. M.S., University of Rochester (presented at the Tilburg University Workshop on Scrambling, October 19, 1990. Available at http://www.indiana.edu/~ykling/Resourcefiles/Publicationpdfs/AntiScrambling2000.pdf).

Klima, Edward S. 1964. Negation in English. In Fodor, Jerry A. & Katz, Jerrold J. (eds.), *The structure of language*. New Jersey: Prentice-Hall.

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

Kuroda, S.-Y. 1970. Remarks on the notion of subject with reference to words like also, even or only. *Annual Bulletin* 4. 127–152.

Larson, Richard K. 1990. Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry* 21. 589–632.

Larson, Richard K. & Wu, Hongchen. 2018. Quantifier Scope and Topicality in Mandarin. Presentation at the *26th Meeting of the International Association of Chinese Linguistics (IACL)*. University of Wisconsin-Madison.

Mahajan, Anoop. 2018. Quantification in Hindi. In Keenan, Edward L. & Paperno, Denis (eds.), *Handbook of Quantifiers in Natural Language*, vol. 2, 383–430. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-3-319-44330-0_7

Matthewson, Lisa. 2004. On the Methodology of Semantic Fieldwork. *International Journal of American Linguistics*, 70(4). 369–415. DOI: https://doi.org/10.1086/429207

May, Robert C. 1977. *The grammar of quantification*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

May, Robert C. 1985. *Logical Form: Its structure and derivation*. Cambridge, MA: MIT Press.

Müller, Gereon. 1999. Optimality, markedness, and word order in German. *Linguistics* 37(5). 777–818. DOI: https://doi.org/10.1515/ling.37.5.777

Müller, Gereon, 2001. Order Preservation, Parallel Movement, and the Emergence of the Unmarked. 2001. In Legendre, Geraldine & Grimshaw, Jane & Vikner, Sten (eds.), *Optimality-Theoretic Syntax*, 279–313. Cambridge, Mass.: MIT Press.

Philipp, Mareike & Zimmermann, Malte. 2021. *On trees blocking roads and cameras recording burglars – an experimental comparison of the availability of inverse scope in English and German* [Manuscript submitted for publication]. Department of Linguistics, University of Potsdam.

Radó, Janina & Bott, Oliver. 2018. What do speaker judgments tell us about theories of quantifier scope in German? *Glossa* 3(1). 1–29. DOI: https://doi.org/10.5334/gjgl.309

Reinhart, Tanya. 1976. *The syntactic domain of anaphora*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Reinhart, Tanya. 2006. *Interface Strategies*. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/3846.001.0001

Sanford, Anthony J. & Garrod, Simon C. 1998. The role of scenario mapping in text comprehension. *Discourse Processes* 26(2–3). 159–190. DOI: https://doi.org/10.1080/01638539809545043

Scontras, Gregory & Polinsky, Maria & Tsai, C.-Y. Edwin & Mai, Kenneth. 2017. Cross-linguistic scope ambiguity: When two systems meet. *Glossa* 2(1). 1–28. DOI: https://doi.org/10.5334/gjgl.198

Szabolcsi, Anna. 1997. Strategies for scope taking. In Szabolcsi, Anna (ed.), *Ways of scope taking*. 109–154. Amsterdam: Kluwer. DOI: https://doi.org/10.1007/978-94-011-5814-5_4

Szendrői, Kriszta. 2022. (Quantifier) Scope judgments. In Sprouse, Jon (ed.), *Oxford Handbook of experimental Syntax*. Oxford: Oxford University Press.

Ueyama, Ayumi & Hayashishita, Junichi R. 2020. Sentence structure and quantifier scope in Japanese: A retrospective and reanalysis. In Jacobsen, Wesley M. & Takubo, Yukinori (eds.), *Handbook of Japanese semantics and pragmatics*, 267–308. Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9781614512073-005

Webelhuth, Gert. 2020. C-command constraints in German: a corpus-based investigation. Ms., University of Frankfurt.

Wilson, Neil L. 1959. Substances without Substrata. *The Review of Metaphysics* 12(4). 521–539.

Winkel, Rico. 2018. *Inverse Lesarten im Deutschen – Eine Experimentelle Studie* [Inverse readings in German – an experimental study]. Potsdam: Universität Potsdam BSc thesis.

Zimmermann, Malte. 2002. *Boys Buying Two Sausages Each: On the Syntax and Semantics of Distance-Distributivity*. Utrecht: LOT Publications.