university of groningen / faculty of science and engineering / artificial intelligence

Long-Distance Dependencies in Dutch: measuring the effect of the quantified pronoun *er* on reading speed

Bachelor's Project Thesis

Daan Kaasjager, s3905845, d.a.kaasjager@student.rug.nl
Supervisor: Dr S. Jones

Abstract: The Dutch word *er* is one of the most versatile and syntactically complex words of the language. This study examined the effects of the presence of the quantitative Dutch word *er* in a sentence on reading speeds. Foundational work by Stowe (1986) shows that a slow-down in reading speed occurs for words that have a Long-Distance Dependency (LDD) in a sentence. A between-subjects design using eye tracking software was set up to test similar behavior for the quantitative pronoun (er_Q). No significant effect of er_Q on reading speed was found. These findings could be explained by the inherent ambiguity of the word but also some methodological measurement issues.

1 Introduction

The word *er* is arguably one of the most interesting words of the Dutch language. It also seen as one of the most difficult words to comprehend for non-native speakers due to its versatility. *Er* is a *multi-functional particle* (a word with little meaning that cannot stand on its own but which can be used in many different contexts) that plays a crucial role in the structure and meaning of sentences in the Dutch language. It can be used as a pronoun or adverb, and its specific function can depend on the context in which it is used.

Odijk (1993) distinguishes between four different uses of the word *er*; (1) existential *er* (er_X), (2) quantitative *er* (er_Q), (3) locative *er* (er_L) and (4) prepositional *er* (er_P). While this study focuses on the effects of er_Q, understanding the difference between the types of *er* is important to fully comprehend the research presented in this paper. Therefore, each use of *er* will be explained in further detail in sections 1.1 through 1.4 of this introduction.

1.1 Existential er_X

According to Odijk (1993) existential er_X, also known as presentative er_X, occurs at the start of a clause when either the subject is indefinite (1) or when it is the subject of an impersonal passive sentence (2). Er_X is required when the first clausal position is occupied by a non-subject constituent or in embedded clauses and there is no explicit subject (3), but it can occur optionally when there is an explicit subject (4). Changing er_X with *daar* or *hier* in the examples below would change the meaning of the sentence.

(1) *Er was niemand op het feestje.*
There was nobody on the party.
'There was nobody at the party.'

(2) *Er is niet gelachen.*
There is not laughed
'There was no laughter.'

(3) *In de rij stond (er_X) een man.*
In the row stood there a man
'In the row stood a man.'

(4) *Gisteren was er_X alles aan*
Yesterday was there everything on
gedaan om het goed te maken.
done to it good to make.
'Yesterday everything was done to make it right.'

1.2 Prepositional er_P

Prepositional er_P, unsurprisingly, is used in combination with prepositions and translates to the English word "it" (5). While it does translate to "it", using the Dutch article *het* instead would be ungrammatical (6). Er_P is also known as pronominal er_P (Jones, 2020) and can be replaced by the Dutch words "daar" and "hier" which translate to "there" and "here" in English respectively. Interestingly, it cannot be placed after the preposition, and neither can the literal translation of "it", "het" (7). Er_P must be placed either immediately before the preposition (combining it into one word such as *ernaast/erbij/erop*) or as an LDD (5). The object after the preposition can be included to specify the location inside the sentence, but this changes the use to er_L (8).

(5) *Jan zat* er_P *tot hij klaar was met*
Jan sat it until he was done with
eten naast.
eating next.
'Jan sat next to it until he was done with eating.'

(6) * *Jan zat het tot hij klaar was met*
Jan sat it until he was done with
eten naast.
eating next.
'Jan sat next to it until he was done with eating.'

(7) * *Jan zat tot hij klaar was met eten*
Jan sat until he was done with eating
naast er/het.
next it.
'Jan sat next to it until he was done with eating.'

(8) *Jan zat* er_L *tot hij klaar was met*
Jan sat it until he was done with
eten naast zijn vader.
eating next his dad.
'Jan sat next to it until he was done with eating.'

1.3 Locative er_L

The locative version of *er* functions as a pronoun and can, just as for er_P be replaced by the Dutch words "daar" and "hier" which translate to "there" and "here" in English respectively (9). Er_L refers to a location or a place and cannot occur at the start of a sentence.

(9) *Jesse heeft* er_L*/daar/hier de hele dag*
Jesse has there/here the whole day
gestaan.
stood
'Jesse has stood there the whole day.'

1.4 Quantitative er_Q

Finally, er_Q functions as a partitive pronoun and is always used in combination with a quantifier (10). However, including a noun after the quantifier changes the meaning of er_Q to a er_L, as in (11). Important to note that therefore the use of *er* remains ambiguous until the word after the quantifier is read.

(10) *Thomas heeft* er_Q *gisteren drie*
Thomas has there yesterday *drie*
opgegeten.
eaten.
'Thomas has eaten three yesterday.'

(11) *Thomas heeft* er_L *gisteren drie*
Thomas has there yesterday *drie*
broodjes opgegeten.
broodjes eaten.
'Thomas has eaten three there yesterday.'

1.5 Long-Distance Dependencies

A Long-Distance Dependency (LDD), is a grammatical relationship between words or phrases that are separated by one or more intervening phrases (Zushi, 2013) and involve complex memory and processing management which create a high cognitive workload (Lakretz et al., 2020).

In English, wh-sentences (sentences containing a wh-phrase, which is a phrase functioning as a question operator) are the most common examples of long-distance dependencies (12). In sentence 12 for example, *Who* (which would be called the *dependent*) is stored in the working memory (Ness & Meltzer-Asscher, 2017) until its canonical thematic position is reached at the word *meet*.

(12) **Who** did you **meet** at the store?

An important note is that in Dutch, a Long-Distance Dependency can occur for er_Q. An example of this can be seen in sentence 10, where the dependent word is er_Q and the tail of the dependency is *drie*. Whether the effects of the LDD on sentence processing can be found for er_Q is explained in more detail below.

1.6 LDDs and reading speed

Research done by Stowe (1986) on parsing wh-sentences showed a slow-down in reading speed whenever the tail of the dependency (also known as the *gap* (Wagers, 2013))(Frazier & Fodor, 1978) of the wh-phrase was reached while reading a sentence. This experiment showed that processing is substantially more difficult for wh-sentences than for declarative sentences, which do not have a Long-Distance Dependency. Furthermore, slow-downs for LDDs can also be seen in languages other than English, such as Japanese (Aoshima et al., 2004), and Russian Sekerina (2003). Whether this effect can be found for the Dutch word er_Q remains to be tested.

Expectation plays an important role here: the processing difficulty only occurs when a gap is expected to occur in the sentence. In other words, if the expectation of a gap did not develop as soon as the dependent word was read, a slow-down would not occur.

This is important in the context of er_Q, due to its ambiguity with er_L as shown in section 1.4. Because there is no expected gap for er_L, it is unlikely there will be a slow-down if participants misinterpret er_Q for er_L at the start of the sentence.

Stowe (1986) shows us that the expectation of the gap created by the LDD is what causes the slow-down in reading speed. From this we can infer that a slow-down in reading speed could provide evidence of LDD processing behavior in a sentence. Other studies have demonstrated that a slow-down in reading speed is often caused by a higher cognitive workload (Just & Carpenter, 1980)(Conklin et al., 2018). Lakretz et al. (2020)'s ideas on the complexity of LDDs are in line with this, as keeping the expectation stored in memory would require complex memory processing management and leads to a high cognitive workload. Measuring the speed at which words are read in a sentence could therefore show if LDD slow-down effects are er_Q, which sets up an LDD.

1.7 Wrap-up effect

It is important to keep in mind that the syntactic position of a word in a sentence is an important modulator of the reading speed. Britton et al. (1982) show there is very often a slower reading speed at the start and at the end of a sentence. This slow-down, named the *wrap-up effect* by Just & Carpenter (1980), is especially measured at the end of a sentence. Many researches have found evidence for a substantial decline in reading speed at the end of a sentence or a clause using a self-paced reading paradigm (Mitchell & Green, 1978) (Just et al., 1982) (Hill & Murray, 2000). Further research using eye-tracking experiments have found similar results, where participants' reading speed also slowed down at the end of the sentence Rayner et al. (1989) (Rayner et al., 2000). When designing experiments where reading time is measured and compared, it is therefore important to keep the *wrap-up effect* in mind, to mitigate the power it has on the effects that are being measured.

1.8 Research question

One of the most impressive papers done on *er* and reading speeds was done by Grondelaers et al. (2009):

In his research on er_X,Grondelaers et al. (2009) shows that er_X can behave as a cue to reduce processing. Er_X functions as a so-called *expectancy monitor,* which reduces reading speed. In this paper er_X is inserted in cases where variables affect the predictability of the subject in a sentence. Grondelaers et al. (2009) found that er_X facilitates the processing of the subject in the sentence when placed in context of abstract location adjuncts, but did not facilitate any processing in the context of concrete location adjuncts. An example is shown in (13) and (14) where the presence of er_X facilitated processing and decreased reading times, especially for unexpected subjects as in (14).

(13) *In het uitstalraam van de juwelier lag*
In the shop-window of the juweler lay
(er$_X$) een halssnoer
there a necklace

"In the jeweler's shop-window (there) was a necklace'

(14) *In het uitstalraam van de juwelier lag*
In the shop-window of the juweler lay
(er$_X$) een sandwich
there a sandwich

"In the jeweler's shop-window (there) was a sandwich'

The question remains whether er_Q will behave similarly to er_X. Er_X does not set up an LDD in a sentence and therefore does not require any of the complex memory processing that LDDs have. As stated in section 1.5, LDDs involve convoluted cognitive memory functions with high cognitive workload. The slow-down in reading speed found by Just & Carpenter (1980) and Conklin et al. (2018) which is often caused by a higher cognitive workload, is in accordance with Stowe (1986)'s work on LDDs for English wh-phrases. A slow-down in reading speed can often be the cause of reaching the tail of the dependency of an LDD in the sentence.

This study addresses a specific area of linguistic cognitive processing, specifically for the Dutch language. The word *er* has been studied for many years by linguists due to its complexity and versatility. The aim of this research is to investigate whether er_Q is processed as an LDD in the brain. Using reading speed, we can investigate if er_Q shows similar slow-downs when the tail of the dependency is reached. This leads to the following research question:

"What is the effect of the presence of the Dutch quantified pronoun er_Q in a sentence on reading speeds?"

The findings can be used to contrast and compare how various languages are processed in the brain and can therefore lead to a better understanding of how languages shape and affect cognitive processes.

Therefore, given that a slow-down in reading speed for an LDD is an effect that has been replicated and reproduced in many languages it is expected to see a slow-down in reading speed when Dutch er_Q is present in a sentence.

2 Method

2.1 Design

This study was a between-subjects (see section 2.3.1) quantitative experimental design, in which participants read sentences that varied in the presence or absence of the Dutch word "er." For each participant the experiment was completed in a single session without breaks and lasted approximately 10 minutes.

2.2 Participants

A total of 24 native Dutch speakers participated in the study. Participants were recruited on campus. Inclusion criteria for the study were as follows: (1) native Dutch speaker, (2) no history of reading difficulties, or learning disabilities and (3) had normal/corrected-to-normal vision. A total number of 16 men and 8 women participated. The age group was 19-25 years. All of the participants received 5 euros for participating in the experiment.

2.3 Materials

The materials for this study included a computer with a 17-inch monitor, a keyboard, an Eyelink Portable Duo Eye Tracker to track the movement of the eyes on the screen and used Open Sesame (Mathôt et al., 2012) as a program for presenting the experimental stimuli and recording the reading times. All of the experiments were performed in the same room.

2.3.1 Sentences

The stimuli for the study were a set of 80 Dutch sentences, 20 target sentences, 20 baseline sentences and 40 filler sentences. The sentences were matched for length and complexity. The target and baseline sentences were paired: target sentences include the word er_Q and baseline sentences are near identical but do not include er_Q. The three types of sentences are explained in more detail below.

For each participant, 20 sentences were randomly selected from a combined pool of the target and baseline sentences and 20 sentences were randomly selected from the filler pool. Participants were paired, e.g. if participant 1 has a target sentence for a certain stimulus then participant 2 has a

baseline sentence for that stimulus and vice versa. This was to ensure that baseline and target sentences were tested the same amount of time.

1. **Target sentences.** These are sentences that contain the word er_Q. An example of this is shown in sentence 15 below. These sentences start off with an adverbial phrase which is shown in red. It consists of a phrase that is intended to prime the correct use of *er* for the participant. In Dutch, when reading (17) for example, when the word *er* is reached it is still unclear whether this *er* is a er_Q, er_P or er_L. To solve the issue of this ambiguity, all target sentences begin with a phrase that indicates that the use of *er* in the sentence is most likely an er_Q, colored in red in sentence 15. Following the priming phrase is the verb and the subject of the sentence, shown in brown. This is then followed by the word er_Q in black. After er_Q each sentence contains an adverbial phrase functioning as a buffer between er_Q shown in teal below. This buffer was consistently kept between 3-5 words and did not contain any prepositions, to avoid any confusion with er_P. The buffer was followed by a quantifier shown in orange which is the tail of the dependency. Finally, to account for the *wrap-up effect* as explained in section 1.7, each sentence contains an adverbial phrase after the quantifier shown in blue.

   (15) *Van de drie kocht Joost* er_Q
   Of the three bought Joost ER
   gisteren op zijn werk twee voor zijn
   yesterday on his work two for his
   oma.
   grandma
   'Out of the three Joost bought two at his work yesterday for his grandma.'

2. **Baseline sentences.** These sentences were structured very similarly to the target sentences mentioned above. Baseline sentences like sentence 16 also start with a priming phrase, colored in red in sentence 16. Just as target sentences, following the priming phrase is the verb and the subject of the sentence, shown in brown. Instead of *er*, these sentences contain the word *al*, colored in black and which translates to the english word *already*. *Al* is followed by 3-5 words as a buffer between *al* and the quantifier shown in teal. A quantifier and a noun shown in orange are placed after the buffer. The added noun plays an important role as simply replacing the word *er* with *al* and not making any further changes to the sentence would make the sentence grammatically incorrect. For this reason a noun is added after the quantifier to correct the syntax. Note, adding a noun after the quantifier for a target sentence would be unsuitable, as er_Q would then become ambiguous with er_L. Finally, as in the target sentences, to account for the *wrap-up effect* each sentence contains an adverbial phrase after the quantifier shown in blue.

   (16) *Van de drie kocht Joost al*
   Of the three bought Joost already
   gisteren op zijn werk twee
   yesterday on his work two
   broodjes voor zijn oma.
   sandwiches for his grandma
   'Out of the three Joost already bought two sandwiches at his work yesterday for his grandma.'

3. **Filler sentences.** Filler sentences were used to mask the goal of the experiment. These sentences were of similar length when compared to the target and baseline sentences. The writing style was also similar to the target and baseline sentences. The main goal of using these sentences was to reduce bias and expectations of the participants while performing the experiment.

2.3.2 Eye tracker

A major design choice in this experiment is whether to use an eye tracker or implement a self-paced reading (SPR) program to collect participant data. While other research on er_P (Thiel, 2021)(Heinhuis et al., 2020) used SPR as a method for recording reading speed and therefore inferring cognitive load and there are a few major advantages to using eye tracking. First of all, eye tracking allows for participants to read more naturally, especially in cases where the participant rereads parts of the sentence.

Furthermore, noise is added to the data because the reaction time of the participant is included into the reading time. An experiment byKung (2021) found that eye tracking yielded more reliable results compared to SPR due to spillover effects.

Although the latter is a reliable method which has provided useful information about sentence processing, it may nevertheless be problematic for the validation of unaccented small words such as *er*.

2.4 Procedure

Participants sat in a room with their head resting on a chin rest in front of the monitor. The eye tracker was placed at an angle out of the line of sight of the monitor but still in front of the participant. Participants first followed auditory instructions to calibrate the eye tracker before the start of the experiment. This was done using the calibration software of the Eyelink Portable Duo Eye Tracker.

After the eye tracker was properly calibrated, participants began with the experiment. Participants were presented instructions for the experiment on the monitor. Each trial of the experiment, participants were presented a visual stimulus. After the participants had read the stimulus, they pressed a key to continue to the next trial. Between trials, participants performed a drift correction, to recalibrate the eye tracker. Drift corrections present a crosshair at the center of the screen that participants must look at, after which the software corrects for shifted calibration. The program will continue if the gaze position is near the center of the screen and stabilized. Participants were not recorded during drift corrections. Once the participant had gone through all the trials, the participant was presented a message on the screen that the experiment was finished.

2.5 Data Processing

Eye tracking data files are large and require processing before any statistical tests can be performed. After converting the files from *.edf* (European Data Format, commonly used to store multichannel biological and physical signals) to *.asc* (files containing ASCII-encoded data) for easier data extraction, the data were segmented into a csv containing the participants fixations during trials using a python script. For each fixation during the experiment, the csv file contained the participant ID, the sentence ID, the duration (ms), and the x- and y-coordinate.

Each of the participants randomized trial data were matched to their output csv file. For all target and baseline sentences, the Region Of Interest (ROI) was set as the quantifier in the sentence and the following two words. This ROI was chosen to address the well-established *spillover effects* of processing time when reading sentences. These effects entail that the processing of a word is not always completed when the eyes have moved to the next word (Rayner & Duffy, 1986)(Ehrlich & Rayner, 1983). The time spent processing a word then spills over to the next word. Rayner (1998) on what the best measure of processing time is using eye-tracking says that "processing time per word is a pale reflection of the reality of cognitive processing". Therefore, to account for the *spillover effects* the ROI includes the quantifier and the next two words. Any slow-down of reading speed in this region indicates increased cognitive processing.

There is a lot of controversy regarding what the best way is to analyze eye movements around the ROI to be able to infer cognitive processing (Rayner, 1998). There is especially heavy discussion about what the best way is to include regressions, which are right-to-left movements along the line of text, in the analysis (Altmann, 1994)(Rayner & Serene, 1994)(Rayner & Sereno, 1994). According to Rayner (1998), most regressions are short regressions and are less than a few letters long. They occur due to an overshot saccade. After the overshot saccade a corrective short regressive saccade is necessary for efficient sentence reading. Furthermore, within-word short regressions could be due to the fact that the reader does not understand or has difficulty reading the current word.

Long regressions are regressions that occur more than 10 letter spaces back and are due to the fact that the reader did not understand the text Rayner (1998). A study on structurally ambiguous sentences by Frazier & Rayner (1982) supports this theory: when readers reach a word that is not in line with their initial analysis of the sentence, a regressive eye movement occurs to the region of ambiguity. In the case of er_Q, which remains ambiguous with er_L until the word after the quantifier has been

read, a regressive eye movement could then mean that the participant initially thought that the use of *er* was er_L and this is then corrected by a regressive saccade.

For some researches however, *processing difficulty* is considered a more general cause for these long regressive saccades (Brysbaert, 1994). In fact, Brysbaert (1994) states that regressions indicate processing difficulties from the region from which they originate and are therefore important to include to the reading time of that region. For LDDs, this means that when readers reaches the gap, a regressive saccade which moves the eyes to the dependent word might occur. Including regressive saccades when researching filler-gap dependencies has been done in other studies as well Taylan (2014), to give useful information on the difficulty of text processing.

In the case of this study, whether regressive saccades occur due to processing of an LDD or are due to the misinterpretation of er_Q for er_L is difficult to determine. However, to not include the regressions altogether would not be representative of the processing time required to read target words in the sentence. Even though it would therefore be difficult to distinguish between regressions caused by misinterpretation of the use of *er* and regressions caused by LDD processing time, including regressions in the data gives a more complete view of the data. On top of that, a priming phrase is used at the start of the sentence to increase the chances of participants assuming the correct use of *er*.

Participants fixations durations were measured and accumulated until the participants fixations moved past the ROI on the x-axis and are called *accumulated fixations*. Cases where participants reread the sentence *after* moving past the ROI were not used for further analysis.

Due to the location of the drift correction fixation cross overlapping with the ROI, the start of each trial includes the starting location of the eyes as a fixation inside the ROI. Therefore for each trial, the first two fixations of that trial were removed for all participants.

3 Results

This study attempted to find the underlying relationship between the presence of er_Q in a sentence and reading speeds. The study was of between-subjects quantitative design. Each of the 24 participants' 20 target sentences were included in the analysis.

Due to the fact that the *accumulated fixation* data was highly skewed, the natural logarithm was taken for the *accumulated fixations*. This helped reduce the impact of outliers on the data set and made it easier to do statistical tests. After taking the natural logarithm, any data points differing more than 2.5 standard deviations from the mean were removed from the data set to reduce the impact of outliers on the underlying patterns.

Target sentences appear to be read at a faster rate at first glance when comparing mean and standard deviation in figure 3.1. The data also seem differently distributed, with baseline sentences having a more concentrated log duration when compared to target sentences. The data are also not very normally distributed but is still slightly skewed, even after taking the logarithm.

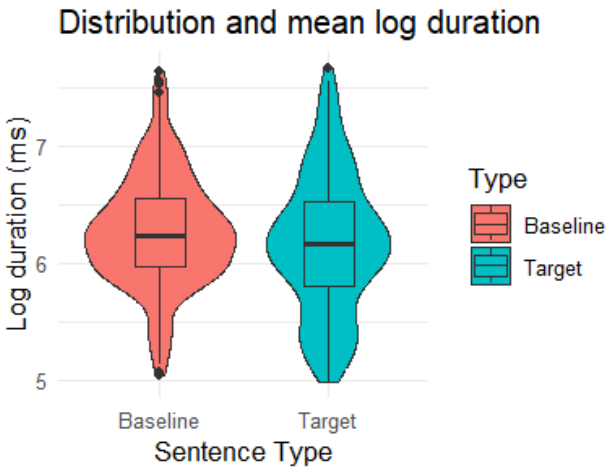

Figure 3.1: A violin plot combined with a boxplot showing the mean accumulated duration of the fixations and the distribution.

To examine the effect of the sentence type on the accumulated fixation duration, a linear mixed effects model was fit to the data. The sentence type was the only fixed effect. When viewing the data there was evidence for systematic noise caused by Sentence ID and Subject Number. Therefore, random effects of the model were the intercepts for participants and the sentence pairs and random slopes

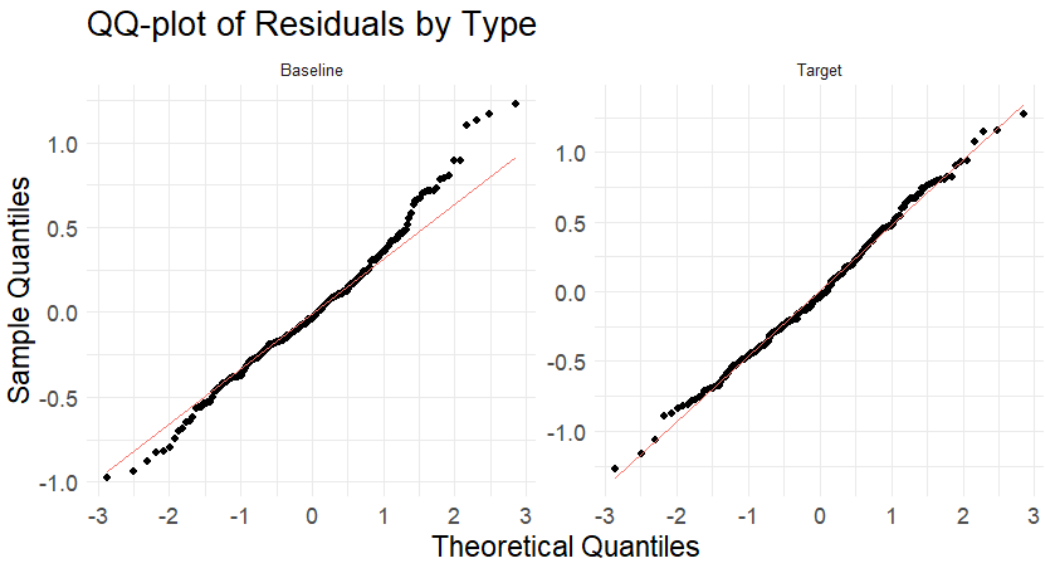

Figure 3.2: QQ plot of the residuals for each sentence type.

for by-participant and by-sentence pairs for the sentence type.

	Effects of erQ on reading speeds		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.26	6.14 – 6.38	**<0.001**
Type (Target)	-0.07	-0.24 – 0.09	0.371
Random Effects			
σ^2	0.21		
$\tau_{00\ Subject}$	0.05		
$\tau_{00\ Number}$	0.01		
$\tau_{11\ Subject.TypeTarget}$	0.05		
$\tau_{11\ Number.TypeTarget}$	0.06		
$\rho_{01\ Subject}$	-0.26		
$\rho_{01\ Number}$	-0.56		
$N_{\ Subject}$	24		
$N_{\ Number}$	20		
Observations	470		
Marginal R^2 / Conditional R^2	0.005 / 0.300		

Figure 3.3: Table showing a summary of the linear mixed effects model

To assess the fit of the model, a QQ-plot was created as can be seen in figure 3.2. The baseline QQ-plot shows us that the data are peaked in the middle, which is line with figure 3.1. The target data appear to be normally distributed. There also do not appear to be any obvious deviations from homoscedasticity.

Furthermore, due to the fact that there were no practice trials in this experiment, abnormalities in the data based on the order of presentation were looked for, but none were found (figure 3.4).

An ANOVA test on the linear mixed effects model showed that the effects of sentence type on accumulated fixation duration were not significant $F(1,23) = 0.800$, $p = 0.371$. The findings therefore fail to reject the null hypothesis that the presence of er_Q in a sentence does not lead to a slow-down in reading speed. Variance caused by the random effects was marginal and the Conditional $R^2 = 0.300$ show the proportion of variance explained by the sentence type (figure 3.3).

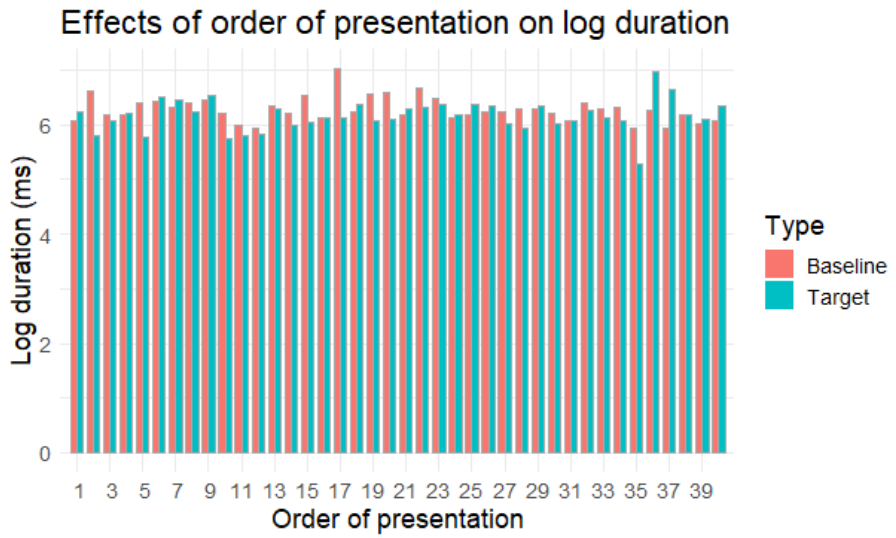

Figure 3.4: A barplot showing the mean accumulated duration of the fixations based on the order of presentation.

4 Discussion

Before the experiment it was hypothesized that the presence of er_Q in a sentence would slow down reading speed around the quantifier, due to the effect of the Long-Distance Dependecy that it has with the quantifier in the sentence. This effect has been replicated often, in English and in many other languages (Stowe, 1986)(Aoshima et al., 2004)(Sekerina, 2003). The lack of evidence for a slow-down in reading speed caused by er_Q, while surprising, could be explained by a few different factors:

First, one of the major issues with researching er_Q, is its inherent ambiguity with er_L. An example of this is shown in sentence 17, where we see that until the tail of the dependency is read in the sentence (which is missing in this example), the use of *er* remains ambiguous.

(17) * *In totaal heeft Daan* er_L/er_Q *tijdens*
In total has Daan ER during
zijn werk -
his work -
'In total Daan has during his work -'

This ambiguity is only resolved once the quantifier has been read and importantly, not immediately followed by a noun, as in sentence 18. Even though these sentences are syntactically sound, these type of sentences are quite uncommon. This might have lead to some confusion around the type of *er* that participants were reading in the experiment.

(18) * *In totaal heeft Daan* er_L/er_Q *tijdens*
In total has Daan ER during
zijn werk drie gezien.
his work three saw
'In total Daan has seen three while he was working.'

Confusion with er_L could have a large impact on the results as we know that expectation plays an important role in the formation of an LDD. Section 1.6 states that if the expectation of a gap did not develop as soon as a dependent word was read, slow-downs would not occur.

It is difficult to determine whether this confusion occurs for participants during the experiment. Even though regressive saccades often occur when readers reach a word that is not in line with their initial analysis of the sentence (which would be the case if the participant misinterpreted er_Q for er_L), we also know that regressive saccades often occur when the tail of the dependency is read in an LDD. For this reason the *accumulated duration* is a tricky measure.

Another issue was the general complexity of the sentences. By accounting for the *wrap-up effect* (see section 1.7) and by including the *priming phrase*, sentences became long and convoluted. Some participants' initial feedback was that some sentences were difficult to comprehend as a result of this. In future research, this could be accounted for by separating the priming phrase from the sentence into two sentences, for which an example can be seen in sentence 19.

(19) * *In totaal heeft hij vier collegas. Daan heeft* er_L/er_Q *tijdens zijn werk drie gezien.*
In total has he four colleagues. Daan has ER during his work three saw
'In total he has four colleagues. Daan has seen three while he was working.'

This splits the sentence which makes it more readable while still functioning as a primer for the quantifier.

Another possible reason for the lack of significance in the findings is that the effect size could be too small and that with a group of 24 participants there is too much variance. The effect size measured for sentence type was notably in the oppposite direction than was hypothesized, but was very small nevertheless. A larger group of participants could reaffirm that there is no effect, or could push the effect size to either direction.

Finally, another explanation for these findings could be that er_Q does not set up a long-distance dependency in the same way other languages do. Even though the slow-down effect for LDDs is a well-established effect, it is possible that *er* is processed in a unique way due to its versatility and ambiguity. Its unique processing could be explained by the fact that *er* isn't as semantically weighty as words with similar meanings (e.g. *daar* or *hier*) and often is not stressed when read. Native Dutch speakers therefore wouldn't require much processing power for the word and would not have to slow down when reading it in a sentence. Er_Q could possibly even speed up reading speed and work as an expectancy monitor as Grondelaers et al. (2009) found for er_X. More research on all of the different forms of *er* could be done for a more complete view of the way it is processed in the brain.

5 Conclusion

The aim of this study was to discover any effects of er_Q on reading speeds. In Dutch, er_Q is dependent on the quantifier in the sentence, when er_Q and the quantifier are separated by one or more phrases, we can speak of a Long-Distance Dependency (LDD) between the two words. In English and many other languages where LDDs occur, a slow-down takes place when reaching the tail of the dependency (Stowe, 1986)(Aoshima et al., 2004). The main goal of this research was to see if LDDs that occur for er_Q behave similarly to LDDs in other language.

The results of this research could tell us more about how Dutch *er* is processed in the brain when reading a sentence as it is such a versatile and dynamic word.

This research used a between-subjects experimental design where participants were asked to read sentences containing er_Q. Participants were also shown sentences without er_Q, as a control group. Using eye-tracking software, the reading speed of the quantifier in a sentence was measured.

A linear mixed effects model was fit to the data and an ANOVA test was performed on the model. Based on the well-established behavior of LDDs in other languages, a similar slow-down in reading speed was expected for er_Q, but no statistically significant evidence could be found for a slow-down caused by er_Q. The variance caused by the predictor as well as the random effects were relatively small, even when accounting for taking the logarithm of the accumulated duration. The conditional $R^2 = 0.300$ shows us that the sentence type was in general not a very good predictor of the reading time of the sentence.

Some possible explanations for these results are readers' confusion with the type *er*, which resulted in misinterpreting er_Q for er_L, as the effect size could be very small it could be due to a lack of participants. Other explanations are that sentences were too complex or that er_Q isn't a heavily stressed word and therefore involves different cognitive processes, perhaps even improving reading speed as it does for er_X.

For future research on er_Q, results might be more reliable if sentence complexity were to be reduced and also first tested for ambiguity with er_L. Furthermore, more research on LDDs for other types of *er* such as er_P might also shed light on how the

relationship is set up for similar uses. Finally, participant feedback might provide important information about the interpretation and expectancy of the sentences used in the experiment.

References

Altmann, G. T. (1994). Regression-contingent analyses of eye movements during sentence processing: Reply to rayner and sereno. *Memory & Cognition*, *22*, 286–290.

Aoshima, S., Phillips, C., & Weinberg, A. (2004). Processing filler-gap dependencies in a head-final language. *Journal of memory and language*, *51*(1), 23–54.

Britton, B. K., Glynn, S. M., Meyer, B. J., & Penland, M. (1982). Effects of text structure on use of cognitive capacity during reading. *Journal of Educational Psychology*, *74*(1), 51.

Brysbaert, M. (1994). Sentence reading: Do we make use of non-audible cues. In *4th workshop on language comprehension, giens* (Vol. 13, p. 14).

Conklin, K., Pellicer-Sánchez, A., & Carrol, G. (2018). *Eye-tracking: A guide for applied linguistics research*. Cambridge University Press.

Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of verbal learning and verbal behavior*, *22*(1), 75–87.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*(4), 291–325.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive psychology*, *14*(2), 178–210.

Grondelaers, S., Speelman, D., Drieghe, D., Brysbaert, M., & Geeraerts, D. (2009). Introducing a new entity into discourse: Comprehension and production evidence for the status of dutch er "there" as a higher-level expectancy monitor. *Acta Psychologica*, *130*(2), 153–160.

Heinhuis, M., et al. (2020). *The dutch worder'and processing times* (Unpublished doctoral dissertation).

Hill, R. L., & Murray, W. S. (2000). Commas and spaces: Effects of punctuation on eye movements and sentence parsing. In *Reading as a perceptual process* (pp. 565–589). Elsevier.

Jones, S. M. (2020). Multifunctional dutch 'er'.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological review*, *87*(4), 329.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of experimental psychology: General*, *111*(2), 228.

Kung, F. (2021). *Comparison of eye tracking and self-paced reading methodologies in relative clause processing* (Unpublished doctoral dissertation). The University of North Carolina at Chapel Hill.

Lakretz, Y., Dehaene, S., & King, J.-R. (2020). What limits our capacity to process nested long-range dependencies in sentence comprehension? *Entropy*, *22*(4), 446.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). Opensesame: An open-source, graphical experiment builder for the social sciences. *Behavior research methods*, *44*, 314–324.

Mitchell, D. C., & Green, D. W. (1978). The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology*, *30*(4), 609–636.

Ness, T., & Meltzer-Asscher, A. (2017). Working memory in the processing of long-distance dependencies: Interference and filler maintenance. *Journal of psycholinguistic research*, *46*(6), 1353–1365.

Odijk, J. E. J. M. (1993). *Compositionality and syntactic generalizations* (Unpublished doctoral dissertation). Katholieke Universiteit Brabant.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, *124*(3), 372.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & cognition*, *14*(3), 191–201.

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology: Section A*, *53*(4), 1061–1080.

Rayner, K., & Serene, S. C. (1994). Regression-contingent analyses: A reply to altmann. *Memory & Cognition*, *22*, 291–292.

Rayner, K., & Sereno, S. C. (1994). Regressive eye movements and sentence parsing: On the use of regression-contingent analyses. *Memory & Cognition*, *22*, 281–285.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton Jr, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive processes*, *4*(3-4), SI21–SI49.

Sekerina, I. A. (2003). Scrambling and processing: Dependencies, complexity, and constraints. *Word order and scrambling*, *4*, 301–324.

Stowe, L. A. (1986). Parsing wh-constructions: Evidence for on-line gap location. *Language and cognitive processes*, *1*(3), 227–245.

Taylan, A. (2014). Eye-tracking analysis on the processing of turkish complex sentences with wh-phrases. *Hacettepe Üniversitesi Edebiyat Fakültesi Dergisi*, *31*(1).

Thiel, A. (2021). *Effects of prepositional'er'on reading speed in link to long-distance dependencies* (Unpublished doctoral dissertation).

Wagers, M. (2013). Memory mechanisms for wh-dependency formation and their implications for islandhood. *Experimental syntax and island effects*, 161–185.

Zushi, M. (2013). *Long-distance dependencies*. Routledge.

A Target sentences

1. Van de vier zag Hendrik er de hele dag drie in zijn tuin.
2. Van de vijf kreeg Johan er gisteren tijdens school geen van zijn docent
3. Van de acht ontving Jan er een week geleden drie van zijn moeder.
4. Van de drie kocht Joost er eergisteren twee voor zijn oma
5. Bij elkaar gerekend at hij er diezelfde dag drie van de bakker.
6. Bij elkaar gerekend dronk Friso er gister 's avonds vijf op zijn verjaardag.
7. Bij elkaar gerekend stal Elise er tijdens het spelen twee van haar zusje
8. Bij elkaar gerekend krijgt Wilco er een week later dertig als hij aardig doet.
9. In totaal melkt Dagmar er gemiddeld per week dertig als ze hard werkt.
10. In totaal koopt Marte er tijdens de wedstrijd drie wanneer ze daar zin in heeft.
11. In totaal drinkt Hanna er tijdens haar studiepauze twee als ze haast heeft.
12. In totaal schrijft Albert er gemiddeld per dag vijf op als hij in een flow zit.
13. Alles bij elkaar genomen schilderde hij er tijdens zijn studententijd vijf voor zijn grootmoeder.
14. Alles bij elkaar genomen koopt Jens er elke donderdag zeven om zelf op te eten.
15. Alles bij elkaar genomen ziet Jasper er doorgaans 's avonds drie tijdens zijn training.
16. Alles bij elkaar genomen eet Esmee er elk paasweekend acht uit de frituur.
17. Bij elkaar opgeteld vangt Colin er gemiddeld per middag drie als hij gaat vissen.
18. Bij elkaar opgeteld beklimt Jet er tijdens het hoogseizoen vier in Zwitserland.

19. Bij elkaar opgeteld pakt ze er wanneer ze gaat koken twee uit haar keuken.

20. Bij elkaar opgeteld ontvangt Marit er per weekend drie waar ze niks aan heeft.

B Baseline sentences

1. Van de vier zag Hendrik al de hele dag drie vogeltjes in zijn tuin.

2. Van de vijf kreeg Johan al gisteren tijdens school geen stickers van zijn docent.

3. Van de acht ontving Jan al een week geleden drie A4'tjes van zijn moeder.

4. Van de drie kocht Joost al eergisteren twee broodjes voor zijn oma.

5. Bij elkaar gerekend at hij al diezelfde dag drie koekjes van de bakker.

6. Bij elkaar gerekend dronk Friso al gister 's avonds vijf biertjes op zijn verjaardag.

7. Bij elkaar gerekend stal Elise al tijdens het spelen twee knikkers van haar zusje.

8. Bij elkaar gerekend krijgt Wilco al een week later dertig euro als hij aardig doet..

9. In totaal melkt Dagmar al gemiddeld per week dertig koeien als ze hard werkt.

10. In totaal koopt Marte al tijdens de wedstrijd drie colaatjes wanneer ze daar zin in heeft.

11. In totaal drinkt Hanna al tijdens haar studiepauze twee kopjes koffie als ze haast heeft.

12. In totaal schrijft Albert al gemiddeld per dag vijf paragrafen als hij in een flow it.

13. Alles bij elkaar genomen schilderde hij al tijdens zijn studententijd vijf aquarellen voor zijn grootmoeder.

14. Alles bij elkaar genomen koopt Jens al elke donderdag zeven donuts om zelf op te eten.

15. Alles bij elkaar genomen ziet Jasper al doorgaans 's avonds drie supporters tijdens zijn training.

16. Alles bij elkaar genomen eet Esmee al elk paasweekend acht garnalen uit de frituur.

17. Bij elkaar opgeteld vangt Colin al gemiddeld per middag drie snoeken als hij gaat vissen.

18. Bij elkaar opgeteld beklimt Jet al tijdens het hoogseizoen vier bergen in Zwitserland.

19. Bij elkaar opgeteld pakt ze al wanneer ze gaat koken twee tenen knoflook uit haar keuken.

20. Bij elkaar opgeteld ontvangt Marit al per weekend drie kadootjes waar ze niks aan heeft.

C Filler sentences

1. Gisteren heeft hij al een aantal keer boodschappen gedaan bij de lokale supermarkt.

2. Jan schoot per ongeluk de bal door het raam van de overburen.

3. Samen met zijn vrienden speelt Lucas 's avonds twee keer per week padel.

4. Geert woont al vanaf zijn vierde in het mooie Beverwijk met zijn oma.

5. Job eet graag per dag ongeveer drie broodjes belegd met hagelslag.

6. Op een goede dag lukt het Sanne soms wel twee keer een hole-in-one te scoren.

7. Om af te vallen gaat Freek vanaf morgen vijf keer per week sporten in de sportschool.

8. Janne heeft vaak veel moeite met opletten bij haar doorgaans saaie Nederlands lessen.

9. Linde werkt meerdere dagen in de week zodat ze haar dure levensstijl kan betalen.

10. Vanaf jongs af aan vindt Teun het erg leuk om te gamen op zijn computer.

11. Over het algemeen heeft Mila geen goede ervaringen met honden gehad in haar leven.

12. Lisa heeft haar tentamen niet gehaald omdat ze te laat is begonnen met studeren.

13. Manon drinkt gemiddeld op een dag vijf glazen water omdat ze denkt dat dat gezond is.

14. Anne eet het liefst elke ochtend een bakje yoghurt met muesli en rozijnen als ontbijt.

15. Voor haar werk moet Mariet elke dag minimaal anderhalf uur met het openbaar vervoer reizen.

16. Martin kan via een uitzendbureau vaak makkelijk terecht bij meerdere werkgevers in een week.

17. Jimmy gaat over een week het bedrijf van zijn vader overnemen als het bestuur instemt.

18. Op zijn werk is Gerben de hele dag bezig met het sluiten van listige deals om geld te verdienen.

19. Felix spendeert het liefst al zijn vrije middagen dicht bij het Paterswoldsemeer met de eendjes.

20. Janneke heeft de laatste tijd meer moeite om haar rekeningen te betalen omdat ze single is.

21. Door de stijgende energieprijzen kan Marieke haar huur steeds moeilijker op tijd betalen.

22. De presentatie die Amalia had gemaakt was niet op tijd klaar omdat ze slecht gepland had.

23. Elke ochtend gaat Jochem in de hoofdstad naar de kerk om te bidden.

24. Op een grote rots in Peru zoekt Jelle al maanden naar edele metalen.

25. Het verzoek van Edwin om samen te gaan wonen met zijn vriendin werd afgewezen.

26. Na de gefaalde voorbereidingen had Jeroen geen vertrouwen meer in de geluidstechnicus.

27. Joep had veel moeite met het afscheid nemen van zijn vriendjes aan het einde van het schooljaar.

28. Cor heeft zijn koeien in verband met het mooie weer de hele week buiten laten grazen.

29. Door het schandaal op zijn werk moest Govert de hele dag journalisten van zich af slaan.

30. Charlotte mocht door haar zorgvuldige werk bij een programmeeropdracht een uur eerder weg.

31. Sinds een paar jaar investeert Olivia met haar pensioenfonds alleen nog maar in groene bedrijven.

32. Evert is erg succesvol met zijn kunstatelier en verkoopt regelmatig kunstwerken voor hoge prijzen.

33. De knuffel die hij ontving op zijn zesde verjaardag koestert hij nog steeds heel erg.

34. Zonder na te denken gooit Barbara vaak eten weg in plaats van het te bewaren.

35. Over het algemeen houdt Tim niet van chic uit eten gaan als hij met zijn vrienden is.

36. Heinz zijn favoriete bezigheid op de zaterdagmiddag is om rugby wedstrijden van zijn zoon te bekijken.

37. Na tien keer proberen gooide Reinier uiteindelijk toch de handdoek in de ring.

38. Kees wil het liefst een militaire opleiding doen als hij klaar is met zijn middelbare school.

39. Maria ergert zich enorm aan de corruptie die voorkomt in veel Afrikaanse landen.

40. Sinds twee weken is Marieke lid van de nieuwe sporthal bij haar om de hoek.