The role of pragmatics in processing of distributivity

Bachelor's Project Thesis

Feline van Aagten, f.e.van.aagten@student.rug.nl,
Supervisors: Dr J.K. Spenader

Abstract: A self-paced reading experiment was conducted to further examine the preference of readers for a collective interpretation compared to a distributive interpretation in ambiguous plural sentences. Specifically the role of pragmatics was examined. Participants were shown items with no disambiguator, the distributive *each* and the collective *all*. All items could be combined with either disambiguator, though the items were forced to be distributive due to pragmatics.Our expecation was that if pragmatics has a strong effect on the reading time of a sentence, it would overrule the influence of the collective preference on the reading time. This difference would appear in sentences with the disambiguator *all* versus sentences without a disambiguator or the disambiguator *each*, with the former being slower to read. We cannot argue that our hypothesis is correct, as we have not found any significant results. However, our findings do slightly seem to hint towards a difference between *each* versus *all* and without disambiguator. There are aspects of the experiment that could be improved, which could have caused the lack of effect found.

1 Introduction

The meaning of plural expressions can often be ambiguous. Take the following sentence for example.

(1) The three girls built a sandcastle.

The sentence can be interpreted in two different ways. The collective interpretation would be that the three girls, together, built one sandcastle. The distributive interpretation would be that there were as many sandcastles built as there are girls, namely three. Each of the girls made their own sandcastle in this case. It is not inherently clear which interpretation is meant for this sentence and this can cause miscommunication. People, however, do not experience plural sentences as a big hindrance in day-to-day life. So, how do people distinguish whether an ambiguous plural sentence is collective or distributive? In previous studies by Frazier et al. (1999) it has been found that readers prefer a collective reading over a distributive reading. What has not yet been researched is the role of pragmatics on this preference. In a self-paced reading experiment, we have compared the processing cost of three variations of a pragmatically distributive sentence. These sentences contain events that are lexically very unlikely to be collective. An example of the test items is shown at (2a,b,c).

(2a) The managers drank a celebratory cocktail after landing the deal.

(2b) The managers **each** drank a celebratory cocktail after landing the deal.

(2c) The managers **all** drank a celebratory cocktail after landing the deal.

Sentence (2a) does not contain a quantifier, (2b) has a distributive quantifier *each*, and (2c) has a collective quantifier *all*. Still, pragmatics forces a distributive reading of these sentences as it is very unlikely that multiple managers, together, drank from one and the same cocktail. However, how would the collective preference manifest in sentences that are pragmatically distributive? Is the collective bias so strong that it overcomes this difference, or does such a strongly pragmatically distributive sentence weaken it or even make it disappear? We hypothesise that if the pragmatics of a sentence has a strong effect, we would expect to see slower reading times for the sentences with a

collective disambiguator *all* (2c) as compared to sentences with no disambiguator (2a) or the distributive disambiguator *each* (2b). Our mixed effects model did not result in any significant findings to substantiate our claims.

1.1 Background

In 1999, a study done by Frazier et al. looked into the mental processing of these kinds of sentences. They used sentences that had a plural noun phrase as subject followed by a verb and object. All of the sentences could be combined with a collective and distributive disambiguator to force the interpretation of the sentence. In their study they used *each* as the distributive disambiguator and *together* as the collective disambiguator. An example of the sentences they used:

(3a) Lou and Deborah danced one tango **each** <u>at our class</u> reunion.

(3b) Lou and Deborah danced one tango **together** <u>at our class</u> reunion.

(3c) Lou and Deborah **each** danced one tango <u>at our class</u> reunion.

(3d) Lou and Deborah **together** danced one tango <u>at our class</u> reunion.

In sentence (3a) and (3b) the disambiguator is placed after the verb and object. While in sentence (3c) and (3d) the disambiguator is placed in front of the verb and object. Frazier et al. (1999) based their hypothesis on the Minimal Semantic Commitment (MSC) hypothesis they had demonstrated in previous work Frazier and Rayner (1990). The MSC hypothesis has two components; the grammatical ambiguity hypothesis and the vagueness hypothesis. The grammatical ambiguity hypothesis states that if a reader encounters ambiguous representations (for example a verb that can be both collective or distributive) they commit to one reading of that representation, until evidence to the contrary is found. The vagueness hypothesis states that a when reader reads an underspecified representation (i.e. there is not enough information to determine what the representation means) they will await disambiguation of that underspecified representation until they are more certain of its meaning. The hypothesis of Frazier et al. (1999) was that the reader would wait with committing to either interpretation until the sentence was completely disambiguated, which follows the vagueness hypothesis. After running the experiment they found that postverbal distributive sentences, such as sentence (3a), were more difficult to understand for readers. The underlined region of the sentence cost more processing time compared to the postverbal collective sentences. They did not find a difference between the reading time of the preverbal sentences (3c) and (3d). This was contrary to what they hypothesised and they concluded that these sentences adhered to the ambiguity hypothesis and not the vagueness hypothesis. Readers commit to one interpretation, even if the sentence is not fully disambiguated yet. Frazier et al. (1999) also concluded that collective readings are preferred over distributive readings.

In a paper by Dotlačil and Brasoveanu (2021), the collective versus distributive processing of plural sentences is being investigated further. They examine why the collective reading of plural sentences is preferred over a distributive reading. They discuss three potential reasons for this preference. First, the structural considerations; the reader picks the simplest version of interpreting the sentence. A distributive reading requires a D operator, such as *each*, which is more complex than the collective reading. The second option is that the preference for collectivity is based on interpretational complexity. The reader will choose the interpretation that has the fewest events and entities. Consider the example sentence (1) above. The interpretation would be that the three girls made one sandcastle all together. The last option Dotlačil and Brasoveanu (2021) suggest is that the distributive disambiguator *each* is the cause of the difference. The preverbal and postverbal *each* are distinct items. He explains that the word *each* differs syntactically and semantically based on the position in the sentence and is therefore more complex, which could be the cause of the collective preference. By changing the distributive disambiguator from *each* to *individually*, which does not syntactically differ based on placement in a sentence, they want to investigate if the wording of the disambiguator has influence on the preference for collectivity.

Dotlačil and Brasoveanu (2021) performed two experiments. The first experiment was based on the

experiment of Frazier et al. (1999), as mentioned they changed the distributive disambiguator from *each* to *individually*. By doing so they were able to test if the collective preference was due to the complexity of the word *each*. They found that the switching of disambiguator did not make a difference in the results, and therefore they eliminated the complexity of the word *each* as a cause for the collective preference. In their second experiment Dotlačil and Brasoveanu (2021) examined the two alternative reasons for a collective preference, namely the structural considerations and the interpretational complexity consideration. They did this by using words that lexically have distributive properties, for example *crying*. The act of crying can be done together, but is essentially one individual with tears dripping from one's eyes. This verb is distributive without needing the D operator. On the other hand, Dotlačil and Brasoveanu (2021) wanted to examine the interpretational complexity, whether or not readers prefer the collective reading because they favor a the simplest explanation in terms of events. These sentences do need a D operator, which is often paraphrased as *each*. An example would be: *The boys carried one suitcase*, which will then be interpreted as *The boys each carried one suitcase.* This would come down to one suitcase per boy.

This first experiment was done by means of a self-paced reading task. The items used were four variations of the same sentence, see (4a,b,c,d).

(4a) The sisters **individually** painted a masterpiece when they were at school.
(EARLY, INDIVIDUALLY)

(4b) The sisters **together** painted a masterpiece when they were at school.
(EARLY, TOGETHER)

(4c) The sisters painted a masterpiece **individually** when they were at school.
(LATE, INDIVIDUALLY)

(4d) The sisters painted a masterpiece **together** when they were at school.
(LATE, TOGETHER)

Dotlačil and Brasoveanu (2021) predict to see two things: if the preference for the collective reading is based on structural considerations they predict to see a preference for collective reading only when a direct object is present, and not when the correct reading of the sentence can be derived from lexical knowledge (i.e. smile). This means they expect a three-way interaction between the type and position of the adverb and the presence of a direct object. On the other hand, if the preference is based on interpretational complexity, the presence of a direct object should not matter, and they would expect to see a preference for the collective reading overall, so even in sentences where the reading can be derived from lexical knowledge. In this case they would expect a two-way interaction between the type and position of the adverb.

This second experiment was also done by means of a self-paced reading task. They combined their first experiment with an extra factor, namely the object of a sentence. Either it was present or it was not. This resulted in eight conditions. An example are (5a,b,c,d,e,f,g,h).

(5a) The students **individually** <u>read a book</u> when they were at the library.
(EARLY, INDIVIDUALLY, OBJECT PRESENT)

(5b) The students **collectively** <u>read a book</u> when they were at the library.
(EARLY, COLLECTIVELY, OBJECT PRESENT)

(5c) The students <u>read a book</u> **individually** when they were at the library.
(LATE, INDIVIDUALLY, OBJECT PRESENT)

(5d) The students <u>read a book</u> **collectively** when they were at the library.
(LATE, COLLECTIVELY, OBJECT PRESENT)

(5e) The students **individually** <u>read</u> when they were at the library.
(EARLY, INDIVIDUALLY, NO OBJECT)

(5f) The students **collectively** <u>read</u> when they were at the library.
(EARLY, COLLECTIVELY, NO OBJECT)

(5g) The students <u>read</u> **individually** when they were at the library.
(LATE, INDIVIDUALLY, NO OBJECT)

(5h) The students read **collectively** when they were at the library.
(LATE, COLLECTIVELY, NO OBJECT)

A third experiment that Dotlačil and Brasoveanu performed was on the influence of pragmatics on the interpretation of the reader. This experiment was not published in their paper however, because they had too few participants, which made the outcome unreliable. We have redone the experiment with more subjects for this paper. In certain sentences the interpretation of a sentence is clearly based on the meaning of that sentence. If we take the following sentence as a example:

(6a) The kids rode a unicycle as part of their final performance.

(6b) The kids **each** rode a unicycle as part of their final performance.

(6c) The kids **all** rode a unicycle as part of their final performance.

Looking at sentence (6a), there is not a distributive or collective quantifier present, yet it is clear what is meant by this sentence because of pragmatics. The principle of a unicycle is clear. It is a bike with one wheel and one saddle, so naturally only one person can ride it at once. When reading this sentence, one will visualise multiple kids all riding their own unicycle. The next two sentences (6b) and (6c) do have a quantifier, respectively *each* (distributive) and *all* (collective). The quantifiers force a certain reading, while the eventual interpretation of the sentence remains the same. We know that the collective reading is preferred. What will happen when we take plural sentences that can be combined with both collective and distributive quantifiers, but are pragmatically distributive. Do pragmatics affect the influence of the collective preference? In the case pragmatics has a strong effect on the reading of the sentence, we expect the effect of the collective preference to weaken or disappear. This would show itself in shorter reading times of sentences with no quantifier or the distributive quantifier each versus sentences with the collective quantifier. If this is not the case, we would expect to see differences between the sentences with a distributive quantifier versus the collective and without a quantifier following Brooks and Braine (1996).

Our hypothesis also follows the findings of the study by Frazier and Rayner (1990) and Frazier et al. (1999) previously mentioned. They examined whether the vagueness hypothesis or the ambiguity hypothesis was adhered to. In their study, Frazier et al. (1999) found that the latter was the case. In sentences where it is unclear yet what the exact interpretation of the sentence should be, they stick to one interpretation until proven otherwise. This falls in line with our hypothesis. Following the ambiguity hypothesis, readers will read the sentence, choose the suggested interpretation based on the disambiguator. After reading the verb and object they will see that their interpretation was chosen prematurely and change their interpretation of the sentence, resulting in a longer reading time. Combined with a natural preference for a collective reading, our pragmatically distributive sentences will potentially counteract their fast reading times.

2 Experiment

2.1 Participants

51 self-reported native English speakers took part. These participants were found through the participant recruitment site Prolific. The age of the participants was between 18 and 71 with a mean of 38. Participants gave consent and received a short explanation on how to perform the experiment. Participants got a small reward of $1.38 for the experiment that took around 10 minutes to complete.

2.2 Design and procedure

An example of the test items in all three conditions used, is shown in (7a,b,c). The three conditions were: (7a) the sentence with no marking, (7b) the sentence with the distributive marking *each* immediately after the subject or (7c) the collective marker *all* also directly after the subject.

(7a) The women read a newspaper just before going to work.

(7b) The women **each** read a newspaper just before going to work.

(7c) The women **all** read a newspaper just before going to work.

The test sentences all contain verbs that can be combined with both a distributive as a collective quantifier. Even though these markers suggest a distributive or collective reading, lexical knowledge tells us the act is done individually. Taking (7a) as an example: one will find that the sentence does not specify how many newspapers are present. Either the women are, together, reading just one newspaper or each of the women is reading her own newspaper. Although it is not explicitly stated, the most sensible situation is the distributive one, where there are as many newspapers as there are women. All the test items are chosen in this particular way. They are pragmatically distributive, yet the verb can also be combined with a collective quantifier without changing the (pragmatically distributive) meaning of the sentence.

In each session 18 test items were presented, which comprised six test items in each of the three conditions plus 30 filler sentences. All were presented in a random order. The distribution of the sentences was based on latin square design principle. The length of test items were between 11 and 18 words. All test sentences started with a few words (region 1), then a predicate (predicate region) which consists out of the the disambiguator plus the verb plus the object. Region 3 are the following three words, unless the third word is the last word of the sentence, in that case region 3 consists of only two words. All remaining words are part of the wrap up. Following Dotlačil and Brasoveanu (2021) we expect to see the effects of the different disambiguators in region 3, which is the reason the test items where longer sentences. In example sentence (8) the regions are seperated with dashes and region 3, our region of interest, is underlined.

(8) Last night, the children/ **all** ate a cookie / after dinner without asking their parents.

The experiment was set up as a self-paced reading task. In this task participants have to read a sentence word-by-word. The participants can press the space button once they have read the word and the next word will then appear until the end of the sentence is reached. The self-paced reading task was cumulative and linear, in the sense that each revealed words was visible until the sentences was completed. Participants knew where to expect the subsequent word as there was a placeholder dash on that exact spot. This approach was chosen because a linear self-paced reading task is more natural for the reader than a centered approach, as it is more like normal reading according to Jegerski (2014). The idea of the self-paced reading task is to measure the amount of time it takes the participant to read a segment, in this case a word. The basic assumption of the self-paced reading is that a higher the reading time implies a higher the processing cost (Just and Carpenter, 1980). With this strategy we can compare different sentences and analyse which types of sentences are more costly and therefore probably less expected by the reader.
After reading a complete sentence a comprehension question is presented and the participant should answer the yes-or-no question accordingly. For example the comprehension question for sentence (7a,b,c) was: "Did the women read a book before going to work?". With these comprehension questions we want to check whether or not the participants really read the sentences.

The experiment was set up as well as hosted on PCIbex by Zehr and Schwarz (2018), a javascript extention of IBEX farms, an online platform to create experiments, by Drummond (2007). Participants were able to partake in the experiment with a PC or laptop.

3 Results

3.1 Preprocessing

After the data was collected it was processed with the goal to remove obvious outliers. No participants were removed based on their reading times, but one participant answered the comprehension question correctly less than 85% of times. This participant was deemed to have not to have participated actively enough in the experiment and their data was removed from our results.

3.2 Statistical analysis

When analyzing our data we will focus on a region of the sentence called region 3. Region 3 is the region directly after the predicate consisting out of

three words, unless the third word is the last word of the sentence. Following Dotlačil and Brasoveanu (2021) we expect to see the effects of the different disambiguators (no, *all*, *each*) in this region. In looking at the data we found that the reaction times were heavily skewed to the right. We therefore used the logarithm of the reaction times to make sure the distribution of the reaction times is closer to a normal distribution. After taking the logarithm of the reaction times the data was normally distributed, the figures can be seen in the appendix (figures B.1 and B.2).

After we calculated the logarithm of the reaction times for all regions in the sentence we then calculated their means and standard deviations. We have plotted these means by sentence version, based on the disambiguator, in figures 3.2 (without errorbars) and 3.1 (with errorbars based on the sd).

In the plots of the mean reaction times (figures 3.2 and 3.1) it can be seen that the mean reaction time for the predicate region for the distributive disambiguator *each* is higher than those of the sentences with no disambiguator or the disambiguator *all*. It should be noted that we removed the disambiguators from the predicate region to get a more realistic comparison between all three versions of the same sentence. The version without a disambiguator would have one word less, influencing the mean. Something else that stands out is that the errorbars for the *each* condition in the predicate region are very large, implying a large standard deviation.

For region 3, the reading time is slightly lower for the disambiguator *each* than those of no disambiguator and the disambiguator *all*. There is also a difference in mean reading time between sentences with the disambiguator *all* and sentences without a disambiguator, but this seems very small. An interesting observation here is that the standard deviation for all three conditions is around the same value.

To analyse our data we used the statistical R Core Team (2020) software. We fitted a linear mixed effects model from the R package "lme4" by Bates, Mächler, Bolker, and Walker (2015). We use a mixed effects model because our subjects gave multiple responses. Our data consists of multiple measurements per subject, this makes the data dependent, for which we have to account. The dependent variable we used was the logarithm of the reaction times on region 3. As we are interested in the effect of the type of disambiguator we chose the version of the sentence as our fixed effect. We wanted to see if the interpretation of plural sentences of a reader is influenced by pragmatics or not. In previous studies by Frazier et al. (1999) it was found that there was a preference for a collective reading. By use of sentences that suggest a distributive interpretation by means of pragmatics. We wanted to see what the influence was on reading times. If we would see comparatively shorter reading times between no disambiguator and *each* versus *all*. This effect would confirm our hypothesis that pragmatics does have an influence. To examine this, we took the three versions as a fixed effect.

We took the subject, through the use of an identifier of the participant, and items, through an identifier of the sentence, as random effects. Both of these factors will contain unintentional individual variation that the model should keep into consideration. These random effect coincided with the effects proposed by Dotlačil and Brasoveanu in their third unpublished experiment. After some testing of different models we indeed found that the best model includes the subject and the item as random effects. Some other random effects that were considered, but seemed not to improve the quality of the model, were the age and gender of a participant.

To select the best fitting model we looked at the AIC score, Akaike (1974), on all combinations of effects. The AIC-score is an estimate of the amount of information lost in fitting the model, thus the lower the AIC-score of a model, the better it performs. The AIC-score of the model we selected was 859.58.

If we look in the appendix at figure B.3 we can see the QQ-plot of our best model. We can see that there are quite a few observations that fall in the higher range of sample quantiles (>1). We used the model to trim our data by removing outliers one standard deviation above and below the residual means. This pruned 24 samples, or about 3% of observations. We then fit a new model using the same effects on this new data. The residuals of this new model are shown in the appendix in figure B.4. We can see that the fit is already better. After removing the outliers the AIC of our model was 397.1, which is big improvement compared to the previous AIC of 859.58.

The model coefficients are shown in table 3.1. In

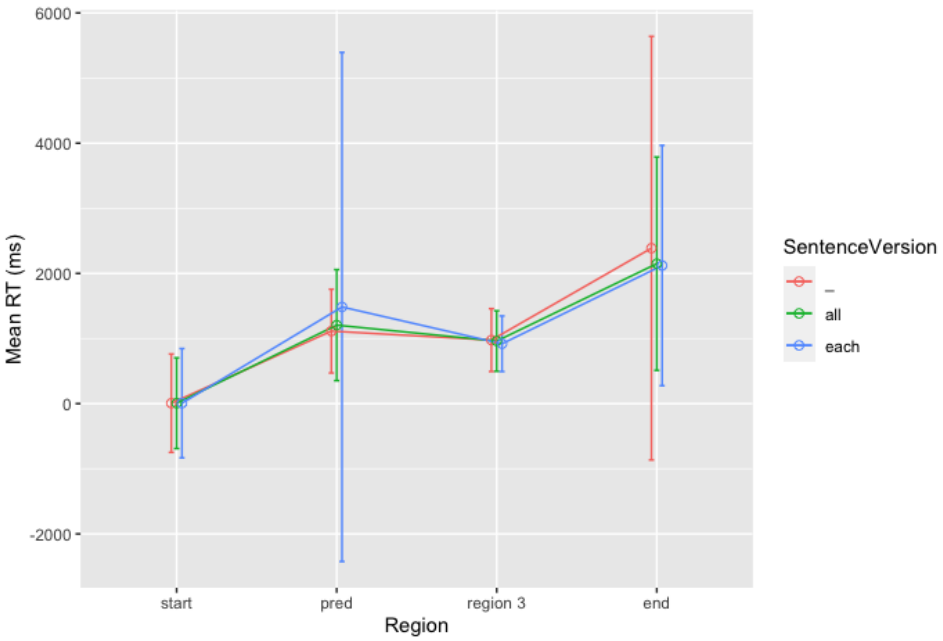

Figure 3.1: Mean RT with error bars

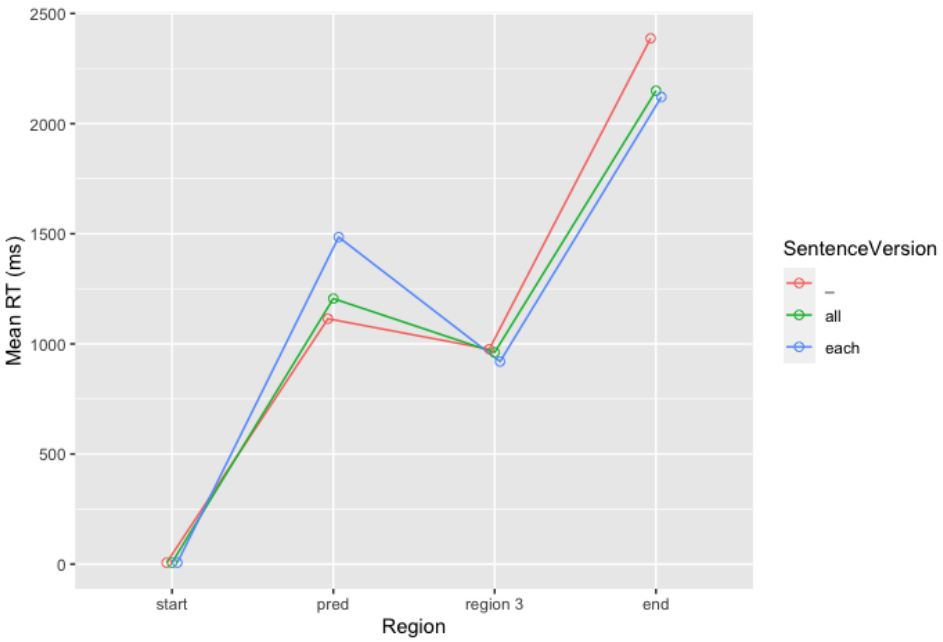

Figure 3.2: Mean reading time

Condition	Estimate	sd	t-value	p-value
-:all	-0.019	0.0229	-0.840	0.4010
-:each	-0.038	0.0229	-1.668	0.0954
all:each	-0.018	0.0229	-0.829	0.4072

Table 3.1: Model coefficients of sentence versions

this table the first collumn is an effect pair, with the effect shown being the effect of the second item of the pair as compared to the first item. It can be seen that reading sentences with the disambiguator *each* is faster than reading sentences without a disambiguator or the disambiguator *each*, and that sentences with the disambiguator *all* are read slightly faster than sentences without a disambiguator.

The strongest effect can be observed in the sentences with the disambiguator *each*, while the effect between the item without disambiguator and the item with the disambiguator *all* is quite a bit smaller. Using the t-values of our model we approximated the p-values of the model using a normal distribution. The R-snippet to do so is given by the following code snippet:

```
coefs <- data.frame(coef(summary(
    model_without_outliers)))

# use normal distribution to approximate p-
    value
coefs$p.z <- 2 * (1 - pnorm(abs(coefs$t.value)
    ))
```

Looking at the t-values and corresponding p-values however shows that these effects are not significant ($p > 0.05$). It should however be stated that the p-value for the disambiguator *each* was $p = 0.095$, while still not significant this effect is the strongest.

4 Discussion

We expected to see a difference between the *all* condition in comparison to the *each* and the items without a disambiguator, where the sentences with the *all* condition would be read comparatively slower than the other versions of the sentence. This would mean that the preference for collective readings was overruled by pragmatics Our results show us that in sentences with the disambiguator *each* region 3 is read slightly faster than in sentences with no disambiguator or the disambiguator *all*. The effect was very small between *each* and *all* however and none of the effects found were significant. We do not believe this contradicts previous studies though. After concluding the experiment and analysing the results we found certain areas that could be improved.

For starters, the word *all* was chosen as the collective disambiguator. Other studies, for example Dotlačil and Brasoveanu (2021), used the collective disambiguator *together*. As stated by Dotlačil and Brasoveanu (2021), the word *together* is a stronger disambiguator than the word *all*, and therefore might have caused a stronger effect on the reading time. As the difference between no disambiguator and the word *each* was closest to being significant, changing the collective disambiguator to *together* might yield interesting results.

Another factor to consider is that we chose our test items without conducting another study on these items. We could have examined our items by means of a pilot or preference study to make sure the test sentences were indeed as pragmatically distributive as we wanted. It could be the case that in the sentences we chose the effect would have been expressed less because they were, for example, too elementary. It could also be the case that our predicates were not distributive enough, for example the predicate "ate an apple" could be less distributive than the predicate "drank an espresso", which in turn could make the effect of the distributor less clear.

Another factor that played a role in our experiment is that we used cumulative reading, showing each new word with a press of the space bar until the whole sentence was visible, in our self-paced reading tasks. Cumulative reading might not be the best way to analyse reading times of participants. According to Ferreira and Henderson (1990) and Just et al. (1982), readers can apply a reading strategy on these sentences. Instead of reading word-by-word, as we intended, readers can click through the sentence very quickly and read the sentence as a whole instead. This cumulative way of displaying words could have had a confounding effect on our data as well. Instead, the experiment could be repeated, only this time set up non-cumulative, where only the new word is visible and the previous words are hidden. Another direction the study could be

taken in is changing the length of the test sentences used. In some cases, region 3 of a test sentence was only two words, leaving very little room for wrap-up effects to be measured clearly in the end region. This could easily be offset by making sure all sentences are long enough for the end region to contain more words.

In this experiment we mainly focused on region 3, however the definition of this region could be changed. We could make it longer, up to the end of the sentence so we capture more of the wrap-up effect of reading the sentence, or shorter and check word by word the difference in reading time after the different disambiguators.

When participants answered their comprehension questions they did not get any feedback on their answer. This could have influenced the way participants approached reading our test sentences. For future experiments it could be interesting to see what happens when participants do get feedback. In the way we have now set up our experiment, there are no direct consequences for the participant when answering incorrectly, potentially leading to less serious behavior overall. Another possibly confounding factor in the comprehension questions is that we did not check the ratio between questions that should be answered with "yes" and questions that should be answered with "no". It is probably better to make sure that the correct answers are equally distributed. With a very uneven ratio of the answers, participants could potentially try to find a pattern, which is not something we want them to do.

Lastly, we recruited our participants through Prolific, an online platform. We had no real way to check how serious the participants for the experiment were and what the conditions of their environment were. Participants could possibly distracted or only interested in earning the reward. The internet connection could also influence the measured reading times for example. All in all, it is difficult to say overall how reliable the measurements are. A way to counteract this would of course be to conduct the study in a more controlled lab setting, which was not possible at the time, but it is a recommendation for the future. A controlled lab setting will have less distractions and can also make participants take the experiment more seriously.

5 Conclusions

In conclusion, we started off by re-conducting an experiment set up by Dotlačil and Brasoveanu (2021). Through a self paced reading task, we investigated the reading times of plural sentences combined with a distributive disambiguator *each*, a collective disambiguator *all* or without a disambiguator. The sentences were all pragmatically distributive dispite the disambiguator. We expected to see a difference between the items with *each* and without disambiguator compared to the items with the disambiguator *all*. Eventually we found a small difference in reading times in the chosen region. The items with the disambiguator *each* were slightly cheaper in region 3 than sentences with no disambiguator or the disambiguator *all*, yet these results where insignificant. Finally it seems that items with the disambiguator *each* took longer to read in the predicate region. There was a lot of variance for this condition in the predicate region as well so it is uncertain how dependable these results are. There were multiple areas in the experiment that could be improved upon, and when done so, it could very well produce interesting results.

5.1 Future research

Apart from controlling for the possibly confounding factors mentioned in the discussion we do have a suggestion for possible future research.

The experiment could be set up in a different way from self paced reading. While many studies before used the self paced reading to measure reading times of words, it could be interesting to explore other measures to determine if there is a preferences for the collective reading. Studies could for example try to measure the effects of an assumed higher processing load associated with forcing a collective reading of pragmatically distributive sentences. This could take the form of a simple psychological experiment in which the cognitive load of participants is measured, for example a memory task. Participants would have to perform this task while reading pragmatically distributive sentences, and we could see how well they performed at the cognitive load task when they are presented with a collective reinterpretation of the pragmatically distributive sentence.

References

H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974. doi: 10.1109/TAC.1974.1100705.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1): 1–48, 2015. doi: 10.18637/jss.v067.i01.

P. J. Brooks and M. D.S. Braine. What do children know about the universal quantifiers all and each? *Cognition*, 60(3):235 – 268, 1996. ISSN 0010-0277.

J. Dotlačil and A. Brasoveanu. The representation and processing of distributivity and collectivity: ambiguity vs. underspecification. *Glossa: A Journal of General Linguistics*, 6(1), 2021.

A. D. Drummond. *Ibex: Internet Based EXperiments*. 2007.

F. Ferreira and J. Henderson. Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of experimental psychology. Learning, memory, and cognition*, 16:555–68, 08 1990. doi: 10.1037//0278-7393.16.4.555.

Frazier et al. Taking on semantic commitments, ii: Collective versus distributive readings. *Cognition*, 70 1:87–104, 1999.

L. Frazier and K. Rayner. Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language*, 29(2):181 – 200, 1990. ISSN 0749-596X.

J. Jegerski. *Research Methods in Second Language Psycholinguistics*, pages 20–49. 01 2014.

M. Just and P. Carpenter. A theory of reading: from eye fixations to comprehension. *Psychological review*, 87 4:329–54, 1980.

M. Just et al. Paradigms and processes in reading comprehension. *Journal of experimental psychology. General*, 111:228–38, 07 1982.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL `https://www.R-project.org/`.

J. Zehr and F. Schwarz. *PennController for Internet Based Experiments (IBEX)*. 2018. doi: 10.17605/OSF.IO/MD832.

A Appendix A

Test items Group A

1. The women read a newspaper in the morning just before going to work.
2. The first-year students had to read a classic novel during the first half of the academic year.
3. After studying the students took an exam in the last week of term.
4. The kids rode a unicycle as part of their final performance.
5. The managers drank a celebratory cocktail after landing the deal.
6. The kids ate an apple after their mother insisted for a third time.
7. During their trip to New York, the parents all ate a hotdog with relish and mustard.
8. The yoga instructors all drank a cup of herbal tea before going to bed.
9. After multiple attempts, the girls all passed their driving test and could finally get a license.
10. The brothers all ate a bagel with cream cheese and salmon at the store.
11. Once in a while, the friends all threw eggs at the abandoned hospital.
12. The boys all drank a Belgian craftbeer at the local brewery.
13. The men each kept their car clean and waxed and vacuumed the inside weekly.
14. During the talk the students each sat in an uncomfortable chair to be near the speaker.
15. Last Monday, the children each sent a postcard for Mother's day.
16. Last summer, the sisters each gave birth to a beautiful baby in the city hospital.
17. Last evening, the guests each had a sip of wine from the 1940 barrel in the cellar.
18. Last night, the children each ate a cookie after dinner without asking their parents.

Test items Group B

1. Last night, the children each ate a cookie after dinner without asking their parents.
2. The women all read a newspaper in the morning just before going to work.
3. The first-year students all had to read a classic novel during the first half of the academic year.
4. After studying the students all took an exam in the last week of term.
5. The kids all rode a unicycle as part of their final performance.
6. The managers all drank a celebratory cocktail after landing the deal.
7. The kids all ate an apple after their mother insisted for a third time.
8. During their trip to New York, the parents each ate a hotdog with relish and mustard.
9. The yoga instructors each drank a cup of herbal tea before going to bed.

10. After multiple attempts, the girls each passed their driving test and could finally get a license.
11. The brothers each ate a bagel with cream cheese and salmon at the store.
12. Once in a while, the friends each threw eggs at at the abandoned hospital.
13. The boys each drank a Belgian craftbeer at the local brewery.
14. The men kept their car clean and waxed and vacuumed the inside weekly.
15. During the talk the students sat in an uncomfortable chair to be near the speaker.
16. Last Monday, the children sent a postcard for Mother's day.
17. Last summer, the sisters gave birth to a beautiful baby in the city hospital.
18. Last evening, the guests had a sip of wine from the 1940 barrel in the cellar.
19. Last night, the children ate a cookie after dinner without asking their parents.

Test items group C

1. Last night, the children ate a cookie after dinner without asking their parents.
2. The women each read a newspaper in the morning just before going to work.
3. The first-year students each had to read a classic novel during the first half of the academic year.
4. After studying the students each took an exam in the last week of term.
5. The kids each rode a unicycle as part of their final performance.
6. The managers each drank a celebratory cocktail after landing the deal.
7. The kids each ate an apple after their mother insisted for a third time.
8. During their trip to New York, the parents ate a hotdog with relish and mustard.
9. The yoga instructors drank a cup of herbal tea before going to bed.
10. After multiple attempts, the girls passed their driving test and could finally get a license.
11. The brothers ate a bagel with cream cheese and salmon at the store.
12. Once in a while, the friends threw eggs at at the abandoned hospital.
13. The boys drank a Belgian craftbeer at the local brewery.
14. The men all kept their car clean and waxed and vacuumed the inside weekly.
15. During the talk the students all sat in an uncomfortable chair to be near the speaker.
16. Last Monday, the children all sent a postcard for Mother's day.
17. Last summer, the sisters all gave birth to a beautiful baby in the city hospital.
18. Last evening, the guests all had a sip of wine from the 1940 barrel in the cellar.
19. Last night, the children all ate a cookie after dinner without asking their parents.

Filler sentences

1. Three students watched the speaker give an shocking presentation.
2. Five girls looked at dresses for their upcoming prom on the webshop.
3. For graduation night, two studens booked a room at an expensive hotel.
4. Over the weekend, four parents baked cookies for the class graduation party.
5. During the past week, ten crimes were solved by the local police.
6. One girl threw the coolest birthday party of the year.
7. In a few months, twenty athletes will compete in the biggest vaulting tournament ever.
8. In the past few weeks, fifteen students tried to finish their degrees.
9. Fourteen cyclists competed for first place in the winter festival race.
10. When they were in Amsterdam, the five friends ate out every single night.
11. During the holidays, the ten girls sent each other letters.
12. Ten big men tried their hardest to win the weight lifting competition.
13. Twenty-one pilots landed their planes safely on the make-shift runway.
14. Sixty elderly people all lived together in the church.
15. In the future, many more people then now will have robotic limbs.
16. Two years ago, the two couples got married on the same day.
17. The five famous actors tried their hardest to finish the movie on time.
18. When in college, the four boys played games every night.
19. The four friends met up every once in a while after they moved.
20. When looking for new applicants, the three recruiters looked for the best candidate.
21. In their search for presents, the girls found themselves running from store to store.
22. In a few weeks, the sisters will get together to celebrate a birthday.
23. The six cousins always watch movies together on sunday.
24. The next six games will determine the winner of the football competition.
25. The seven girls gave the best presentation in the senior class.
26. When reading the difficult paper, the students took a lot of notes.
27. During Christmas, the three couples visited each of their families.
28. In the coming period, the students will have to pay more attention during class.
29. The eighteen cooks worked very hard to make dinner for all the guests.
30. When on their trip, the two families hardly interacted.
31. During the fight, two of the students used their books as weapons.

Comprehension Questions

1. Did the women read a book?
2. Did the students have to read a classic novel?
3. Did the students take an exam?
4. Did the children play with a balloon?
5. Did the managers drink coffee?
6. Did the kids eat an apple?
7. Did the parents eat a pizza?
8. Did the yoga instructors drink tea?
9. Did the girls pass their test?
10. Did the brothers eat a bag of chips?
11. Did the brothers throw toilet paper?
12. Did the boys drink beer?
13. Did the men keep their bike clean?
14. Did the students sit on a couch?
15. Did the children send a card?
16. Did the sisters have a baby?
17. Did the guests drink beer?
18. Did the children have dinner?
19. Did the students watch a movie?
20. Did the girls look in a magazine?
21. Did the students book an expensive room?
22. Did the parents bake cookies?
23. Did the police solve crimes?
24. Did a boy throw the coolest birthday party of the year?
25. Will twenty artists compete?
26. Did the students try to finish their degrees?
27. Did the cyclists compete for third place?
28. Did the friends eat out every night?
29. Did the girls send each other video's?

30. Did the men try to bake cookies?
31. Did the pilots crash their planes?
32. Did the elderly people live in the elderly home?
33. Will more people have robotic limbs?
34. Did the couples get married on the same day?
35. Did the actors try to finish the movie on time?
36. Did the boys study every night?
37. Did the friends meet up after they moved?
38. Did the recruiters look for bad candidates?
39. Did the girls visit only one store?
40. Will the sisters celebrate a holiday?
41. Do the cousins watch movies alone on sunday?
42. Will the games determine the winner of the rugby competion?
43. Did the girls give the best presentation?
44. Did the students take notes?
45. Did the couples visit only one family?
46. Will the students have to pay more attention during class?
47. Did the eighteen chefs work hard?
48. Did the families interact a lot?
49. Did two students use their books for reading?

B Appendix B

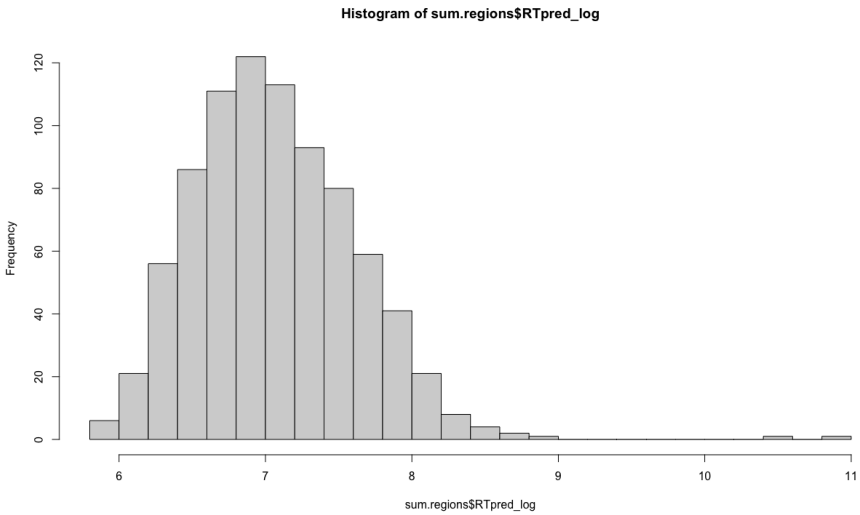

Figure B.1: Histogram of the logarithmic reaction time of predicates

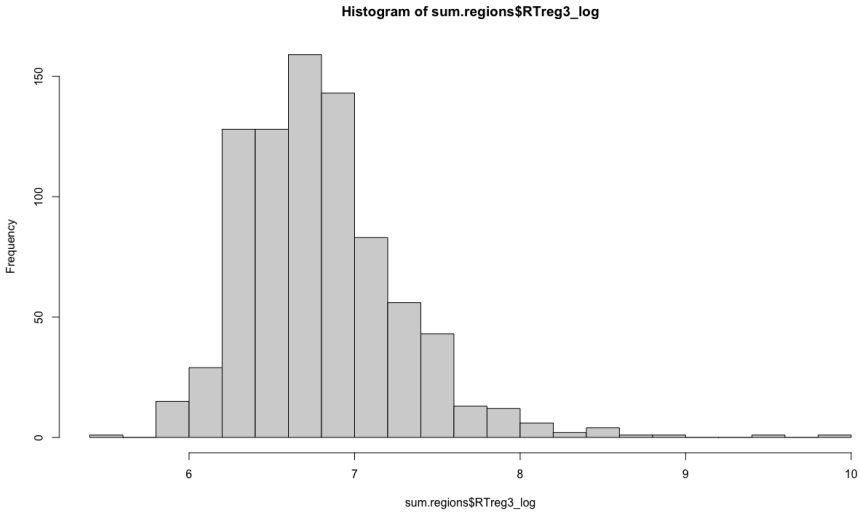

Figure B.2: Histogram of the logarithmic reaction time of region 3

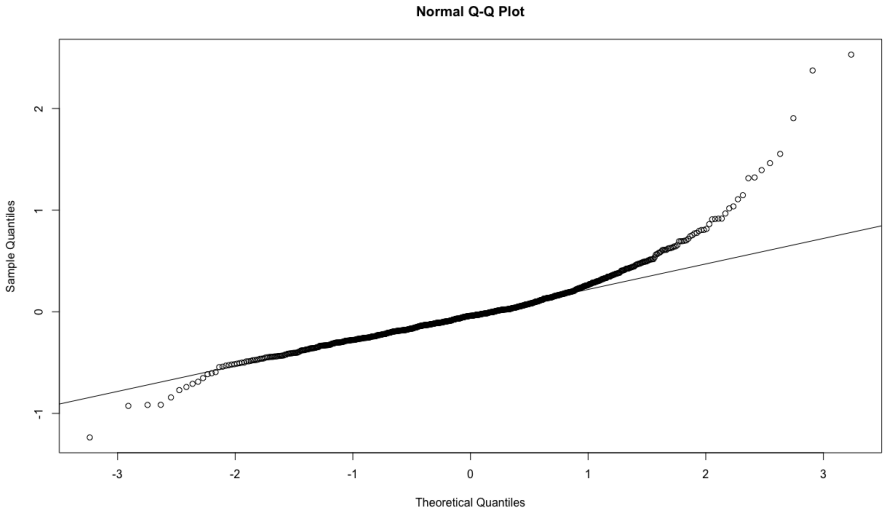

Figure B.3: QQ-plot of fitted model with outliers

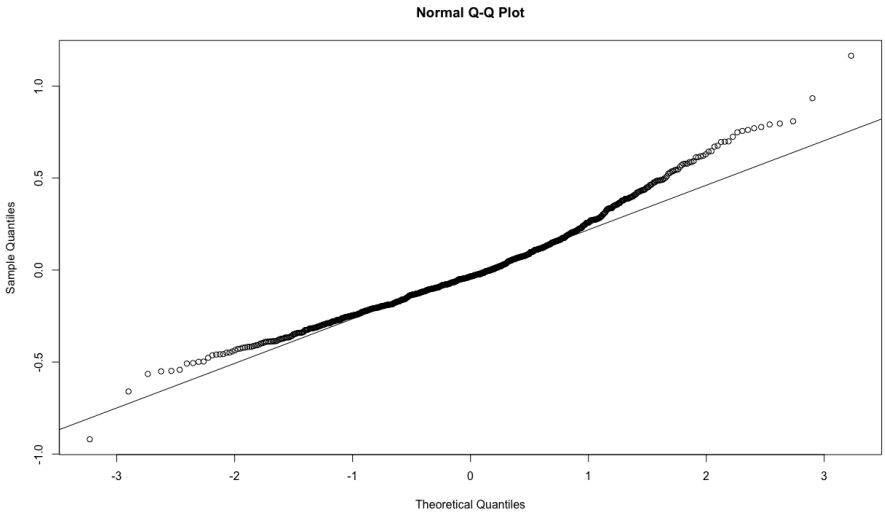

Figure B.4: QQ-plot of fitted model with outliers removed