Distributive markers and their exhaustivity requirements - Evidence from German

Bachelor's Project Thesis

Maximilian Velich, s2762927, max@velich.me,
Supervisors: Ana Bosnić & Jennifer Spenader

Abstract: This study is about analysing the interpretations of the German distance-distributive quantifier *jeweils*. In contrast to *each*-like individual distribution, it additionally allows for event distribution, meaning a distribution over space and time. Existing semantic analyses of this has not been empirically verified cross-linguistically. We tested different generalisations made by Bosnić (in prep.) for the Serbian marker *po* with the modification of marking the object of a transitive sentence. The first experiment consisted of a Truth Value Judgement Task with a 3x2 design conducted with 64 participants. Since the results showed ambiguity with regards to the crucial condition, we conducted a modified follow-up experiment with the intention to prime participants. It was found that individual distribution is significantly more preferred than spatial distribution. The rest of the results were unexpected, but gave rise to new ideas for further research. The goal is to disambiguate what is considered a restriction of the quantifier and to what extent the results are generalisable cross-linguistically.

1 Introduction to *jeweils*

This paper will focus on the German *distance-distributive* marker *jeweils*. While *every* and other similar quantifiers have a clear distributive interpretation, *jeweils* can distribute over temporal or spatial events that are not specified by the given sentence. Such *Anti-Quantifiers* (Choe, 1987) are found in typologically diverse languages: Gil (1990) and Oh (2001) offer accounts for Japanese and Korean; Faller (2001) elaborates on a marker in Quechua; and Balusu (2006) the marker of Telugu. An example of *jeweils* is illustrated in (1). Depending on the context, it can be translated to either a, b or c.

(1) *Jeweils* ein Affe springt.
  a) "Each monkey is jumping."
  b) "A monkey is jumping at each place."
  c) "A monkey is jumping at each time."

This is impossible for the regular quantifier *each*, because it offers only one interpretation. The nominal phrase that contains *each* forces a relationship of distributivity (Gil, 1995). This means, that the nominal phrase marked by *each* is the set of which all members need to be engaged in the activity denoted by the verb (e.g. *jumping*). This requirement of universal quantifiers is best described as exhaustivity (Balusu, 2006; Zimmermann, 2002). However, *jeweils*-like markers have multiple interpretations. Because of this, their analysis is difficult to explain.

While there are many theoretical accounts of semantic analyses of *jeweils* and its equivalents, there is little empirical evidence to support claims made about them. In an approach to provide empirical data for this, Bosnić (in prep.) challenged two different theories regarding the Serbian marker *po*. Her goal was to find the precise exhaustivity requirement of *po* (Bosnić, in prep.).

In this paper, we will follow the approach of Bosnić (in prep.) of the Serbian marker *po* and extend it to the German marker *jeweils*. We will use a similar experimental setup and we will explain why we turn to transitive sentences immediately. Furthermore, we will incorporate an analysis by Faller (2001) and argue why it offers a plausible explanation for our results.

2 Previous research

Terminology The universal quantifier *each* can occur in different positions in English. This is illustrated in (2). It can occur as determiner, in adnominal or adverbial position. However, the three sentences are essentially synonymous (Champollion, 2012). Therefore, the meaning does not depend on the position of the marker *each* and can only be interpreted as: there is a set of monkeys and every monkey in this set ate two bananas.

One can identify two different concepts, introduced by Choe (1987). The set that is distributed (bananas) is called the scope or the *Distributive Share*. The set that is distributed over (the monkeys) is called the restrictor set or the *Sorting Key*. For a distributive quantifier (like *each*), it is crucial to understand that the Sorting Key needs to be exhaustively distributed over, while the Distributive Share does not have this requirement (Balusu, 2006). For example, there could still be bananas (Distributive Share) left in (2) without the sentences being false. However, if there is even one monkey in the set of monkeys (Sorting Key) who is not eating two bananas, all three sentences in (2) would be considered false, because the exhaustivity requirement of the Sorting Key is not met. In short, determiner-like distributive quantifiers (like *each*) syntactically form constituents with the argument that is the Sorting Key, regardless of the position of the marker.

When translating the three synonymous sentences in (2) into German, one needs to use two separate markers. The marker *jed-* is used as a determiner as the head of the subject (e.g. from *Each monkey*). It behaves in the same way as *each*, meaning it distributes over individuals by selecting the Sorting Key as a syntactic argument. Balusu (2006) calls this reading participant key reading. The translation of 2a can be seen in 3a. In the adnominal or adverbial position (2b/2c), the marker *jeweils* is used (3b). The full translation can be seen in 3b (Champollion, 2012). This type of marker is called an *Anti-Quantifier* by Choe (1987) or a *distance-distributive* marker by Zimmermann (2002), because it features additional syntactic complexity in comparison with regular distributive markers. A difference to *each*-like quantifiers is that *jeweils* chooses the Distributive Share (i.e. two bananas) as argument. While *jeweils* can be interpreted in the same way as *each*, it can additionally distribute over temporal or spatial occasions, if the surrounding context offers a set of such occasions (Zimmermann, 2002). For example, the sentence listed in 3b could mean both 4a and 4b, if you imagine a situation where there were monkeys separated by spatial or temporal events that ate two bananas in each of those events.

Reading 4b is only interpreted as such if the Sorting Key (the set of occasions) is saliently offered by the context. For example, by providing a context for 4b that entails that there were multiple occasions throughout a time interval where monkeys ate two bananas, people can judge the sentence to be true, because they presumably exhaust these time intervals (Champollion, 2012). According to our German informants, if not offered a context that specifically explains such occasions, people will either reject the sentence or express their discomfort with accepting it.

(2)	a)	Determiner:	**Each** monkey ate two bananas.
	b)	Adnominal:	The monkeys ate two bananas **each**.
	c)	Adverbial:	The monkeys **each** ate two bananas.
(3)	a)	Determiner:	**Jeder** Affe hat zwei Bananen gegessen. *Each* monkey has two bananas eaten.
	b)	Adnominal/Adverbial:	Die Affen haben **jeweils** zwei Bananen gegessen. The monkeys have DIST two bananas eaten.

(4) a) *Each* of the monkeys ate two bananas.
b) The monkeys ate two bananas in *each time interval/space*.

In intransitive sentences, *jeweils* can only attach to the subject. This by definition gives rise to declaring the subject as the Distributive Share. The Sorting Key therefore needs to be provided covertly by the context.
In a sentence like (5), our German informants report that they imagine a scenario where there are multiple monkeys that form groups and in each there is a monkey jumping.

(5) *Jeweils* ein Affe springt.
DIST one monkey jumps.

Therefore, also here the interpretation is ambiguous and depending on the context. People feel the inherent need to define a set that is in need of exhaustivity, so either an overt or covert Sorting Key.

In transitive sentences, *jeweils* can attach to the subject or the object of the sentence, or both. This is exemplified in (6).

(6) **attached to the subject**
Jeweils zwei Affen essen drei Bananen.
DIST two monkeys eat three bananas.

attached to the object
Zwei Affen essen jeweils drei Bananen.
Two monkeys eat DIST three bananas.

In the two cases here, an individual distributive reading (*each*-like) or an event distributive reading is possible. In case of *jeweils* being attached to the subject, the object could potentially be the Sorting Key (if no other Key is given), or vice versa. Therefore, testing transitive sentences might be more revealing in terms of what is considered to be the Sorting Key. While *each*-like individual distributions have clear exhaustivity requirements, the additionally allowed spatial or temporal interpretation of distance-distributive markers (like *jeweils*) makes the requirement less clear. People might prefer one interpretation over the other.

The Serbian marker *po* In an attempt to identify generalisations about distance-distributive exhaustivity, Bosnić (in prep.) experimentally tested two different claims made about the Serbian marker *po* and its exhaustivity requirement. She observed that: *A set of relevant event participants defines a spatial unit and for every set of event participants, event participants have to be exhausted.* She based this conclusion on the results of a Truth Value Judgement Task. Participants were shown images of spatially scattered animals in a zoo-like environment with corresponding intransitive sentences using the marker *po* on the subject. This is exemplified in Figure 1 with the corresponding sentence *"Pleše po jedan majmun"* (in English: one monkey is dancing at different locations). This sentence would be rejected with Figure 1.

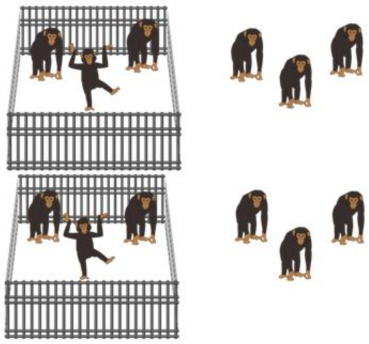


Figure 1: Example of rejected condition

This result showed that participants delimited relevant spatial units by groups of animals, regardless of whether or not they were in a cage (which was believed to maybe define a spatial unit). This approach was therefore an attempt to precisely define what is considered the Sorting Key in intransitive sentences, or in other words, what needs to be exhausted (cages as spatial units or event participants). It was found that the spatial units are defined by event participants.

Furthermore, the approach from Bosnić (in prep.) aims at offering a cross-linguistic semantic analysis with generalisations about exhaustivity. For this reason, it should also show similar results for German. Following her approach, this paper will adapt to the established observations of Bosnić (in prep.). However, as described above, the use of transitive items might be more revealing than intransitive ones, because the distributive marker can be placed on the object. Therefore, we turn directly to transitive sentences.

External Keys Another reason to focus directly on transitive sentences is suggested by Faller (2001). She states that in typical distributive readings there is a clear relation between the Sorting Key and the Distributive Share. This relation is mostly described by the verb of the sentence. However, the Sorting Key can be provided externally by the context, in which case the relation needs to extent from the internal Share to an external Key. Faller (2001) illustrates this with an analysis offered by Link (1998) for the German marker *je* (which is the short form of *jeweils* (Zimmermann, 2002)). The sentence in (7) offers an example.

(7) *Je* drei Äpfel ware faul.
Je three apples were rotten.
"Three apples per Y were rotten."

If one presents this sentence to a German speaker she will express confusion, but provided with context, it becomes less ambiguous. Faller (2001) gives as example for such context a farmer's market where three apples per basket are rotten. Here, the sentence in (7) would be felicitous. People seem to perceive this scenario as "multiple groups of apples in threes". The plurality lies in the multiple groups and not multiple events as in a typical distributive reading. To make this distinction clear Faller (2001) calls this interpretation: *group reading*. Such readings are less preferred than typical distributive readings with internal Keys, according to both Faller (2001) and also our own German informants.

According to Faller (2001), group readings can be paraphrased as *n X per Y*, where *n X* would translate to *three apples* and *per Y* to *per basket*. Here, *Y* is considered the covert Sorting Key (*the baskets*), that is not given by the sentence itself. If the key is not provided internally, the relation between the Sorting Key and the Distributive Share is also not given by the sentence (i.e. the verb). Therefore, there must be an additional relation for the example in (7). From there, she reasons that group readings derive from a relation that extends to an external Key, while regular distributive readings are provided with Keys internally. For this paper, the external Key could be provided by spatial units following Bosnić (in prep.). From this, we can see that conducting an experiment that tests intransitive sentences for *jeweils* like in the experiment by Bosnić (in prep.) for *po* is likely unnecessary, because the Sorting Key would be unambiguously provided externally. In transitive sentences, the Key could be provided internally instead, so an empirical account here would probably reveal more.

Research Questions In an attempt to test the observations of Bosnić (in prep.) and incorporating the group-reading analysis by Faller (2001), we chose the following research questions:

1. *Is a spatially distributed reading accepted when a Sorting Key is provided internally on the subject of transitive sentences where* jeweils *marks the object?*
2. *Is an individually distributed reading more widely accepted than a spatially distributed reading when* jeweils *marks the object of a transitive sentence?*

Hypotheses The expectations for the above stated questions are explained in what follows.

1. Following the research of Faller (2001), we expect that declaring the object as the Distributive Share (by attaching *jeweils*), people will find the Sorting Key either in the subject of the sentence, thus internally; or in the spatial groups that are defined by the context, thus the external Key. We distinguish between two cases: a scenario without visually distinct groups and a scenario with visually distinct groups (groups of event participants).

In the first case (no groups), people will find a regular distributive relation with the subject and the event participants. Since there are no other clues for other possible (external) Keys provided by the context, people will most probably choose a participant Key reading that essentially behaves like a universal quantifier: if the Share is exhausted in all event participants (Sorting Key), the sentence is accepted; if not, it is rejected.

In the second case (distinct groups), people will find an event-distributive relation between the Share and an external Key. Here, the Key is provided by the sets of event participants that form spatial groups, following Bosnić (in prep.). People opting for this group reading will probably accept the sentence in case of the spatial units being exhausted. However, people choosing the internal key

will most likely reject the sentence even in the exhaustive case, because not *all* participants are exhausted. They again, would choose universal quantification. Since the group reading is said to be generally less preferred, we expect a lower acceptance for this particular case.

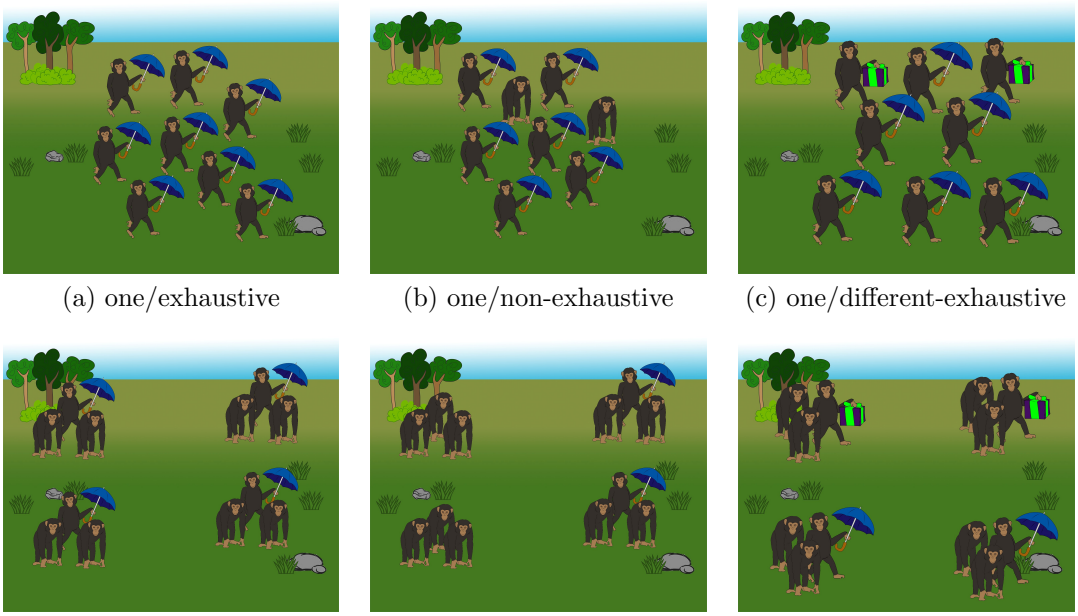
2. In the previous hypothesis we assumed that individually distributed readings (participant readings) are generally less ambiguous, and therefore likely to be accepted more. The spatially distributed readings are generally assumed to be more ambiguous, therefore likely to be accepted less. Thus, we argue that it is indeed the case that individually distributed readings are more likely to be accepted than spatially distributed readings.

3 Experiments

To test the hypotheses, we devised an experiment. Following the approach of Bosnić (in prep.), we adopted the zoo images with animals. Her observations include that Distributive Share markers come with exhaustivity requirements and that event participants play a major role in defining spatial units. Therefore, we use similar scenarios that however exclude cages or otherwise defined spaces, except the spatial separation of the animals. Bosnić (in prep.) used intransitive sentences with the marker attached to the subject, thus a possible Sorting Key needed to be provided by the picture as a salient occasion. In this paper's case, the marker is attached to the object of a transitive sentence, which means that the subject can also be considered the Sorting Key. An example sentence can be found in (8). Here, the Distributive Share is *one umbrella* and the Sorting Key *the monkeys* (if no context is provided). To allow for event-distributive readings (i.e. an external Sorting Key), a spatial separation has to be given by the context.

For clarification reasons, we introduce the six conditions of the experiment now. They can be viewed in Figure 2.

(8)	Die	Affen	halten	*jeweils*	einen	Regenschirm.
	The	monkeys	holding	DIST	one	umbrella.



(a) one/exhaustive (b) one/non-exhaustive (c) one/different-exhaustive

(d) four/exhaustive (e) four/non-exhaustive (f) four/different-exhaustive

Figure 2: The six conditions of Experiment 1

Since we want to test whether individual distribution is more widely accepted than spatial distribution, we need one independent variable to be the distribution of the animals in the scenario. We will either depict them in one big group (see Figure 2 a/b/c) or four spatially distinct groups of three animals each (see Figure 2 d/e/f). Since the plural subject refers to the animal and a possible Sorting Key, we only allow one kind of animal per condition, thus we do not mix them in the same picture.

Furthermore, the second variable is the exhaustivity. We need to differentiate between an exhaustive case (see Figure 2 a/d) and a non-exhaustive case (see Figure 2 b/e). Also, we include a different-exhaustive case (see Figure 2 c/f) that fully exhausts the relevant units but with different objects (e.g. umbrellas and gifts instead of only umbrellas).

The conclusions of Bosnić (in prep.) was: *A set of relevant event participants defines a spatial unit and for every set of event participants, event participants have to be exhausted.* We can project this statement to the example in (8). According to this, a group of monkeys defines a spatial unit if the monkeys are considered to be relevant event participants. In order to be exhausted, one monkey needs to hold one umbrella in every relevant spatial unit. These spatial units defined by event participants serve as Sorting Key. Thus, if we visually define four groups of monkeys and if there is one monkey per group who is holding an umbrella (Share), participants should interpret this as exhaustive, and therefore accept it.

3.1 Experiment 1

3.1.1 Method and design

Participants Sixty-four native German speakers from Germany, Austria and Switzerland (38 women and 26 men) ranging in age from 19 to 58 years old (mean: 29.64 years), voluntarily participated. They were recruited over various social connections. The selection process was restricted to only native German speakers. They did not receive a reward. Participants gave their informed consent. If a participant terminated the experiment half-way through, this data was excluded from the analysis. The same held for issues regarding server related problems. The data was also excluded when participants gave wrong answers for the control items (see below).

Design and procedure A Truth Value Judgment Task (TVJT) was conducted. The test had a 2x3 incomplete counterbalanced design that included 30 target items, 5 control items and 20 filler items, in 6 lists. An item consisted of one picture and one corresponding sentence. The participants were asked to verify whether the sentence *possibly* described the scenario in the given picture.

The two independent variables were explained before. The dependent variable is the answer the participants gives on the TVJT representing their interpretation of *jeweils* with the given picture. In order to provide enough target items, other scenarios with animals and objects were created, these include for instance: bears and barrels, elephants and pianos, owls and flower, and more. The full list can be seen in Appendix C.

A within subject design is used, meaning that all participants are exposed to all levels of the independent variables. Using a Latin Square, the six conditions are distributed over six lists.

An example of a NO-answer control item can be seen in Figure 3 with a corresponding sentence in (9). An example of a filler item can be seen in Figure 4 with a corresponding sentence in (10). These fillers should deceive the purpose of the test.

Before the test started we collected basic personal but not identifiable data of the subject. The consent form and a short instruction followed. It was indicated that the pictures were originally designed for children. In addition, the participant was asked to evaluate whether it is *possible* to describe the scenario with a given sentence. This was asked in order to overcome people's inherent desire to phrase explanations *optimally*. Finally, the subject was introduced to some

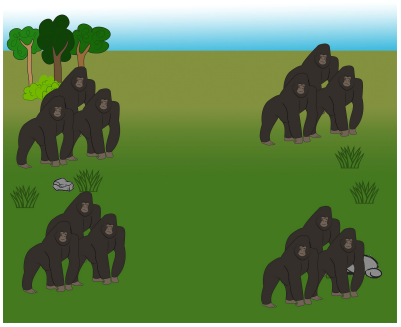


Figure 3: Control item example

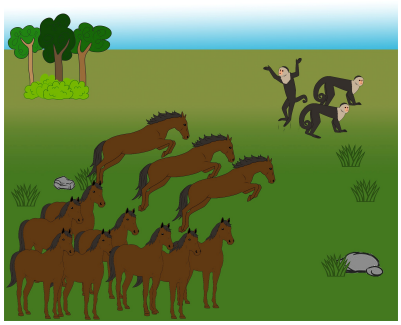


Figure 4: Filler item example

(9) Die Gorillas zeichnen *jeweils* einen Kreis.
The gorillas draw DIST one circle.

(10) Viele Pferde springen.
Many horses jump.

fictitious context. The actual test followed. Each participant was shown 55 items, one item at a time.

The data was analysed using R Studio (Version 1.0.136). Since the dependent variable was binary, a mixed effect logistic regression model was chosen. Using a linear combination of independent variables, the log odds of the results are modelled.

3.1.2 Results

The results of the one-group and the four-group condition are in Figure 5 and Figure 6, respectively. The standard error bars are plotted into the graphs.

Considering the individual distribution case of the one-group condition (Figure 5), one can see that results suggest a clear rejection of the non-exhaustive and different-exhaustive case, with a mean acceptance of 14% (SE: 2.0%) and 8% (SE: 1.5%) respectively. The exhaustive case shows an overwhelming mean acceptance of 93% (SE: 1.4%).

Considering the spatial distribution case of the four-group condition (Figure 6), one can see that results suggest a clear rejection of the non-exhaustive and different-exhaustive case again, with a mean acceptance of 6.9% (SE: 1.4%) and 3.1% (SE: 1.0%) respectively. The exhaustive case shows a mean acceptance of 43.8% (SE: 2.78%).

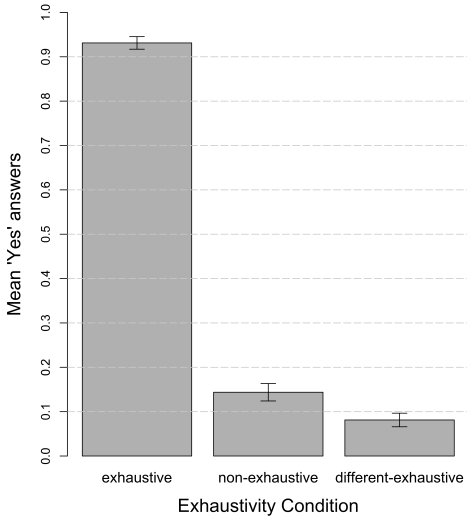


Figure 5: One-group results

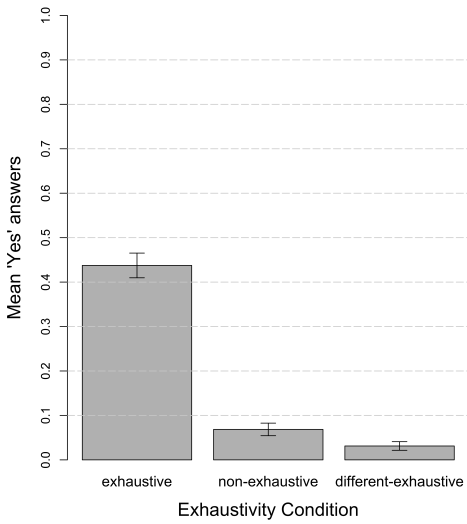


Figure 6: Four-groups results

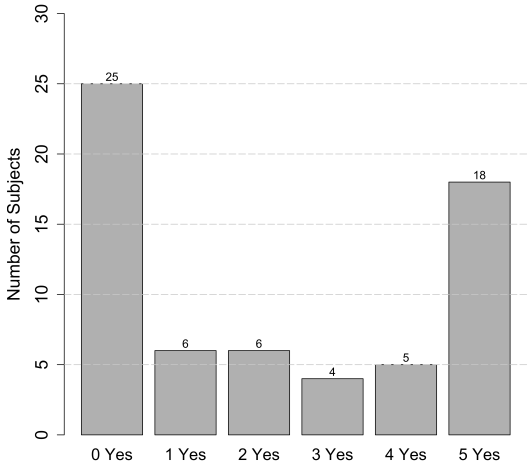


Figure 7: Distribution - *four/exhaustive*

In Figure 7, one can see the distribution of the amount of YES-answers per participant for the crucial condition (*four-groups/exhaustive*). One can read the plot as follows: 25 participants systematically rejected the condition, while 18 systematically accepted it. If one allows one mistake per participant, this formulation turns to: 31 participants rejected it consequently and 23 accepted it consequently. This leaves 10 participants (15.6% of all participants) who were inconsistent.

Statistics Using the R (R Core Team, 2014) package *lme4* (Bates, Mchler, Bolker, & Walker, 2015) a repeated measures mixed-effects logistic regression was performed to analyse the data. Multiple models were tested and their fit was evaluated by the comparison of the AIC scores. The best model turned out to be a full model of the predictor for visual distribution (the groups) and the predictor for the type of exhaustivity. It also included a random slope with a by-Subject effect of the exhaustivity and another random slope with a by-Item effect of the exhaustivity. The reference level of the exhaustivity factor was the condition *different-exhaustive*. The reference level of the groups was *four-groups*. The model formula and the R output can be viewed in Appendix A. There was a significant difference between the conditions *four/exhaustive* and *four/different-exhaustive* ($\beta = 8.445$, $z = 4.191$, $p < 0.05$). There was also a significant difference between *one/different-exhaustive* and *four/different-exhaustive* ($\beta = 2.117$, $z = 2.856$, $p < 0.05$). Lastly, the difference between *one/exhaustive* and *four/exhaustive* was significantly different from the difference between *one/different-exhaustive* and *four/different-exhaustive* ($\beta = 3.483$, $z = 3.738$, $p < 0.05$).

3.1.3 Discussion Experiment 1

When comparing the significantly higher acceptance rate of the one-group condition with the four-groups condition in the exhaustive case, we see that we have grounds to answer the second research question. While *one/exhaustive* got accepted almost uniformly, *four/exhaustive* seems to be rejected by more than half of the participants. Therefore, we can conclude that individual distribution is indeed more widely accepted than spatial distribution in *jeweils*-sentences.

When we turn to the crucial case of *four/exhaustive* we see a conflicting story. While the mean acceptance rate reveals an ambiguous account, we see a clear split in participants' interpretations when looking at Figure 7. This split suggests that individual participants were rather clear and consistent in their interpretations and answers. There was uncertainty *between* participants, thus there were two distinct populations.

The first research question asked whether a distributional reading is accepted, when the Sorting Key is provided internally. A possible Key was provided internally (the plural subject) and externally (the spatial units), and about half of the participants accepted the sentence. This uncertainty was hypothesised. However, it does not provide a clear answer to the question.

To tackle the problem of this inconclusive statement, we asked ourselves if it is possible to change the experiment in such a manner that participants would answer no, therefore ignoring the external Key. We decided on trying to prime participants with a second sentence.

The reason why the subject of the transitive sentence is allowed as Sorting Key, is because it is plural. By changing it to a singular form, the internal Key would be omitted. This would theoretically only leave the external Key as a possibility. Therefore, participants would only either choose the group reading or would just reject the sentence. However, since the internal Key is omitted, we hope

to increase the acceptance rate of this singular-subject sentence. The external Key would still not provide the optimal interpretation (according to Faller (2001)), but by leaving no alternative Key, people could decide to allow it more frequently.

By providing a highly accepted singular-subject sentence before the original condition (plural-subject), we hope to prime the participants. If they initially see the singular-subject sentence which is less ambiguous than the plural-subject sentence, they might opt for an easier rejection of the plural-subject sentence. Basically, the preference for the internal Key (and its participant reading) could be amplified against the less preferred external Key (and its group reading).

This line of reasoning is further supported by reports of participants after the experiment. When they were asked about their answer of the crucial condition (*four/exhaustive*), some reported that they answered no, because they were not entirely sure. Others reported that they answered with yes, because they felt a tendency towards it being felicitous. Apparently, even though most participants answered consistently, there was a general feeling of uncertainty that they reported. Also, some subjects were ready to change their mind if they got presented the singular-subject condition before the plural-subject condition. When presented the idea of an additional sentence (as explained above) they reported that if they would have to choose between the plural-subject sentence (*The monkeys ...*) *and* a singular-subject sentence (*One monkey ...*), they would reject the plural case and accept the singular case. Their explanation was that the singular-subject case felt "more correct" than the plural-subject case. Thus, if they accepted the singular-subject condition first, they were inclined to reject the previously accepted plural-subject condition.

For this reason, we devised a follow-up experiment that includes a second sentence with a singular subject. Since the results of the previous experiment were unequivocal regarding the one-group condition (the individual distribution) and the different-exhaustive condition, we decided to drop them from the follow-up experiment. That left us with two pictures, the four-groups condition either exhaustive or not exhaustive. The rest of the experiment was done in the same way as before.

To summarise the expectations: we expect that participants who initially said yes to the crucial condition (*four/exhaustive*) to change their mind and reject the condition instead. This could happen, because they are offered a less ambiguous sentence to describe the situation in the picture beforehand which leads to a priming effect. If this results in a more uniform rejection of the crucial condition, we could draw a more substantial conclusion.

3.2 Experiment 2

3.2.1 Method and design

This experiment was a modification of the previous one. This subsection will therefore be less detailed.

Participants Twenty-three native German speakers from German, Austria and Switzerland (16 women and 7 men) ranging in age from 20 to 60 years old (mean: 29.8 years), voluntarily participated in this experiment. They were recruited by the experimenters over various social connections. The selection process was restricted to only native German speakers. They did not receive a reward. Participants gave their informed consent.

Design and procedure A Truth Value Judgement Task was conducted. The test had a 2x2 incomplete counterbalanced design that included 24 target items, 5 control items and 20 fillers, in 4 lists. The fillers and controls were copies of the previous experiment.

The pictures were kept the same as before, see 3.1.1 for a description of those stimuli. However, as mentioned before, instead of one sentence, this experiment featured two types of sentences: singular-

(11)	Ein	Affe	hält	jeweils	eine	Banane.
	One	monkey	holding	DIST	one	banana.
(12)	Die	Affen	halten	jeweils	eine	Banane.
	The	monkeys	holding	DIST	one	banana.

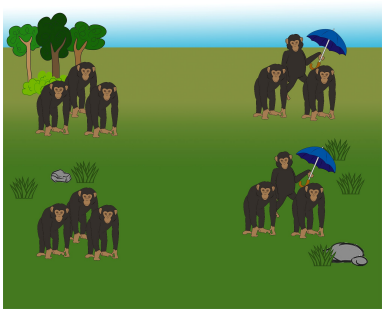


Figure 8: Example - non-exhaustive

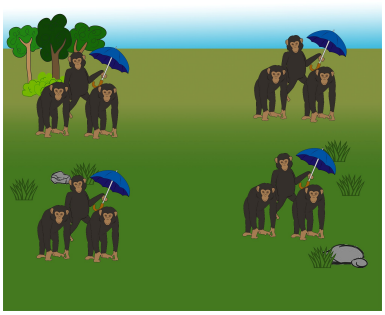


Figure 9: Example - exhaustive

subject sentences and plural-subject sentences. Examples for those can be seen in (11) for the singular case and in (12) for the plural case.

There were two independent variables. The first one is the level of exhaustivity. There are two possibilities: non-exhaustive and exhaustive (exemplified in Figure 8 and 9 respectively). The second independent variable was the sentence type, which was either singular or plural.

Therefore, in total there are four conditions (2x2). The dependent variable is the answer the participants gives representing their interpretation of *jeweils* and the given picture. The same target items of the previous experiment were used here.

The rest of the design was as previously, except for one minor change: for the priming the first item a participant got to see was always the condition *singular/exhaustive*. Thus, every experiment started with a random version of this particular item. The data was analysed as before.

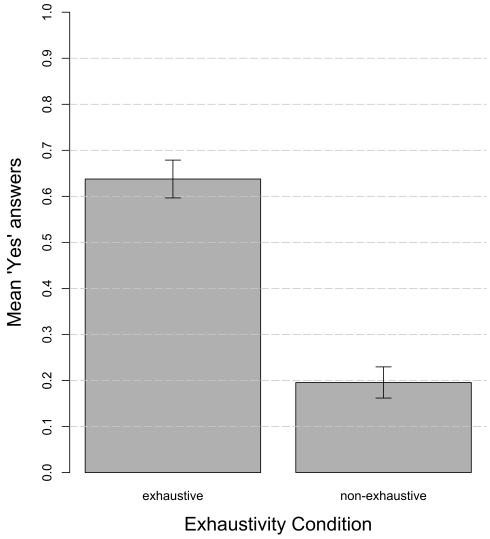


Figure 10: Singular-subject results

3.2.2 Results

The results for the singular-subject and the plural-subject condition can be seen in Figure 10 and Figure 11, respectively. The standard error bars are plotted into the graphs.

Considering the singular-subject conditions (Figure 10), one can see that results suggest a rejection of the non-exhaustive condition, with a mean acceptance of 19.6% (SE: 3.4%). The exhaustive case shows mean acceptance of 63.8% (SE: 4.1%).

Considering the plural-subject conditions (Figure 11), one can see that results suggest a rejection of the non-exhaustive condition, with a mean acceptance of 13.0% (SE: 2.9%). The exhaustive case shows a mean acceptance of 52.2% (SE: 4.3%).

In Figure 12, one can see the distribution of the amount of Yeses per participant for the same crucial condition of the last experiment (*plural/exhaustive*). One can read the plot as follows: 4 participants consequently rejected the condition, while 5 consequently accepted it. If one allows one mistake per participant, this formulation turns to: 6 participants rejected it consequently and 6 accepted it consequently. This leaves 11 participants (47.8% of all participants) who were inconsistent.

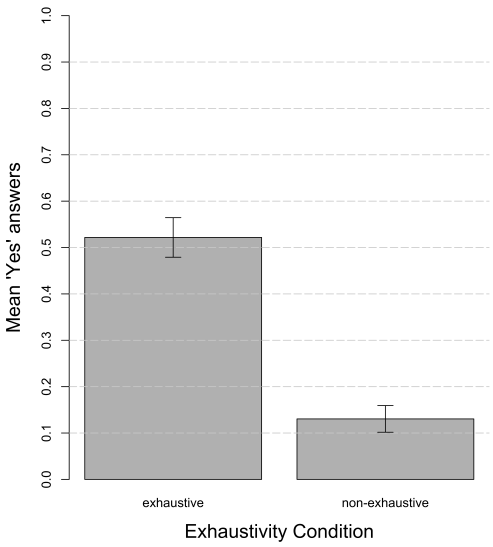


Figure 11: Plural-subject results

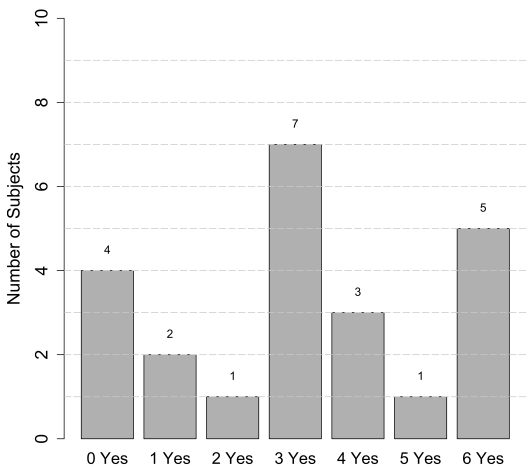


Figure 12: Distribution - *plural/exhaustive*

Statistics Using the R (R Core Team, 2014) package *lme4* (Bates et al., 2015) a repeated measures mixed-effects logistic regression was performed to analyse the data. Multiple models were tested and their fit was evaluated by the comparison of the AIC scores. The best model turned out to be a full model of the predictor for the type of exhaustivity (exhaustive or non-exhaustive) and the predictor for the type of sentence (singular or plural). It also included a random slope with a by-Subject effect of exhaustivity and a random intercept with a by-Item effect. The reference level of exhaustivity was *exhaustive*. The reference level of the sentence type was *plural*. The model formula and the R output can be viewed in Appendix B. There was a significant difference between the conditions *plural/non-exhaustive* and *plural/exhaustive* ($\beta = -6.299$, $z = -3.417$, $p < 0.05$). There was also a significant difference between *singular/exhaustive* and *plural/exhaustive* ($\beta = 0.7450$, $z = 2.391$, $p < 0.05$).

3.2.3 Discussion Experiment 2

The follow-up experiment did not have the expected results. The mean acceptance rate is even closer to a 50% split than before. Furthermore, even though the acceptance rate of the exhaustive singular-subject condition was higher than its plural equivalent, it was not as high as expected.

Even though the results are not the desired ones, the mixing of the singular sentences into the plural ones had definitely some effect on the distribution of the crucial case. When comparing the distribution figures of the crucial conditions (Figure 7 and Figure 12), one will see that participants turned from mostly systematically answering, to mostly irregular patterns. Almost half of the participants answered inconsistently.

In terms of the first research question, we can still not draw a decisive conclusion. It seems that participants perceive *jeweils* as even more vague than initially expected. Apparently, mixing a similar sentence with the crucial one leads to overall inconsistencies and uncertainty. The distribution plot (Figure 12) basically tells us that the participants were literally confused what to answer.

4 General Discussion

Conclusions From the results of the experiments we can state some observations:

- The individual distribution conditions seem to be interpreted uniformly depending on the level of exhaustivity.
- When presented with a single isolated condition of distance-distributive marker sentences, participants form distinct groups: systematic accepters and systematic rejecters.
- When presented with mixed conditions of distance-distributive marker sentences, participants do not form groups and answer inconsistently overall.
- For about half of the participants *jeweils* is exhausting sets of event participants, provided that they choose the external Key, thus the group reading.

With respect to the research questions and hypotheses, we can state the following:

1. *Is a spatially distributed reading accepted when a Sorting Key is provided internally in the subject of transitive sentences where* jeweils *marks the object?* As argued before, we cannot yet answer this question. We have reasons to believe that it is indeed the case that the spatially distributed reading is accepted, but not preferred, which results in partial rejection.
2. *Is an individual distributed reading more widely accepted than a spatially distributed reading when* jeweils *marks the object of a transitive sentence?* We answered this question in the discussion of the first experiment. The results were unambiguous here. It is indeed the case that an individual distributed reading is more widely accepted (preferred) than a spatially distributed reading.

In the beginning we stated that we would like to help develop a unified cross-linguistic theory about distributive markers. Unfortunately, our results are not clear. They emphasise how vague the notion of exhaustivity of external Sorting Keys actually is. Our results showed that even native speakers apparently are easily confused by a not-so-uncommon distributive marker.

Criticism and Speculation The participant reading shows much higher availability and therefore acceptance in general. Let us assume that all participants followed the instructions closely and they evaluated the sentences with regards to a *possible* (not *optimal*) accordance with the provided pictures. Then, one might speculate that the availability of the participant reading is in fact so high that half of the people do not even consider looking for an external Key in the context.

In fact, this leads us to one of the major drawbacks of the experimental setup. If roughly 44% of the participants are willing to accept the group reading, which stems from an external Sorting Key, then we have reason to speculate that at least a part of the rejecters might not have followed the instructions with enough care. This is because 44% is not an insignificant amount of people and there must be some reason for this. Even though, it was clearly stated in the instructions, we should have paid even closer attention to communicate this crucial information with the subjects as clear as possible. People are generally used to avoid ambiguous statements for obvious reasons. Therefore, some of the subjects who were less inclined to meticulously follow the experiment might have disregarded the focus on the difference between *possible* and *optimal*. Thus, we have speculative reasons to believe that the acceptance rate will be higher, if participants are more severely confronted with this specific line in the instructions. Note that this argument can also be turned around: participants could have also lazily accepted the sentences because they did not take the time to fully contemplate, therefore accepting it upon perceiving its validity as a syntactically correct sentence. While this could be the case, we believe that the former situation is much more likely.

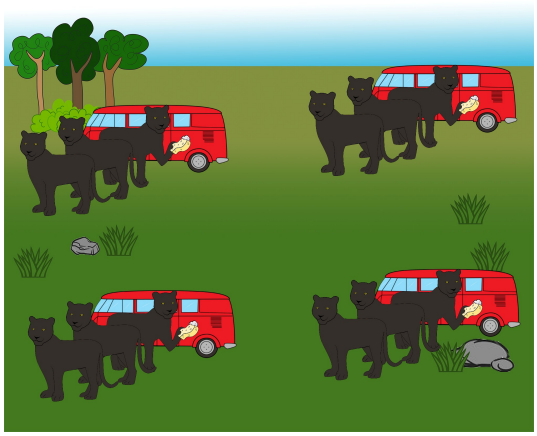


Figure 13: Group activity illustrated

Another interesting observation is that participants reported that they rejected the crucial condition because: "the activity was not a group activity". Their reasoning for rejection was that e.g. *holding an umbrella* is not a group activity as much as e.g. *washing a van* is. If one monkey is *holding an umbrella*, clearly the other monkeys in the spatial unit are not also *holding an umbrella*. Thus, the activity is projected on a single individual. However, if there is a group of panthers next to a van, and only one of them is e.g. holding a sponge to wash that van, the activity of *washing a van* can still be projected onto the whole group (i.e. spatial unit). The other panthers might be currently resting from *washing a van*, which would not exclude them from the general event of *washing a van*. This particular example can be seen in Figure 13.

The problem here is obvious. If participants interpreted these few conditions in the latter way, they would not exhaust spatial units, but just participants, because according to that logic all panthers are washing a van (maybe not all at the same time and not all with a sponge, but they are all *in the act* of washing). This nuance between the activities was not considered originally, and could lead to skewed data. However, there are only few items that could be interpreted as such. Although, it should be considered next time, we think it is not crucial problem now.

Future Research As a follow-up experiment we propose the use of a Likert scale to offer participants more choices. This could allow for a more thorough analysis of participants' interpretations. The reason is given by an interview with one participant. She reject the crucial condition (*four/exhaustive*) in the experiment systematically. Afterwards, she was asked to rate the crucial condition (*four/exhaustive*) on a scale from 1 to 10 (10 being 100% appropriate). She chose a 6. When asked if she would therefore accept the sentence (even though it was not a strong yes), she denied and replied that she would still reject the sentence next time. She would only accept the sentence when she would give it a 7 or higher. This again adds to the overall apparent ambiguity of the *native* speakers. Providing a Likert scale could solve this communication problem to some extend by allowing for a deeper analysis.

Another possible follow-up is a modification to the second experiment. Previously, we tried to prime participants. However, the priming was very subtle: only the first item (out of 49) was a singular-subject sentence. The rest of the items were randomised like in the first experiment. Therefore, we propose changing the experiment more significantly. With the current setup, participants see a picture with a corresponding sentence, which is either singular or plural. The changed setup would include the same picture but both sentences (singular-subject and plural-subject) side-by-side. Participants would

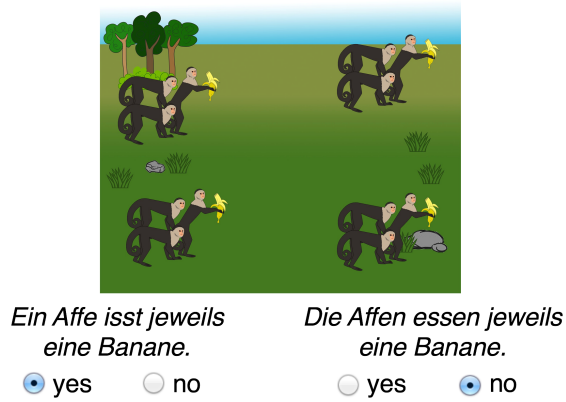


Figure 14: Sketch of the modified test

have to opportunity to evaluate both sentence at the same time. This test scenario is illustrated in the sketch in Figure 14 to help understand the idea.

Being forced to evaluate both sentences next to each other, the participants are implicitly forced to contemplate on both sentences longer than before. The expected outcome here is that people will tend towards accepting the singular-subject sentence (on the left in Figure 14), because it features only one Key. If they decide to do so, they might reject the plural-subject sentence (on the right in Figure 14), with one internal and one external key, with greater ease, since they already accepted the left one. This could further shine light on the ambiguity of the external Key described by Faller (2001).

As one can see, there are still a lot of opportunities for further research to close the gaps in understanding what could become a uniform cross-linguistic theory of distributive quantifiers.

References

Balusu, R. (2006). Distributive reduplication in telegu. *Proceedings NELS, (Conf 36)*, 39–52.

Bates, D., Mchler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, Articles*, *67*(1), 1–48. Retrieved from `https://www.jstatsoft.org/v067/i01` doi: 10.18637/jss.v067.i01

Bosnić, A. (in prep.). Spatial distribution and exhaustivity requirements for distributive markers - evidence from Serbian.

Champollion, L. (2012). Each vs. jeweils: A cover-based view on distance distributivity. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7218 LNCS*, 251–260.

Choe, J.-W. (1987). *Anti-Quantifiers and a Theory of Distributivity* (PhD thesis). University of Massachusetts, Amherst.

Faller, M. (2001). The problem of Quechua -nka - Distributivity vs. group forming.

Gil, D. (1990). Markers of Distributivity in Japanese and Korean. In H. Hoji (Ed.), *Japanese/korean linguistics* (pp. 385–393). Stanford: The Center for the Study of Language and Information.

Gil, D. (1995). Universal quantifiers and distributivity. , 321–362.

Link, G. (1998). Algebraic semantics in language and philosophy. *Stanford: CSLI Publications*, *74*.

Oh, S.-R. (2001). Distributivity in an event semantics. *Semantics and Linguistic Theory*, *11*, 326–345. doi: 10.3765/salt.v11i0.2844

R Core Team. (2014). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `http://www.R-project.org/`

Zimmermann, M. (2002). *Boys buying two sausages each: On the syntax and semantics of distance-distributivity* (PhD thesis). University of Amsterdam.

Appendices

Appendix A

The best fitted model is expressed in the following formula:

$$\text{Answer} \sim \text{Exhaustivity} * \text{Group} + (1 + \text{Exhaustivity} \mid \text{Subject}) + (1 + \text{Exhaustivity} \mid \text{Item})$$

The summary of the fixed effects follows in the table.

	Estimate	**Std. Error**	**z-value**	**Pr** $(> \lvert z \rvert)$	
Intercept (four/diff-exhaustive)	−8.82181	1.97379	−4.469	$7.84e^{-06}$	***
exhaustive	8.44547	2.01526	4.191	$2.78e^{-05}$	***
non-exhaustive	1.16520	2.24694	0.519	0.604058	
one-group	2.11651	0.74095	2.856	0.004284	**
exhaustive:one-group	3.48275	0.93167	3.738	0.000185	***
non-exhaustive:one-group	−0.05428	0.89221	−0.061	0.951488	

Appendix B

The best fitted model is expressed in the following formula:

$$\text{Answer} \sim \text{Exhaustivity} * \text{Sentence} + (1 + \text{Exhaustivity} \mid \text{Subject}) + (1 \mid \text{Item})$$

The summary of the fixed effects follows in the table.

	Estimate	**Std. Error**	**z-value**	**Pr** $(> \lvert z \rvert)$	
Intercept (plural/exhaustive)	0.1497	0.4525	0.331	0.74707	
non-exhaustive	−6.2990	1.8434	−3.417	0.0006333	***
singular-sentence	0.7450	0.3116	2.391	0.016811	*
non-exhaustive:singular	0.4368	0.6226	0.702	0.482970	

Appendix C

Die Schimpansen putzen jeweils ein Fenster.
Die Nashörner schieben jeweils ein Auto.
Die Eulen halten jeweils einen Blumenstrau.
Die Affen essen jeweils eine Banane.
Die Koalas bemalen jeweils eine Vase.
Die Panther waschen jeweils einen Van.
Die Bären schieben jeweils ein Fass.
Die Zebras halten jeweils einen Ballon.
Die Nilpferde essen jeweils eine Portion Eis.
Die Tiger bemalen jeweils ein Ei.
Die Elefanten waschen jeweils ein Boot.
Die Affen schieben jeweils einen Kinderwagen.
Die Pandas halten jeweils einen Ball.
Die Löwen essen jeweils eine Keule.
Die Gorillas streichen jeweils ein Brett.
Die Papageien tragen jeweils ein Geschenk.
Die Panther ziehen jeweils einen Schlitten.
Die Bären schieben jeweils ein Fass.
Die Löwen bauen jeweils ein Lego Haus.
Die Tiger trinken jeweils eine Packung Milch.
Die Eulen tragen jeweils einen Hut.
Die Papageien ziehen jeweils ein Spielzeugauto.
Die Nashörner heben jeweils eine Hantel.
Die Pandas bauen jeweils einen Schneemann.
Die Gorillas trinken jeweils eine Flasche Coca Cola.
Die Schimpansen halten jeweils einen Regenschirm.
Die Nilpferde ziehen jeweils einen Baum.
Die Elefanten heben jeweils ein Klavier.
Die Zebras bauen jeweils eine Sandburg.
Die Koalas trinken jeweils einen Saft.