university of groningen / faculty of science and engineering

Cumulative Readings in English and Dutch

Bachelor's Project Thesis

Carynna De Zilva, s2752573, c.e.u.de.zilva@student.rug.nl,
Supervisors: Prof Dr J. K. Spenadar & A. de Koster

Abstract: This paper investigates a number of questions with respect to cumulative readings in order to determine if cumulative readings are treated similarly to collective readings. We conducted a series of experiments in order to answer them. First, to determine if the acceptability of the distributive marker *each* was sensitive to verb type, a truth value judgment task was conducted in Dutch. We found that cumulative interpretations were similar to collective interpretations in that they were accepted more with independent verbs. The second question to answer was whether there was a significant difference in acceptance rate for cumulative interpretations with *each* between Dutch and English. By replicating the experiment in English, the results showed no difference in acceptance rates between the two languages. Furthermore, we investigated if the Dutch *elke* was more similar to English *every* than the English *each* as it's occasionally claimed to be. The results showed that this appears to be true. Finally, we explored how cumulative and collective interpretations might differ in acceptability by conducting an acceptability judgment task and self paced reading task to judge complexity through response time.

1 Introduction

Sentences containing plural DPs can be interpreted in several ways. There are distributive interpretations, collective interpretations, and cumulative interpretations. Take sentence (1):

(1) The girls are brushing a goat

In the distributive interpretation (see figure 1.1), each of the girls is brushing one goat each. In both the collective and cumulative interpretations, this is a collaborative task. But while in the collective interpretation (see figure 1.2), all three girls are brushing one goat, in the cumulative interpretation (see figure 1.3), where one girl brushes one goat, the other two are jointly brushing another. Plural sentences are commonly ambiguous between these three interpretations. The focus of this thesis will be on cumulative interpretations.

Now consider the following sentences:

(2) Each girl is bushing a cow

(3) The girls are brushing a cow

In previous research, it was found that *each* is preferred by adults for distributive interpretations, but disliked in collective readings, therefore, making it a marker for distributive interpretations (Pagliarini et al. (2012); de Koster et al. (2017)). A definite plural subject such as in (3) is semantically compatible with all three readings, but adults have been shown to disprefer it in distributive readings (Pagliarini et al. (2012); de Koster et al. (2017)). For the definite plural subject, research has shown that participants have a strong preference for collective (Frazier et al. (1999); Kaup et al. (2002)) and cumulative interpretations (Ussery, 1998) but it was marginal in distributive interpretations.

Despite *each* being a marker for distributivity, there have been unexpected results found, and particularly in Dutch, regarding distributive and collective interpretations. Dutch adults show unexpectedly high acceptance rates of the distributive quantifier *each* describing collective situations (Rouweler and Hollebrandse (2015); de Koster et al. (2017)). De Koster et al. (2017) found a very high rate of acceptance of around 36% despite it being predicted to be overwhelmingly rejected. This is problematic if *each* is claimed to be a distribu-

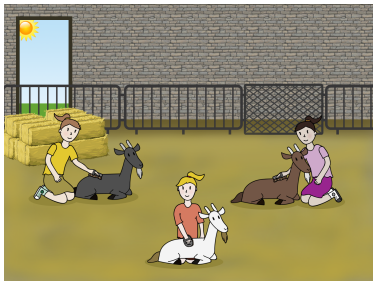

Figure 1.1: Distributive

Figure 1.2: Collective

Figure 1.3: Cumulative

tive marker. This led to claims that perhaps the Dutch *elke (each)* is not really distributive, or less distributive than its English counterpart *each*.

To examine this issue more closely, De Koster et al. (2020) proposed that cases where *each* was highly acceptable with collective situations may have been due to the types of actions denoted by the verbs used in the Dutch experiments. It was predicted that some verbs are more compatible with distributive than collective interpretations. They found that *each* can be used to describe collective situations especially if the verb implied independent actions. When the verb implied a group action that required teamwork from dependent actions, collective situations become marginal in combination with *each*.

In this work, we hope to answer two questions with respect to cumulative interpretations in order to determine if people treat cumulative interpretations similarly to collective interpretations. The first question being, is there a difference in acceptance rates between independent and dependent verbs with cumulative interpretations? Furthermore, if this is similar to the difference found with collective interpretations. The second is whether there will be a significant difference between English and Dutch. We investigate these two research questions in a series of three experiments. We find that both the Dutch and English *each* is accepted more with independent verbs than with dependent verbs, as also seen in collective interpretations. However, we found no significant difference between the acceptance of *each* with cumulative situations between English and Dutch.

2 Background

In the literature on distributivity, collective and cumulative interpretations are often combined rather than compared. Kratzer (2007) proposed a theory that collective and cumulative interpretations are derived from the same syntactic structure, and that there is no significant processing preference for either interpretation. In contrast, Sternefeld (1998)'s theory proposed that they are all derived from distinct structures in which collective interpretations are the most structurally simple. For Sternfeld, cumulative interpretations are more complex due to it having one instance of movement, while distributive are the most complex as they involve two instances of movement. Similarly, on Kratzer's account, distributive interpretations are also the most complex as distributives involve additional structure. Lee (1997) defines cumulative readings as a scope independent reading where each member of either set of members is connected to at least one member of the other set. He found children to have a lower acceptance rate with cumulative readings than in adults when in combination with numerically quantified expressions. Another study by Musolino (2009) showed that adults have a relatively low acceptance rate of cumulative interpretations with the quantifier *each* as compared to children (who have high acceptance rates). His experiment tested participants' interpretation of sentences containing numerically quantified expression (e.g. three boys are holding two balloons) by using a truth value judgment task. The sentences differed only with respect to the type of quantified object (e.g. two N vs each N). They were shown a short animated vignette, accompanied with a statement. They were then tasked to determine if the state-

ment matched the vignette. The task was identical for both adults and children. The results showed that with the quantifier *each*, adults had a 17.1% acceptance rate while children had a 54.6% acceptance rate.

De Koster et al. (2020) proposed that the acceptance of *each* in collective situations is affected by the type of verbs, and this could explain the unexpected high acceptance rates of *each* in collective situations. They distinguished two different types of verbs: *dependent* and *independent*. As discussed by de Koster et al. (2020), *independent verbs* are verbs that describe situations that permit interpreting each agent as acting in a separate subevent, regardless of the number of objects (see figures 1.1 to 1.3. Meanwhile, *dependent verbs* are compatible with situations where the agents are depending on each other to complete the action, and thus, cannot be interpreted as separate subevents (see figure 2.1).

Figure 2.1: Dependent-Collective

Consider figure 2.1, where a barrel carrying event is shown. The picture implies that the action cannot be done alone (and thus, is collective) as the agents are relying on each other. It may be possible that if one of the boys leaves, then the barrel is too heavy for the remaining two boys to carry the barrel together. Therefore, it may not be possible to use a distributive marker like *each* to describe the scene, as it requires every boy to carry their own barrel, thus, cannot be accepted. However, when using a verb like 'brush', as in figure 1.2 (showing three girls brushing a single goat), if one of the girls leaves, it does not affect the brushing actions of the other girls. Therefore, we can say that each individual girl brushes a goat. This event seems to allow for independent subevents without explicitly displaying a 1-to-1 pairing of the goats and girls. This therefore suggests that independent verbs will allow the use of the distributive marker *each* in this 'collective' scene.

De Koster et al. (2020) also tested the effects of verb type on the acceptance of definite plurals with distributive interpretations. The study found no significant difference in verb type for *the* in distributive interpretations. So how would the acceptance rates differ in a cumulative interpretation of the same events? When discussing the interpretation of sentences containing plural DPs, collective, distributive, and cumulative readings are commonly discussed. While the former two have been highly researched, the latter has very little experimental work investigating them despite the many claims that have been made about it. We hope to contribute valuable findings by exploring this difference (or lack of difference). In this work, the acceptance rates of *each* and *the* is studied by investigating the role of verb semantics in cumulative interpretations.

In Experiment 1, this was tested in Dutch speaking adults with a between subject design as not to influence the participants by minimising the learning effects across conditions. With several experimental results showing that Dutch participants were more accepting of distributive markers in collective situations than English participants (Rouweler and Hollebrandse, 2015; de Koster et al., 2020), it had been claimed that Dutch distributive markers are less distributive than English ones. This also leads to the question, will there be a significant difference between Dutch and English in cumulative situations? This hypothesis was tested in Experiment 2 by testing English adults with the same procedure as in Experiment 1. In Experiment 1 and 2, cumulative interpretations were studied with truth-value judgments and a preference task.

3 Experiment 1

3.1 Participants

57 native Dutch speaking participants took part in Experiment 1. The Dutch participants were recruited from Prolific and were paid for their participation. The distribution of the participants and

their main features are reported in table 3.1.

Table 3.1: Number of Dutch adult participants per participant group.

Verb Type	# of participants	Mean age
Dependent	30	27.6
Independent	27	26.1

3.2 Design and procedure

Experiment 1 is a truth value judgment task in which participants were shown a picture (such as 1.1) on a computer screen and a sentence (such as (1)) presented aurally. They had to judge whether the picture and the sentence matched by clicking yes or no. They were also asked to rate how confident they were with their choices by choosing their confidence on a scale from 1-5. Where in some cases, they had to explain their choices.

There are two factors to experiment 1, **subject type** and **verb type**. The experiment includes four experimental conditions. The pictures depicted a cumulative situation (figures 1.3 and 3.1) and the sentences were all of the form Subject-Verb-Indefinite object, beginning either with the definite plural 'de' (the) or the quantifier 'elke' (each). Two different verb types were also distinguished: dependent and independent verbs.

Figure 3.1: Dependent-Cumulative

For the dependent verbs, the four verbs used were 'tillen', 'dragen', 'vasthouden' and 'trekken' (in English: 'lift', 'carry', 'hold' and 'pull'). The four verbs used for the independent verbs were 'aaien', 'borstelen', 'kammen' and 'wassen' (in English: 'pet', 'brush', 'comb', and 'wash'. An example of each type of sentence in Dutch and English are:

Dependent

(4) De jongens droegen een vat.
The boys were carrying a barrel.

(5) Elke jongen droeg een vat.
Each boy was carrying a barrel.

Independent

(6) De meisjes borstelden een geit.
The girls were brushing a goat.

(7) Elk meisje borstelde een geit.
Each girl was brushing a goat.

Participants received 24 target items (6 per condition), 20 filler items, used to mask the goal of the experiment, and 6 control items to check if the participants were paying attention, giving 50 items in total. Each participant was given 4 practice items in order to get used to the experiment (which were relatively easy). We have 24 distinct situations with a balanced design across all participants. The experiment was a latin square design and participants did not see the same conditions for each situation. The items were distributed over four lists and were presented randomly.

3.3 Predictions

Although Experiment 1 tests the Dutch *each 'elke'* and *the 'de'*, we will refer to them as *each* and *the* in section 3 for clarity. The four conditions in Experiment 1 are each-dependent, each-independent, the-dependent, and the-independent. Due to the previously found higher acceptance rate for items of *each* in collective situations for independent verbs than dependent verbs, the prediction is that people will reject the each-dependent verbs, and thus, we expect a higher acceptance rate for items of condition *each* for independent verbs than for dependent verbs. This follows the claims that cumulative is more similar to collective than distributive (Kratzer (2007); Ussery (1998)). Furthermore, as it was found that there was no significant difference in verb type found for *the* in collective situations (de Koster et al. (2020)), we expect to see the same with *the* here.

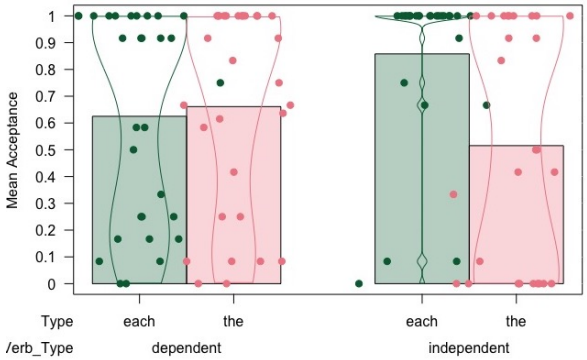

Figure 3.2: An RDI plot of the results from Experiment 1 (with the bars showing mean acceptance for dependent and independent verbs for both subject types in Dutch. Outlines show the variation in distribution and dots show individual mean responses.)

3.4 Results

The descriptive statistics can be seen in table 3.2.

Table 3.2: Proportion of 'yes' responses (with standard deviation (SD)) in Experiment 1.

Condition	Verb Type	Results
Each	Dependent	0.63 (0.49)
	Independent	0.86 (0.35)
The	Dependent	0.66 (0.47)
	Independent	0.52 (0.50)

Sentence type *each* is accepted more with the independent verb type, than the dependent verb type. Additionally, with sentence type *the*, it is more accepted with dependent verb type then independent. However, the difference is smaller with *the* than with *each* sentences. Overall the *each* condition with an independent verb type had the highest acceptance rate amongst all conditions.

In figure 3.2 we see an RDI plot of the data. There is a difference between each-independent and the rest of the conditions, with it being accepted the most. The other conditions show a hourglass shape, showing a weak binomial distribution. The-independent is especially binomial. Its mean is approximately 50% due to the fact that it was either accepted or rejected by most participants, rather than because it was accepted half the time by the participants, with only a small number of people in between.

Generalised mixed effect modelling with R was used to analyse the data. The full model can be seen in table A.1 in Appendix A. For the models, a random intercept was included for the participants. There was a main effect found with verb type ('independent'), and also for sentence type ('the'), as well as an interaction between the two ($\beta = -4.148; z = -9.528; p < 0.01$).

The post-hoc Tukey multiple comparisons showed that each-independent was accepted significantly more than each-dependent ($\beta = 2.832; z = 3.062; p < 0.01$), and the-dependent ($\beta = 2.573; 2.781; p = 0.02$). Lastly, the-independent was accepted significantly less than each-independent ($\beta = -3.889; z = -10.076; p < 0.01$). There was no significance between *the* with independent and dependent verbs.

3.5 Discussion

The distributive marker *each* was highly accepted with cumulative actions. As predicted, *each* was accepted more with independent verbs than with dependent verbs. As compared to de Koster et al.'s 2020 study in which the same pictures were used with distributive and collective situations, *each* was accepted with collective readings at a rate of 21% for dependent verbs and 51% for independent verbs and fully accepted with distributive readings for both verb types. These results show that cumulative readings appear to be more similar to collective than distributive readings in their acceptance of a distributive marking as we find acceptability of *each* at 51% for dependent verbs and 78% for independent verbs. This is also consistent with previous theories proposed by Kratzer (2007) and Ussery (1998) claiming that cumulative interpretations are more similar to collective than distributive interpretations.

Furthermore, we see that cumulative readings appear to have the same sensitivity to different verb types as collective readings seen in previous research in Dutch. The results showed that there was a significant difference in verb type for sentence type *each*, while sentence type *the* showed no significant difference in verb type. We want to explore whether this pattern that cumulative is more similar to collective than distributive in the English language. Specifically, if it has the same sensitivity

to the difference in verb type. Thus, we replicated Experiment 1 in English in Experiment 2.

4 Experiment 2

Previous research with Dutch participants often found that Dutch *elke* was being accepted with collective interpretations at an unexpected high rate (Rouweler and Hollebrandse, 2015; de Koster et al., 2017, 2020). This research then suggested that the Dutch *elke* might be a weaker distributive marker than the English *each* in order to explain the high acceptance rate of *elke* in collective situations. Experiment 1 showed that the acceptance of Dutch *elke* in cumulative situations strongly correlated with the verb types used in the predicates, similar to previous research seen in collective situations. Therefore, adding to new explanations about the distributive characters of *elke* and *each*. Experiment 2 is a replication of Experiment 1, but with native English speakers, where we expect to also see a difference in the acceptance rate of the different verb type. This follows de Koster et al. (2020)'s results in which they saw similar results in their Dutch and English experiments. The participants received the same pictures as in Experiment 1, and all sentences were translated to English.

4.1 Participants

52 native English speaking participants took part in Experiment 2. The participants were recruited from Amazon Mechanical Turk and were paid for their participation. The distribution of the participants and their main features are reported in table 4.1.

Table 4.1: Number of English adult participants per participant group.

Verb Type	# of participants	Mean age
Dependent	25	36.4
Independent	27	36.6

4.2 Results

The descriptive statistics can be seen in table 4.2 below:

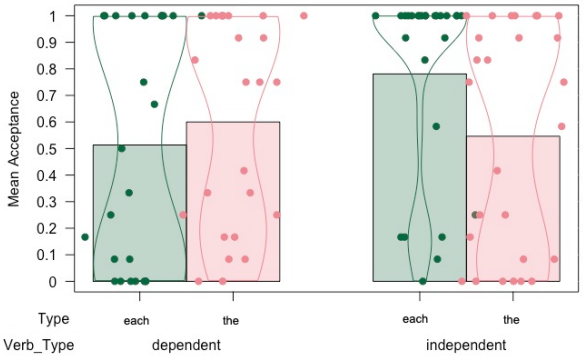

Figure 4.1: An RDI plot of the results from Experiment 2 (with the bars showing mean acceptance for dependent and independent verbs for both subject types in Dutch. Outlines show the variation in distribution and dots show individual mean responses.)

Table 4.2: Proportion of 'yes' responses (with SD) in Experiment 2.

Condition	Verb Type	Results
Each	Dependent	0.51 (0.50)
	Independent	0.78 (0.41)
The	Dependent	0.60 (0.49)
	Independent	0.54 (0.50)

Subject type *each* is accepted more with the independent verb type, than dependent verb type. While subject type *the* does not seem to have too much of a difference with acceptance rate in either verb types. So it did not matter what the verb type was when looking at the-cumulative conditions. Overall the *each* condition with an independent verb type had the highest acceptance rate amongst all conditions.

In figure 4.1 we see an RDI plot of the data. As in Experiment 1, there is a clear difference between each-independent and the rest of the conditions, with it being accepted the most. For the other three conditions, we see a hourglass shape, showing a weak binomial distribution. The-independent and each-dependent are especially binomial. Their means are approximately 50% due to the fact that it was either accepted or rejected it by most participants, rather than because it was accepted half the time by the participants

Using generalised mixed effect logistic modelling,

the data was analysed to check for a significant difference. The full model can be seen in Table A.2 in Appendix A. For the models, a random intercept was included for the participants. There was a main effect found with verb type 'independent', and also for sentence type *each*, as well as an interaction between the two ($\beta = 3.360; z = 8.344; p < 0.01$).

The post-hoc Tukey multiple comparisons showed that each-dependent sentences was accepted significantly less than the-dependent sentences ($\beta = -0.6130; z = -2.849; p = 0.0167$). Each-independent was accepted significantly more than each-dependent $\beta = 2.613; z = 2.808; p = 0.0188$), and lastly, each-independent was accepted significantly more than the-independent ($\beta = 2.747; z = 8.099; p < 0.01$). There was no significance between *the* with independent and dependent verbs.

4.3 Discussion

The distributive marker *each* was highly accepted in this experiment. As predicted, subject type *each* was accepted more with independent verbs than with dependent verbs. The results showed that there was a significant difference in verb type for condition *each*, while *the* seemed to not be affected by verb type. These results are similar to those found in Experiment 1.

Something notable to add is that even with independent verb types, *each* was not fully accepted in either experiment (1 and 2). The same outcome was seen by de Koster et al. (2020) in collective situations. They provided two possible explanations for this. The first being, that aside from requiring distributivity, *each* carries another requirement. Tunstall (1998) had argued that a sentence containing *each* can only be true if each individual in the set is associated with a subevent that can be differentiated from one another. This could either be with a differentiable time/location of each sub-event or different entities. It is unclear if this condition is satisfied. In the case of cumulative situations (such as figure 1.3), there are two entities (two goats), but three girls, who are all at the same place at the same time. Thus, the predicate may not be understood as applying to each individual member in the quantified set. Furthermore, each-cumulative may not be fully accepted due to the participant's different individual interpretations of what a sub-event is. As the results were very similar between the Dutch and English adults, it can be assumed that they will have the same explanations. So what Tunstall (1998) observed for English *each* should also apply for Dutch *each* 'elke'. The other explanation given was that it was possible that condition each-collective with independent verbs is not fully accepted because some speakers differ in what they consider sub-events, with some speakers assuming that the predicate has to hold for the whole group, which goes against the distributivity requirement of each. This explanation is not as obvious in this case of cumulative situations as the scene does not present a single whole group. However, with the difference of sizes of the groups (one group having two girls, and the other, one girl), some participants may have assumed the predicate has to hold for the two groups, going against the distributivity requirement of each.

Overall, these results were extremely similar to the results found with the use of *each* in collective situations, making it seem that people behave similarly in cumulative and collective readings. The largest difference though, being that in each-collective, the independent verb type had a higher acceptance rate in English than in Dutch, while we found in each-cumulative, Dutch had a higher acceptance rate than English.

4.4 Language

In order to see how accurate the claim was that *elke* is a weaker distributive marker than *each*, the results of both experiments were combined and placed into the same data set. Again, generalised mixed effect modelling was used to analyse this data. The full model can be seen in table A.3 in Appendix A. Random intercept was included for the participants. Here, we see a main effect with type *each* ($\beta = 1.286; z = 4.754; p < 0.01$), with the independent verb type ($\beta = -1.587; z = -6.528; p < 0.01$), as well as the interaction between the two ($\beta = 2.862; z = 9.249; p < 0.01$).

The post-hoc Tukey multiple comparisons showed that each-independent in English was accepted significantly less than each-independent in Dutch ($\beta = -0.838; z = -3.066; p < 0.05$). Despite this, the mixed model showed that there was no significant difference between English and Dutch.

Furthermore, with the combination of the data,

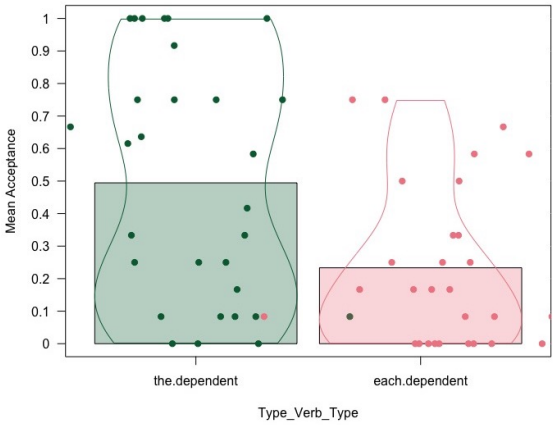

Figure 4.2: RDI plot of responses for participants who answered 'yes' less than 80% of the time (with the bars showing mean acceptance for dependent verbs for both subject types. Outlines show the variation in distribution and dots show individual mean responses.)

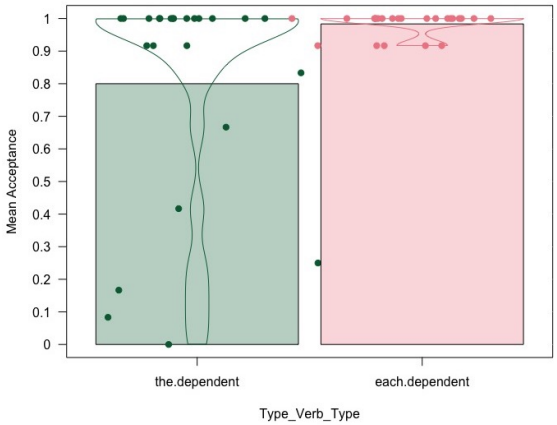

Figure 4.3: RDI plot of responses for participants who answered'yes' more than 80% of the time (with the bars showing mean acceptance for dependent verbs for both subject types. Outlines show the variation in distribution and dots show individual mean responses.)

we want to look at the population responses for the each-dependent cases. It was predicted that this condition would be disliked, however it was still accepted over 50% of the time. In figure 4.2 we see an RDI plot of the proportion of 'yes' responses when the participant answered yes less than 80% of the time in the each-dependent condition while in figure 4.3 we see an RDI plot of the proportion of 'yes' responses when the participant answered yes more than 80%. The graph also includes what these participants answered for the other condition that saw (the-dependent).

It is apparent that participants who were highly accepting of each-dependent were also more accepting of the-dependent. They had an acceptance rate of about 80% while participants who had low acceptance rates of condition each-dependent had an acceptance rate of about 50% in condition the-dependent. These participants appear to be sensitive to the verb type, so we would expect that these are the type of participants that accept each-independent and the-dependent. These participants were somehow not also preferring *the* with dependent verbs as there were a number of participants who rejected it.

In 4.2, we see a hourglass shape in the-dependent, showing a weak binomial distribution, with a mean of about 50%, while we do not see the same type of distribution for the-dependent in 4.3. There we see that the participants who were rather insensitive to the verb type, are also largely showing full acceptance of condition the-dependent. It could be that these participants are just very accepting, and tend to always say yes while those who were rather sensitive to verb type appear to have more of a range of responses to the-dependent.

It is interesting to note that of the 30 scores that were considered low, 15 of them were from the Dutch experiment, while the other 15 were from the English experiment. However of those, there were only 2 Dutch participants who never responded 'yes' at all for each-dependent, while there were 6 English participants who never responded yes. From the graph, we get a visual representative of how the participants answered.

5 Experiment 3

Each and *every* are both universal quantifiers. Sentences with *each* and *every* as their subject make a claim about all the members in the set, making them both distributive markers. Tunstall (1998) claimed that they are somewhat interchangeable, albeit differ in a number of ways. The paper presents a linguistic analysis of the differences be-

tween each and every with respect to distributivity, in which she claims that *each* is more strongly distributive than *every*. While *each* directs attention to the individual in *each*'s restrictor set, *every* emphasizes its restrictor set as a whole, or at least does not stress the separate individuals which comprise the set.

Furthermore, the English *every* has sometimes been claimed to be like the Dutch *elke*. The next section of this paper aims to investigate how acceptability in cumulative readings will differ with the different distributive marker *every* and if it will show similar results to the Dutch *elke*.

5.1 Participants

21 native English speaking participants took part in Experiment 3. They were all recruited from Prolific and paid for their participation. The distribution of the participants and their main features are reported in table 5.1.

Table 5.1: Number of English adult participants per participant group for Experiment 3.

Verb Type	# of participants	Mean age
Dependent	10	30.6
Independent	11	34.5

5.2 Design and procedure

The design and procedure of Experiment 3 were carried out in the same manner as Experiment 1 (and Experiment 2), apart from the fact that only one **subject** was tested (*every*). Therefore, participants only received 12 target items, and so 38 items in total. The participants received the same pictures as in Experiment 1 and Experiment 2.

5.3 Predictions

The two conditions in Experiment 3 are every-dependent and every-independent. Following the results from Experiments 1 and 2, the prediction is that *every* will have a higher acceptance rate with the independent verb type than the dependent verb type. Furthermore, we predict the results of Experiment 3 to be more similar to the Dutch *elke* than the English *each* due to claims that the Dutch *elke* is a weaker distributive marker than its English counterpart as well as *every* being a weaker distributive marker than English *each*.

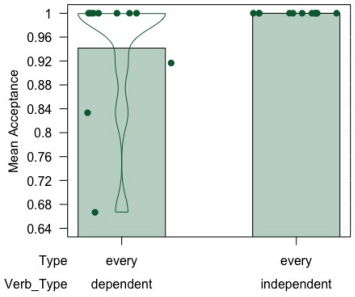

Figure 5.1: An RDI plot of the results from Experiment 3 (with the bars showing mean acceptance for dependent and independent verbs for *every* in English. Outlines show the variation in distribution and dots show individual mean responses.)

5.4 Results

The descriptive statistics can be seen in table 5.2 below:

Table 5.2: Proportion of 'yes' responses (with SD) in Experiment 3.

Verb Type	Results
Dependent	0.94 (0.24)
Independent	1 (0)

Acceptability in condition every-cumulative is extremely high with the verb type 'independent' being fully accepted by all of the participants and 'dependent' being accepted 94% of the time.

In figure 5.1 we see an RDI plot of the data. It is clear that *every* independent is fully accepted while *every* dependent showed a little more range in responses but still almost fully accepted. A somewhat similar pattern seen in the different verb types in *each* although the acceptance rate for both verb types in *every* is much higher than in *each*.

5.4.1 Comparing Every with the Dutch 'elke'

The results from this experiment were combined with the results from condition *each* in Experiment 1 in order to examine the claim that *every* was more like the Dutch *elke*. Generalised mixed effect modelling with R was used to analyse the data. For the model, a random intercept was included for the participants. The full model can be seen in Table A.4 in Appendix A. There was a main effect found with the quantifier type with *every* ($\beta = 3.270; z = 7.489; p < 0.01$).

The post-hoc Tukey multiple comparisons showed that every-dependent was accepted significantly more than elke-dependent ($\beta = 2.334; z = 4.619; p < 0.001$). But no significance in acceptance of *every* and *elke* with independent verbs.

5.5 Discussion

The results of *every* in cumulative situations were as predicted in that its acceptance rate showed more of a similarity with *elke*, than with *each*. *Every* was accepted much higher in both verb types than with *each* in both the English and Dutch experiment. However, there is still more of a similarity between *every* and *elke*, than with *each* as seen by their acceptance rates in Experiments 1 and 2. Furthermore, *every* appears to show a sensitivity to verb type, as seen by the other quantifiers. The results also showed a significant difference between the quantifier type *every* and *elke*, therefore, we cannot say that they are alike.

The results follows Tunstall (1998)'s claims that *each* is more distributive than *every*. The lower acceptance rate of *each* in cumulative situations as compared to *every*, show that people dislike using *each* much more than *every*. Furthermore, the full acceptance of *every* follows the claims that *every* only requires a partially distributive event structure, which explains why *every* was fully accepted in cumulative situations.

6 Experiment 4

A question that arose from the results of Experiments 1-3, is how are cumulative and collective interpretations different. It is evident that although the results from cumulative interpretations seen in this work are very similar to the results from collective interpretations seen in de Koster et al. (2020)'s work, they are not similar enough to show that they are the same. In this section, we investigate if collective interpretations are more basic that cumulative. We also examine if distributive interpretations are harder to process, as claimed by Sternefeld (1998) and Kratzer (2007). This is studied by testing participant's response time and acceptability judgment. A longer response time could suggest higher complexity due to more time needed to process the situation.

6.1 Participants

48 participants took part in Experiment 4. However, as 12 participants said that English was not their native language, their results were removed from the data set. The remaining 36 native English-speaking participants had a mean age of 33.25. They were all recruited from Prolific and paid for their participation.

6.2 Design and procedure

Experiment 4 is a self-paced reading task that also tests acceptability judgment. This experiment was a Latin square design with items being distributed over 4 different lists that were presented randomly. Each list contained 3 practice items, 4 test items, and 2 fillers. Participants were shown a short passage, and were asked to rate the passage by choosing a number between 1 and 5. With 1 being the worst, and 5 being the best. The following is an example of one of the passages:

> Three American and three Japanese astronauts met up to upgrade a section of the international space station. The American astronauts- Emma, John, and Mark brought three control panels. The Japanese astronauts- Mai, Shoko, and Aya replaced seven of the outdated monitors. Mai and Shoko replaced five monitors while Aya replaced two. Everyone was pleased with the updated features.

After rating, a comprehension sentence appeared on a new page based on the passage. The four test sentences for this story were as follows:

A The American astronauts brought one control panel and everyone was pleased.

B Each American astronaut brought one control panel and everyone was pleased.

C The Japanese astronauts replaced seven monitors and everyone was pleased.

D Each Japanese astronaut replaced seven monitors and everyone was pleased.

All sentences began with the definite plural *the* or the quantifier *each* and the structure of the four test sentences were the same for all four stories. The participants were therefore asked to judge whether the sentence was true or false. We label each type of test item as A, B, C, or D as seen above.

6.3 Prediction

Due to the claims seen in previous literature, the prediction is that participants take less time to judge the collective sentences, while distributive sentences take the longest. From the results seen in Experiments 1-3, the cumulative sentences will have more similar reaction times to the collective sentences than the distributive sentences. However, we expect it to be slower. This follows claims from (Sternefeld, 1998) that cumulative interpretations are more complex than collective interpretations, but less complex than distributive interpretations.

6.4 Results

After removing outliers (greater than 2.5 SD away from the mean), the descriptive statistics can be seen in table 6.1 below:

Table 6.1: Proportion of 'yes' responses (with SD) and response time (in ms).

Sentence	Results	Response time (ms)
A	0.31 (0.48)	5763.11 (3491.00)
B	0.61 (0.50)	8641.11 (7864.25)
C	0.81 (0.40)	7759.38 (6743.87)
D	0.22 (0.42)	5239.12 (3663.05)

Sentence type 'C' had the highest acceptance rate while sentence type 'D' had the lowest acceptance rate. Meanwhile, sentence types 'A' and 'D' had the fasted reactions times while 'B' had the slowest.

Using generalised mixed effect logistic modelling with respect to acceptance, the data was analysed to check for a significant difference. The full model can be seen in table A.5 Appendix A. A random intercept was included for the participants. A main effect was found for sentence type 'B', and 'C'.

With respect to response time, the estimate of the standard deviations of the random effects for the intercept is 4154 ms. The fixed-effects coefficient β for the intercept is 5953.06ms ($p < 0.001$). The model also showed a main effect with sentence 'B' ($\beta = 2933.76; p = 0.003$), and sentence 'D' ($\beta = 1998.29; p = 0.0046$). The full model can be found in table A.6 in Appendix A.

The post hoc Tukey test showed that sentence type 'B' was significantly slower than 'A' ($\beta = 2933.8; z = 2.947; p = 0.017$) and that sentence type 'D' was significantly faster than type 'B' ($\beta = -3455.7; z = -3.438; p = 0.003$).

6.5 Discussion

Sentence 'B' was predicted to have the slowest reaction time as it contains the distributive marker *each* and suggesting a distributive reading. The results followed our prediction. Despite sentence 'D' also having the distributive marker, it was predicted to be treated as cumulative due to its cumulative suggestion. And as predicted, its reaction time was faster than sentence 'B'. Sentence type 'A' was a predicted to have a relatively fast reaction time, due to suggesting a collective interpretation. Although it was fast, it did not have the fastest reaction time. Sentence 'C' was also as predicted in that it was faster than 'B' but slower than 'A' as it was also a sentence that was treated as cumulative.

With this, the results appear to agree with claims from Sternefeld (1998) and Kratzer (2007) in that distributive interpretations are hardest to process (determined by its slow response time). Furthermore, the the most cumulative sentences 'C' showed response times higher than the collective 'A' and the distributive 'B', which also suggests that Sternefeld' claims was right. Sternefeld (1998) argued that distributive interpretations were the most complex while collective interpretations were the most basic.

7 Conclusion

In this paper, we investigated a number of questions regarding distributive markers in cumulative readings. The first being how verb type affected the acceptability of cumulative readings with distributive markers, and how this compared to collective readings. The second being whether there is a significant difference in sensitivity to verb type between the Dutch and English language. Our results followed our predictions and showed that cumulative readings appear to have a similar sensitivity to verb type, having a higher acceptance with independent verbs, than dependent verbs. Our results also showed no significant difference in acceptance between the Dutch distributive marker *elke* to its English counterpart.

Furthermore, we investigated the claim that *every* was similar to the Dutch *elke* due to them both being weaker distributive markers. The results showed that our prediction was true, with *every* having a higher acceptance, that is closer to the Dutch *elke*.

Lastly, we looked at how cumulative and collective interpretations might be different by comparing readings times. The results showed that distributive interpretations are the most complex, and also seem to suggest that cumulative interpretations are more complex than collective interpretations, alluded through response time. However, to get more conclusive results about the the complexity of cumulative and collective interpretations, and if collective is really more basic, more research needs to be carried out on a larger scale.

References

de Koster, A., Dotlačil, J., and Spenader, J. (2017). Children's understanding of distributivity and adjectives of comparison. In *Proceedings of the Annual Boston University Conference on Language Development*, volume 1, pages 373–386. [1, 5, 6]

de Koster, A., Spenader, J., Dotlačil, J., and Hendriks, P. (2020). A multiple cue explanation of collective interpretations with each. In *Proceedings of the 41st Annual Boston University Conference on Language Development*, pages 252–265. [2, 3, 4, 6, 7, 10]

Frazier, L., Pacht, J. M., and Rayner, K. (1999). Taking on semantic commitments, ii: collective versus distributive readings. *Cognition*, 70(1):87–104. [1]

Kaup, B., Kelter, S., and Habel, C. (2002). Representing referents of plural expressions and resolving plural anaphors. *Language and Cognitive Processes*, 17(4):405–450. [1]

Kratzer, A. (2007). On the plurality of verbs. *Event structures in linguistic form and interpretation*, 269:300. [2, 4, 5, 10, 11]

Lee, T. H.-t. (1997). Scope and distributivity in child mandarin. In *Proceedings of the 28th Child Language Research Forum. Stanford, CA: Stanford Linguistics Association, Centre for the Study of Language and Information*, pages 173–182. [2]

Musolino, J. (2009). The logical syntax of number words: Theory, acquisition and processing. *Cognition*, 111(1):24–45. [2]

Pagliarini, E., Fiorin, G., and Dotlačil, J. (2012). The acquisition of distributivity in pluralities. In *Proceedings of the Annual Boston University Conference on Language Development*, volume 2, pages 387–399. [1]

Rouweler, L. and Hollebrandse, B. (2015). Distributive, collective and "everything" in between: Interpretation of universal quantifiers in child and adult language. *Linguistics in the Netherlands*, 32(1):130–141. [1, 3, 6]

Sternefeld, W. (1998). Reciprocity and cumulative predication. *Natural language semantics*, 6(3):303–337. [2, 10, 11]

Tunstall, S. L. (1998). *The interpretation of quantifiers: semantics & processing*. PhD thesis, University of Massachusetts at Amherst. [7, 8, 10]

Ussery, C. (1998). Processing plural dps: Collective, cumulative, and distributive interpretations. *Amherst.: University of Massachusetts*. [1, 4, 5]

A Appendix

Table A.1: Experiment 1: Model = Response ~ Type * Verb_Type + (1 |Subject_Nr)

Predictor	Estimate	Standard Error	z-value	p-value
(Intercept)	1.3266	0.6157	2.154	0.0312 *
Typethe	0.2590	0.2004	1.293	0.1962
Verb_Typeindependent	2.8324	0.9253	3.061	0.0022 **
Typethe:Verb_Typeindependent	-4.1478	0.4353	-9.528	<2e-16 ***

Table A.2: Experiment 2: Model = Response ~ Type * Verb_Type + (1 |Subject_Nr)

Predictor	Estimate	Standard Error	z-value	p-value
(Intercept)	1.1066	0.6543	1.691	0.09076 .
Typeeach	-0.6129	0.2151	-2.849	0.00439 **
Verb_Typeindependent	-0.7461	0.9120	-0.818	0.41328
Typeeach:Verb_Typeindependent	3.3595	0.4026	8.344	<2e-16 ***

Table A.3: Language: Model = Response ~ Type * Verb_Type * Language + (1 |Subject_Nr)

Predictor	Estimate	Standard Error	z-value	p-value
(Intercept)	1.2856	0.2704	4.754	1.99e-06 ***
Typeeach	-0.2073	0.1791	-1.158	0.2469
Verb_Typeindependent	-1.5868	0.2431	-6.528	6.67e-11 ***
LanguageEnglish	-0.2830	0.2070	-1.367	0.1716
Typeeach:Verb_Typeindependent	2.8623	0.3095	9.249	<2e-16 ***
Typeeach:LanguageEnglish	-0.2945	0.2656	-1.109	0.2675
Verb_Typeindependent:LanguageEnglish	0.5845	0.3292	1.775	0.0758 .
Typeeach:Verb_Typeindependent:LanguageEnglish	-0.8456	0.4146	-2.039	0.0414 *

Table A.4: Experiment 3: Model = Response ~ * Verb_Type + (1 |Subject_Nr)

Predictor	Estimate	Standard Error	z-value	p-value
(Intercept)	1.3079	0.4029	3.247	0.00117**
Type [Every]	3.2702	0.4367	7.489	<0.0001***
Verb_Type [independent	0.1437	0.4762	0.302	0.76283

Table A.5: Experiment 4 (acceptance): Model = Response ~ Type + (1 |Subject_Nr)

Predictor	Estimate	Standard Error	z-vale	p-value
(Intercept)	-0.8446	0.4002	-2.110	0.034819 *
type [scene.b]	1.4097	0.5469	2.578	0.009946 **
type [scene.c]	2.4888	0.6467	3.849	0.000119 ***
type [scene.d]	-0.5940	0.5760	-1.031	0.302347

Table A.6: Experiment 4 (Response time): Model = Time ~ Type + (1 |Subject_Nr)

Predictor	Estimate	Standard Error	p-value
(Intercept)	5953.06	988.1	<0.001
type [scene.b]	2933.76	995.5	0.003
type [scene.c]	1998.29	999.8	0.046
type [scene.d]	-521.97	999.8	0.602

B Appendix

B.1 Experiment 4

Story 1

Three boys and three girls need to prepare for their school play. Alice, Amanda, and Angela designed three costumes. Ben and Bob built three props and Barry built two props. Everyone was very impressed.

Test items

A The girls designed one costume and everyone was very impressed.

B Each girl designed one costume and everyone was very impressed.

C The boys built five props and everyone was very impressed.

D Each boy built five props and everyone was very impressed.

Story 2

Three American and three Japanese astronauts met up to upgrade a section of the international space station. The American astronauts- Emma, John, and Mark brought three control panels. The Japanese astronauts- Mai, Shoko, and Aya replaced seven of the outdated monitors. Mai and Shoko replaced five monitors while Aya replaced two. Everyone was pleased with the updated features.

Test items

A The American astronauts brought one control panel and everyone was pleased.

B Each American astronaut brought one control panel and everyone was pleased.

C The Japanese astronauts replaced seven monitors and everyone was pleased.

D Each Japanese astronaut replaced seven monitors and everyone was pleased.

Story 3

The country Blueland is known for its great lake. The North and South are divided by this great lake. Both the North and South have three towns each. The mayors of the Northern towns- Ned, Norman, and Natalie built three schools in the North. The mayors of the Southern towns- Sarah, Sam, and Sandy hosted five charity events last year. Sarah and Sam hosted three events while Sandy hosted two. Everyone in Blueland is very happy.

Test items

A The mayors in the North built one school and everyone was very happy.

B Each mayor in the North built one school and everyone was very happy.

C The mayors in the South hosted five charity events last year and everyone was very happy.

D Each mayor in the South hosted five charity events last year and everyone was very happy.

Story 4

The local high school is hosting a swim meet. The school asked some of the mums to bring refreshments for the swimmers. Sandra, Karen, and Debbie provided three cases of sports drinks. Matt, Henry, and Sean participated and won five races. Matt and Henry won three races and Sean won two. Everyone enjoyed themselves.

Test items

A The mums brought one case of sports drinks and everyone enjoyed themselves.

B Each mum brought one case of sports drinks and everyone enjoyed themselves.

C The swimmers won 5 races and everyone enjoyed themselves.

D Each swimmer won 5 races and everyone enjoyed themselves.