How Does the Predictability of the Subject in a Sentence Change the Preference for the Use of the Word 'er' in Dutch Language?

Bachelor's Project Thesis

Matthew Melcherts, s3772713, m.b.melcherts@student.rug.nl

Supervisors: Dr S. Jones

Abstract: This thesis focuses on the preference of Dutch language users for the use of the Dutch word 'er' in sentences where it can be omitted. It aims to look at the disparity between Dutch prescriptivists saying that 'er' should be omitted, and Grondelaers et al. (2009) saying it has some sort of 'expectancy marker' function for unexpected sentence subjects. To look further into whether the Dutch word 'er' has such an 'expectancy marker' function in a sentence and measure the preference of Dutch language users, three experiments were conducted. First, to obtain expected sentence subjects, a cloze task was conducted in which the participants filled in an expected sentence subject in a gap sentence. Second, to obtain unexpected sentence subjects which are considered plausible within the sentence, a plausibility experiment was conducted in which the participants ranked sentence subjects based on their plausibility within the sentence. Third, a preference survey was conducted in which participants were asked whether they preferred a sentence with or without 'er', containing either an expected or unexpected sentence subject. The results show no significant difference in preference between sentences with expected and unexpected sentence subjects. Furthermore, the results show that participants were very likely to prefer sentences without 'er' over sentences with 'er' when given the choice. The difference between the results of the preference survey and earlier research is discussed with multiple possible explanations, such as plausibility, relation of 'er' to the verb, the grammaticality of 'er' and the separation of cognitive systems.

1. Introduction

The Dutch word 'er', which literally means 'there' is an unstressed form of the Dutch word 'daar' in origin, as described by Donaldson (2008). The Dutch word 'er' has four distinct pronominal functions in Dutch language. Using the categorisation by Odijk (1993), those functions are existential, locative, prepositional and quantitative functions. Furthermore, many sources describe the distributions of 'er' and its syntactic constraints (Bech, 1952; Bennis, 1986; van Riemsdijk, 1978; Odijk, 1993; Neeleman and van de Koot, 2006; Donaldson, 2008; Grondelaers et al., 2009; Klooster, 2014; Webelhuth and Bonami, 2019), a strong constraint of 'er' is that it occurs only once and maximally twice within a single clause. Also, this means that one instance of 'er' can carry more than one function and 'er' may refer to more than one distinct antecedent.

1.1. Existential 'er'

The existential 'er' is also described as the the presentative 'er' or repletive 'er'. This function of 'er' is used at the beginning of a clause where the subject is indefinite or as the subject of an impersonal passive sentence (Jones, 2020). Existential 'er' can often be replaced with the Dutch words 'daar' or 'hier', 'there' in English. Finally, existential 'er' is often not needed to form a correct sentence, it can often be omitted in Dutch language. See example 1.1 for an existential 'er' sentence with an indefinite sentence subject and see example 1.2 for an impersonal passive sentence with an existential 'er':

Er	*staat*	*een*	*man*	*op*	*de*	*weg.*
There	stands	a	man	on	the	road.

"There stands a man on the road". (1.1)

Er *wordt* *gedanst.*
There is dancing.
"There is dancing." (1.2)

1.2. Locative 'er'

The locative 'er' is used when 'er' replaces a locative phrase, it replaces part of a phrase containing an argument or adjunct. The locative 'er' is replaceable by either 'daar' or 'hier'. If the locative is the only function of 'er', then the 'er' can not occur at the beginning of the sentence, else it can occur anywhere in the sentence. See example 1.3 for a sentence with a locative 'er':

De auto staat ***er*** *geparkeerd.*
The car stands there parked.
"The car stands parked there." (1.3)

1.3. Prepositional 'er'

The prepositional 'er' is also described as the pronominal 'er'. The prepositional 'er' replaces the pronouns 'it' and 'them' which refer to non-humans only. This 'er' is always combined with a preposition like 'erop', 'on it' in English. The prepositional 'er' can always be replaced by 'daar' or 'hier'. Finally, like the locative 'er', if prepositional is the only function of 'er', then it can not occur at the beginning of the sentence, instead it needs to be replaced by 'daar' or 'hier'. See example 1.4 for a sentence with the prepositional 'er':

Ik heb ***er****mee* *de deur geopend.*
I have with-it the door opened.
"I opened the door with it." (1.4)

1.4. Quantitative 'er'

The quantitative 'er' is also described as the partitive 'er'. This 'er' is used with numerals and adverbs of quantity (Donaldson, 2008). It often corresponds to the French word 'en' and in English it means 'of them/it'. The quantitative 'er' is the only function which can not be replaced by 'daar' or 'hier'. Finally, it can not occur at the beginning or end of the sentence. See example 1.5 for a sentence with the quantitative 'er':

Ik at ***er*** *tien.*
I ate of-them ten.
"I ate ten of them." (1.5)

1.5. Background research

In a Dutch writing blog by Van Alst (2016), it is stated that in some cases the use of existential 'er' is unnecessary and sometimes even incorrect. The blog gives examples of Dutch sentences in which this is the case and argues that the 'er' in these sentences has no function.

In contrast, research by Grondelaers et al. (2009) conducted reading experiments using similar sentences. In their research they constructed sentences focussing on the post-verbal variant of existential 'er', which is the variant of existential 'er' which follows immediately after the verb within a clause. Grondelaers et al. combined this post-verbal variant of existential 'er' with sentences containing expected sentence subjects and unexpected sentence subjects to find whether this variant of 'er' would function as a marker for sentence subject predictability, meaning that 'er' would help the reader read unexpected sentence subjects within a clause. With expected and unexpected sentence subjects Grondelaers et al. mean sentence subjects that either follow logically from the context of the sentence or not.

To obtain expected sentence subjects, Grondelaers et al. first conducted an experiment where they asked native Dutch speakers to complete sentence fragments with a missing word or item. From that experiment they obtained nouns that the participants filled in the sentences, which were evaluated on the percentage of participants that completed the sentence fragment with a certain noun. The most often used nouns were deemed as expected sentence subjects by Grondelaers et al. and were then used for their reading experiment.

Grondelaers et al. then selected unexpected sentence subjects from words that were never answered in the sentence fragments, and were matching the expected sentence subjects on length and frequency. These unexpected subjects were considered carefully by Grondelaers et al. in terms of real-world plausibility, such that they did not violate real-world properties.

They found that this variant of existential 'er' had no significant effect on the reading times of the participants when there was an expected sentence subject, but when there was an unexpected sentence subject, then there was a

significant effect on the reading time of the participants, meaning that the participants actually read sentences with unexpected sentence subjects faster if 'er' is present. Therefore, they concluded that the post-verbal variant of existential 'er' had a function as a marker for sentence subject predictability.

However, I find that the way of selecting unexpected sentence subjects based on their plausibility, as done by Grondelaers et al., is somewhat lacking. Grondelaers et al. base their expected sentence subjects on the results of the participants, but not their unexpected sentence subjects. Furthermore, Grondelaers et al. (2009) state in their own results that they did not find a main effect of 'er', or an interaction between subject predictability and the presence of 'er'. This makes me believe that there is more to be found about their conclusion that the post-verbal variant of existential 'er' had a function as a marker for sentence subject predictability.

1.6. Research question and hypothesis

I form two big questions from my background research. First, does the post-verbal variant of 'er' have this predictability marker function? Second, do Dutch speakers prefer the use of this post-verbal variant of existential 'er'? This leads to the research question of this thesis: "How does the predictability of the subject in a sentence change the preference for the use of the word 'er' in Dutch language?"

My hypothesis is mostly formed by the research done by Grondelaers et al. (2009), wherein they show that this variant of existential 'er' increases the reading times of participants when there is an unexpected sentence subject. Therefore, when given sentences with unexpected subjects, Dutch speakers should prefer the use of 'er', more so than in sentences with expected sentence subjects.

2. Methods

To test the hypothesis that native speakers of Dutch prefer the use of the word 'er' in Dutch language when it is followed by an unexpected subject in the sentence, three experiments were conducted. First, a cloze task experiment, which was done to obtain expected subjects in sentences for use in the following experiments. It tries to answer what expected subjects in sentences are, it functions as the first step to finding the difference between expected and unexpected subjects within a sentence.

Second, a plausibility ranking experiment, which was done to determine which subjects gathered from the cloze task experiment were most expected in the sentences and which subjects were least expected in the sentences. Furthermore, the plausibility ranking experiment was also done to ascertain whether the subjects were plausible in the sentence or not. It tries to answer what the difference is between expected subjects and unexpected subjects, and it tries to answer whether the subjects are plausible or not. It functions as the final step to determining the difference between expected and unexpected subjects, and it functions as a method to determine the plausibility of those subjects.

Finally, a preference survey was conducted, which was used to measure the preference in the use of the word 'er' in Dutch language. It tries to answer whether native Dutch speakers prefer the use of the word 'er' in Dutch language when it is followed by unexpected subjects and expected subjects. These unexpected subjects and expected subjects are also referred to as most probable and least probable subjects in the context of the cloze task experiment and the plausibility ranking experiment. It functions as the final experiment which answers the research question of this thesis.

2.1. The cloze task experiment

The cloze task experiment that was conducted consists of 30 similar grammatically structured sentences, in which the post-verbal variant of the existential 'er' could either be used or omitted. A cloze task is a test in which a portion of language has certain items or words removed, participants are then asked to replace the missing item or words. The cloze task sentences used in this experiment have a missing word, which is indicated by a gap. The participants were asked to fill in this gap with a sentence subject that they would expect to fit in the context of the sentence. For example, the following sentence (see example 2.1) was used to assess which sentence subjects fit in the sentence and thus are expected in the sentence.

Onder de	*zee*	*zwemt een*	____	*op*	*jacht*	*naar*	*prooidieren.*
Under the	sea	swim a	____	on	the-hunt	for	prey.

"Under the sea swims a ____ on the hunt for prey."

(2.1)

All sentence subjects provided by participants are considered to be expected sentence subjects in the sentence. The cloze probability was used to measure how probable the provided sentence subjects are in comparison to other provided sentence subjects within the context of the sentence. This cloze probability is measured by the percentage of times participants provided a certain sentence subject. Using the previous example sentence 2.1, the sentence subject "shark" was provided by participants in 80 percent of all the answers. Therefore "shark" has a cloze probability of 0.8, which is 4 times as probable as all the other answers combined, which would have a combined cloze probability of 0.2.

The cloze task experiment was conducted with a small group of 6 participants, all students from ages 21 to 24. Using the results from this cloze task experiment, 4 sentences were filtered out due to not having a sentence subject with a cloze probability above 0.2. Therefore, 26 sentences with expected sentence subjects were obtained, 22 of which had a sentence subject with a cloze probability of at least 50 percent. These 26 sentences and their expected sentence subjects were used in the following plausibility ranking experiment.

2.2. The plausibility ranking experiment

To determine which sentence subjects are most expected, least expected, and whether the sentence subjects are plausible in the context of the sentence (congruent within the real-world context), an experiment was conducted in which participants were asked to sort sentence subjects within one of two groups, then to rank those sentence subjects within the groups. This plausibility ranking experiment seeks to gauge which sentence subjects are considered plausible in the sentence by participants.

The 26 sentences obtained from the cloze task experiment were used together with 5 sentence subjects that could fill the gap within the 26 cloze task sentences. These 5 sentence subjects consisted of the expected sentence subjects provided by participants during the cloze task experiment, as well as randomly chosen subjects with a similar word frequency ranking to the provided sentence subjects from the cloze task, word frequency rankings were compared using the word frequency ranking list by Brysbaert & Keuleers (2010). This word frequency list is based on subtitles from Dutch films and television shows, and is one of the most extensive and freely available databases for the word frequency of Dutch language.

The plausibility ranking experiment displays the 26 cloze task sentences together with the 5 potential sentence subjects that could fill the gap, it then asks the participants to sort the sentence subjects in one of two categories, 'plausible in the sentence' or 'breaks real-world context of the sentence'. Then the participants rank the sentence subjects that were put in the 'plausible in the sentence' category based on the probability of the sentence subjects being in the sentence. For example, using sentence 2.2 with the subjects: "vogel, aap, vliegtuig" (bird, ape, plane), plane could be sorted into the 'breaks real-world context of the sentence' category, bird could be ranked most probable and ape second most probable in the 'plausible in the sentence' category. Thereby ranking sentence subjects on their probability to be in the sentence and sorting them to be plausible or not. The results were used to determine the most probable sentence subjects and the least probable, but still plausible sentence subjects in the sentence.

Op	*de*	*hoogste*	*tak*		*in*	*de*	*boom van*	*de*
	buurman	*zit*	*een*	_____		*te*	*zingen.*	
On	the	highest	branch		in	the	tree of	the
	neighbour	sit	a	_____		to	singing	

"On the highest branch of the neighbour's tree a ____ is singing."

(2.2)

The plausibility ranking experiment was conducted with the same small group of 6 participants as the cloze task experiment. The results of the plausibility ranking experiment were evaluated to determine which sentence subjects were most probable in the sentences, and which sentence subjects were the least probable in the sentence without breaking real-world context too often. A total of 10 sentences were selected for the following preference survey, the selected sentences were those with clear rankings of most probable and least probable sentence subjects, which the participants sorted in the 'plausible in sentence' category in at least 80 percent of the answers.

2.3. The preference survey

The 10 sentences with the most probable and least probable sentence subjects were used to construct 20 pairs of sentences, each pair containing either the most probable subject or the least probable subject, with one of the two sentences using the word 'er' and the other omitting the word 'er'. A latin square design was implemented in which the 20 pairs of sentences were divided over 2 surveys, such that no survey had the same sentence pairs with different sentence subjects, also each survey had 5 pairs with the most probable sentence subjects and 5 pairs with the least probable sentence subjects.

The two surveys both contained the same 30 filler sentence pairs which consisted of sentence pairs utilising the Dutch words "die" and "dat" (both translate to 'that', but are used grammatically differently used in Dutch), sentence pairs utilising "diens" and "zijn" (both translate to 'his', but "diens" is more formally used than "his"), and sentence pairs utilising "welk" and "welke" (both translate to 'which', but are grammatically differently used in Dutch).

The participants were given one of the two surveys at random, then at the beginning of this survey the participants were instructed to read the sentence pairs out loud and indicate which of the two sentences they preferred. The sentence pairs were shown to the participants in random order, with the sentences in the sentence pair also appearing in random order.

The survey was distributed online using qualtrics.com and responses were gathered from 145 participants. Each of these 145 participants were native Dutch speakers ranging from ages 20 to 80. From the 145 participants about 37 were excluded because of response recording issues, and 24 were excluded because of the survey not being fully completed by the participants. The remaining 84 responses were used to obtain the results of this thesis.

3. Results

The results of the preference survey are shown in Figure 3.1, the figure shows on its x-axis the frequency that participants preferred the use of the word 'er' in Dutch language. The y-axis shows the amount of participants that answered that they did prefer the use of the word 'er' in that frequency in Dutch language in the 10 sentence pairs.

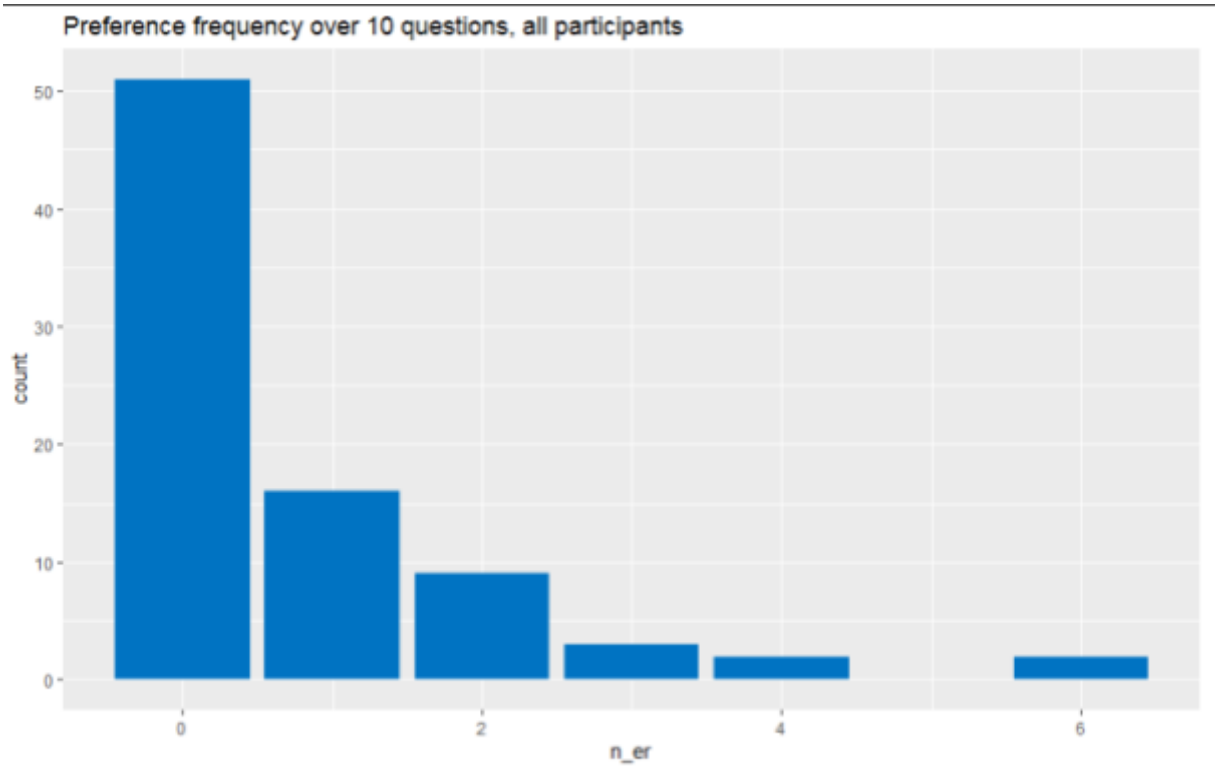

Figure 3.1: Plot of the preference frequency for using the post-verbal variant of existential 'er' for all participants.

Over all the individual participants there is a mean of 9.240964 times that they preferred to not use the word ‘er’ in the 10 sentence pairs of the survey. Of the 84 participants, 2 preferred the use of ‘er’ 6 times, 2 preferred the use of ‘er’ 4 times, and 3 preferred the use of ‘er’ 3 times. The

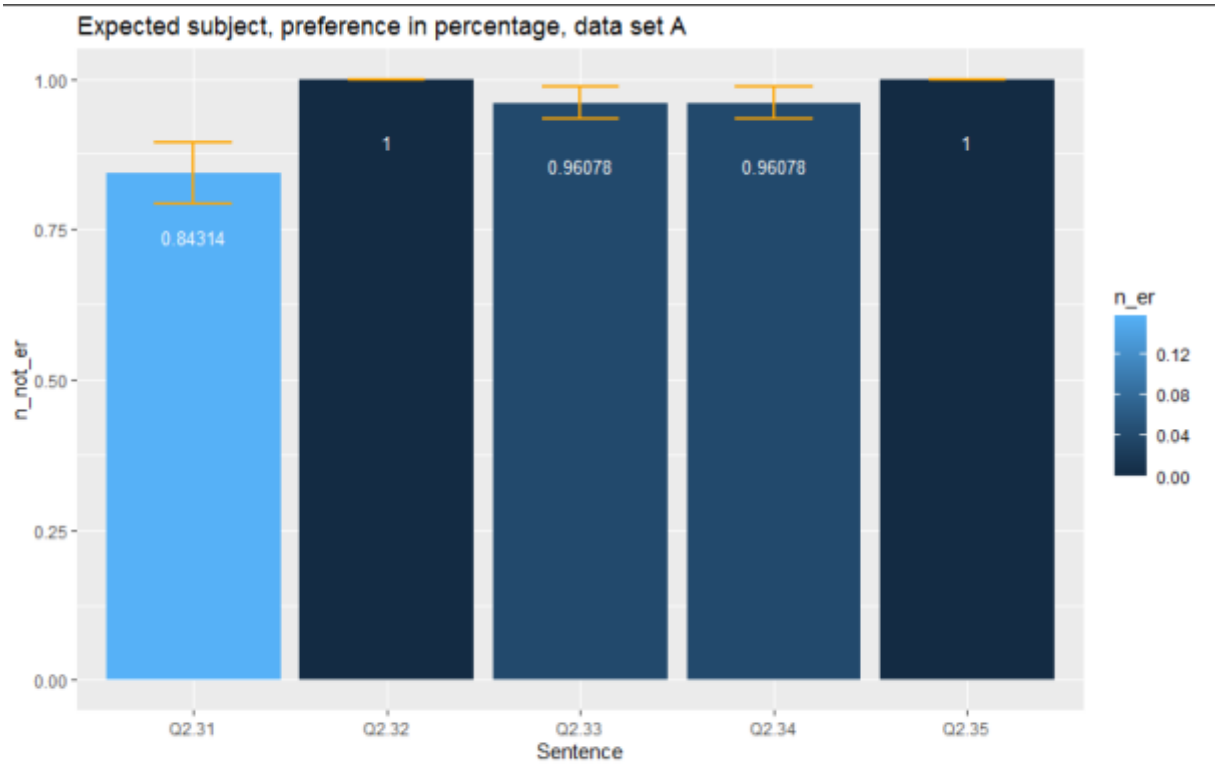

Figure 3.2: Preference in percentage for the first five sentences with expected sentence subjects.

Figure 3.3: Preference in percentage for the first five sentences with unexpected sentence subjects.

remaining 77 participants preferred the use of ‘er’ 2 times or less.

The results of the preference survey for the first five sentence pairs containing ‘er’ are shown in Figures 3.2 and 3.3. Figure 3.2 shows the preference for the use of the word ‘er’ for all participants in the first five sentence pairs containing ‘er’ with an expected subject, standard error is shown by the orange error bars. On the y-axis, the figure shows the preference in

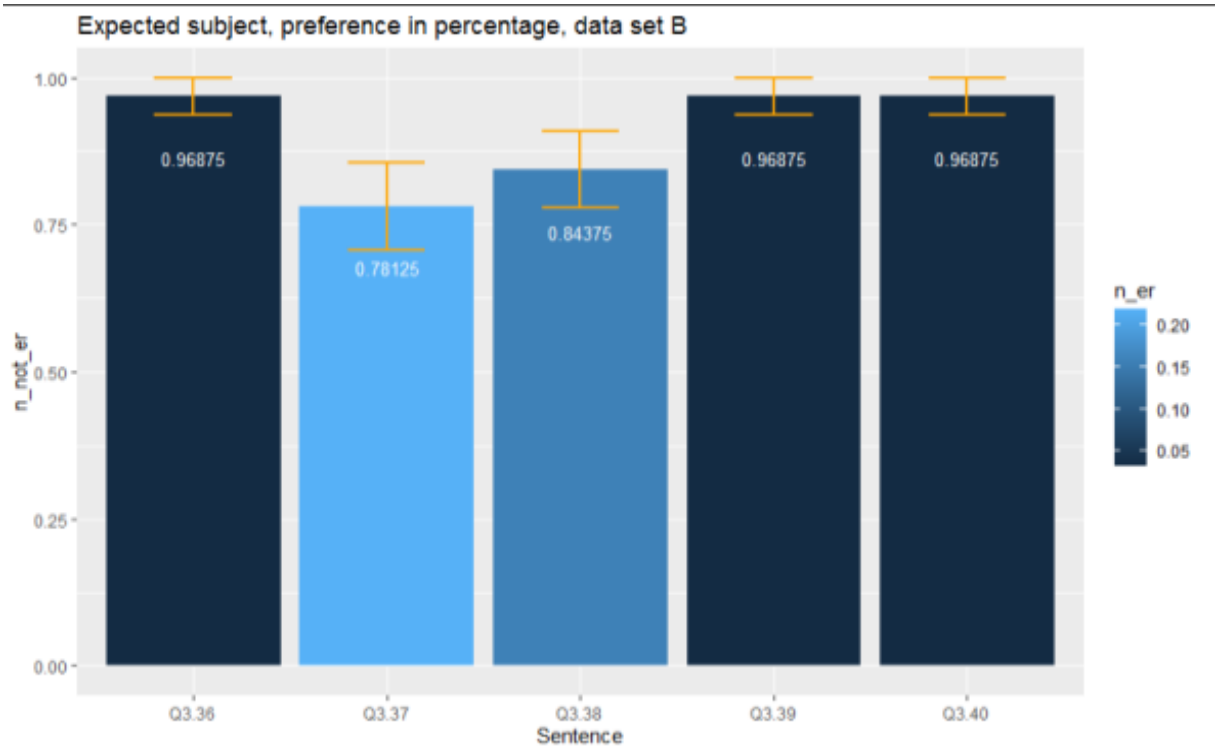

Figure 3.4: Preference in percentage for the last five sentences with expected sentence subjects.

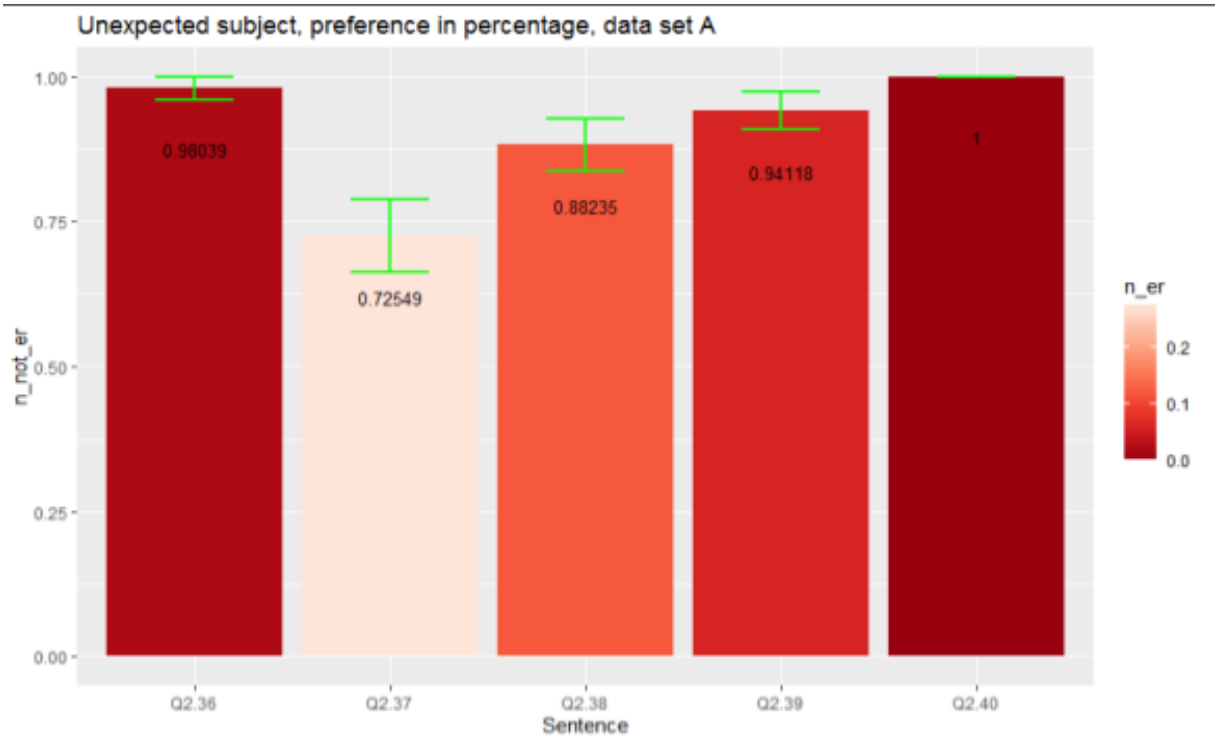

Figure 3.5: Preference in percentage for the last five sentences with unexpected sentence subjects.

percentage of the participants for not using ‘er’, with the percentage also being shown by the number within each bar in the plot. The brightness of blue of each bar in the figure shows the preference in percentage of the participants for using ‘er’.

Figure 3.3 shows the preference for the use of the word ‘er’ for all participants in the first five sentence pairs containing ‘er’ with an unexpected subject, similar to Figure 3.2, except it uses green for the error bars and the brightness of red is used in each bar in the figure to show the preference in percentage of the participants for using ‘er’.

The results of the preference survey for the last five sentence pairs containing ‘er’ are shown in Figures 3.4 and 3.5. Figure 3.4 shows the preference for the use of the word ‘er’ for all participants in the last five sentence pairs containing ‘er’ with an expected subject. Figure 3.4 is plotted similarly to Figure 3.2, it uses orange for the error bars and the brightness of blue is used in each bar in the figure to show the preference in percentage of the participants for using ‘er’.

Figure 3.5 shows the preference for the use of the word ‘er’ for all participants in the last five sentence pairs containing ‘er’ with an unexpected subject. Figure 3.5 is plotted similarly to Figure 3.3, it uses green for the error bars and the brightness of red is used in each bar in the figure to show the preference in percentage of the participants for using ‘er’.

When evaluating the error bars for each sentence pair, there are two sentences of which the sentences with the expected sentence subjects and the sentences with unexpected sentence subjects fall outside the error bars. These are sentence pairs 32 and 35. Since they fall outside the error bars, it means that there is a significant difference between these sentences with expected sentence subjects and unexpected sentence subjects.

To evaluate the difference between sentences with expected sentence subjects and sentences with unexpected sentence subjects a paired t-test was used. The results of the paired t-test show that there is not a significant difference between the sentences with expected sentence subjects and sentences with unexpected sentence subjects (t = 1.1829, p-value = 0.2672).

Table 3.1: Percentage of preference over all questions of all participants based on the subject expectability and presence of ‘er’.

Subject	**‘er’**	**Not ’er’**
Expected	0.07040441	0.92959559
Unexpected	0.08455882	0.91544118

Table 3.1 shows the percentage of preference over all questions with the conditions of sentence subject expectability and the presence of ‘er’. From it we can see how small and insignificant the difference between the sentence having an expected sentence subject and the sentence having an unexpected sentence subject is. A Wilcoxon signed rank test was used to evaluate whether there is a significant difference between the results of the sentences with expected sentence subjects and sentences with unexpected sentence subjects. The results of which show that there is no significant difference between the two sentence variables (V = 36, p-value = 0.4145).

To evaluate what influenced the preference of the participants in the survey, a model was constructed consisting of variables based on the following: the preference, the sentence, whether ‘er’ occurs in the sentence, and whether the sentence has an expected or unexpected sentence subject. The model was then evaluated using an one-way analysis of variance test, see Table 3.2, which indicated that there was a significant effect on the preference of the participants only for whether the sentence contains the word ‘er’ or not (F = 1138, p < 2e-16).

Table 3.2: Summary of one-way ANOVA test.

	Df	**Sum Sq**	**Mean Sq**	**F**	**Sig**
‘er’	1	7.141	7.141	1138	<2e-16
Question	1	0	0	0	1
Expected	1	0	0	0	1
Residuals	36	0.226	0.006		

An one-way analysis of variance chi square test was also used to evaluate the model, see Table 3.3, it lead to the same conclusion that only the sentence containing the word ‘er’ has a significant effect on the preference of the

participants (df = 37, p < 6.612e-09). Since there is no significant effect of the expectedness of the sentence subject on the preference for the use of 'er', it means that my hypothesis that unexpected sentence subjects would increase the preference for the use of 'er' over sentences with expected sentence subjects is false.

Table 3.3: Summary of one-way ANOVA chi square test.

	Df	Deviance Resid.	Df	Resid. Dev	Sig
NULL			39	36.711	
Question	1	0.000	38	36.711	1
'er'	1	33.646	37	3.065	6.612e-09
Expected	1	0.000	36	3.065	1

4. Conclusion and discussion

From Figure 3.1 we can see that, in most cases, the participants in this survey answered that they did not prefer the use of 'er' when given the option between a sentence with and without 'er'. From looking at this figure and its mean, we can conclude that when given the choice between a sentence with the post-verbal variant of existential 'er' and the same sentence without the post-verbal variant of existential 'er', the participants generally prefer the sentence without the use of the post-verbal variant of existential 'er'.

From Figures 3.2, 3.3, 3.4 and 3.5, we can see that there were only three sentences which had the use of 'er' preferred in more than 10 percent of the answers. Since all sentences used in the survey were similarly structured grammatically, the only difference seems to be in the words and their relations within the sentence. However, it is notable that the difference in these three sentences is very unlikely to be due to the sentences having an expected sentence subject or an unexpected sentence subject, because the results show that the difference in preference for the use of the word 'er' between the three sentences with expected sentence subjects and the three sentences with unexpected sentence subjects is small.

Furthermore, the figures show that in half of the sentences, the preference for not using the word 'er' of the participants is higher with an unexpected sentence subject than a sentence with an expected sentence subject. However, the differences in preference for not using the word 'er' between the sentences with expected sentence subjects and sentences with unexpected sentence subjects are mostly insignificant. The largest difference in preference is 6.25 percent and the figures show two sentences which have this difference in preference for not using the word 'er' of 6.25 percent, these sentences are also the only sentence pairs of which the results from the expected sentence subjects and the unexpected sentence subjects fall outside the error bars. Notably, these two sentences are the only sentences which share the Dutch verb for sleeping, aside from sharing the same grammatical structure, the sentences have nothing else in common.

As noted before, in Figures 3.2, 3.3, 3.4 and 3.5, we can see that the differences in preference for not using the word 'er' between the sentences with expected sentence subjects and sentences with unexpected sentence subjects are insignificant, most fall within the error bars. For these sentences, it might mean that the preference for the use of the word 'er' between sentences is not due to whether the sentence contains an expected or unexpected sentence subject, rather the difference in words used in the sentence might explain the difference of preference for the use of the word 'er' between sentence pairs. This is further strengthened by the statistical tests, which tell us that whether the sentence contains 'er' or not is the only significant difference within a sentence pair.

From the results and conclusions it becomes clear that the hypothesis that there would be a higher preference for the use of the word 'er' in sentences with unexpected sentence subjects than sentences with expected sentence subjects is false. From the results and conclusions it becomes clear that the predictability of the sentence subjects has little influence on the preference of the use of the word 'er' in the sentence. Furthermore, it seems that the words used within the sentence might have more of an influence on the preference of the use of the word 'er' in the sentence than the predictability of the sentence subject.

A future study might attempt to find a relationship between certain words and the

preference for the use of the word 'er'. It could start with an analysis of Dutch language texts which use 'er', then using the analysis, develop a list of words which are often used with 'er'. Using such a list, the study could research whether some words lead to a higher preference for the use of 'er'. More specifically, since the preference for the use of the post-verbal variant of existential 'er' could be influenced from the verb in the sentence after which 'er' follows, future studies could experiment with different verbs and research whether there is a strong influence from the verb in the sentence opposed to the sentence subject.

There is a difference between the preference for the use of the post-verbal variant of existential 'er' and the research by Grondelaers et al. (2009). This difference could be explained by the measure of plausibility within my experiment compared to the research by Grondelaers et al. The research by Grondelaers et al. selected unexpected sentence subjects matching their expected sentence subjects in both length and frequency, they state in their research that the unexpected sentence subjects were carefully considered in terms of real-world plausibility, in an example they gave they say that a reading room of a library could contain a trout, instead of an atlas. Whether this example is plausible or not is why this thesis conducted the second experiment, the plausibility ranking experiment, which might have led to sentence subjects that are more plausible than the examples given by Grondelaers et al. Thereby leading to unexpected sentence subjects that are more expected in comparison to the unexpected sentence subjects in the research by Grondelaers et al.

From the results of my second experiment, the plausibility ranking experiment, it becomes clear that the participants had different views on which sentence subjects were plausible within the sentence and which sentence subjects were not plausible within the sentence. This led to a sort of ranking based on how often a sentence subject was answered to not be plausible within the sentence. A future study could expand on this, experimenting with a larger list of sentence subjects and asking a large number of participants to state which sentence subjects are plausible within a sentence. Based on my results, it should create some sort of gradient list of the plausibility of the sentence subjects. This list could then be further experimented with, for example, the list could be used in a reading experiment to see how plausibility of the sentence subject influences the effect of 'er' on reading speed as discussed by Grondelaers et al. (2009).

Another explanation to why there is a difference between the preference for the use of the post-verbal variant of existential 'er' and the research by Grondelaers et al. (2009), is discussed in one of Grondelaers recent research papers (2020). In which Grondelaers discusses how the use of the word 'er' is often considered ungrammatical in Dutch language even though experimental data and corpus data show that the use of 'er' is grammatical and even beneficial to Dutch speakers. Furthermore, Grondelaers discusses the differences in the use of 'er' between Netherlandic and Belgian Dutch. In Belgian Dutch, the use of 'er' is considered to be grammatical, but more notably, the benefit of the post-verbal variant of existential 'er' differs from Netherlandic Dutch. The beneficial effect on reading from 'er' in Belgian Dutch is a processing advantage in the subject noun phrase, whereas in Netherlandic Dutch, the processing advantage is for anything that follows from the sentence. As Grondelaers describes it, 'er' has a late effect in Netherlandic Dutch.

A future study might further on the research by Grondelaers (2020), the survey in this thesis only had Netherlandic Dutch speakers, therefore doing something similar to the survey from this thesis to measure the difference in preference between Netherlandic Dutch speakers and Belgian Dutch speakers could tell more of the differences within Dutch language. Furthermore, since the effects of 'er' are different between the Netherlandic Dutch and Belgian Dutch, an experiment could be designed to test the limits of this difference within Dutch language. Since the effect of 'er' in Belgian Dutch is in the processing advantage of the noun phrase, it might be more telling to experiment with the predictability of sentence subjects using Belgian Dutch speakers.

As previously described, Grondelaers (2020) describes a late effect of 'er' in Netherlandic Dutch. Which makes for an

interesting future study in which this late effect of 'er' is experimented with. How far does this late effect of 'er' reach, does it carry over sentences? Does the effect retain its influence over the following sentences or does it diminish? This could lead to a better understanding of the effects of the word 'er' within Netherlandic Dutch.

In a paper by Lewis and Phillips (2015), they discuss the question of whether grammatical theories and language processing models describe separate cognitive systems. In the paper they argue mostly in favor of a one-system view, but when we view grammatical theories and language processing models as separate cognitive systems, then the differences between the reading experiments by Grondelaers et al. (2009) and the results from the preference survey might be explained on a cognitive level. Lewis and Phillips discuss in their paper two groups, one focusing on offline data, which are judgements made under ideal conditions, and the other focusing on online data, which are more time-sensitive and linked to reading.

Using the explanation from Lewis and Phillips, the explanation for the difference between the reading experiments by Grondelaers et al. (2009) and the results from the preference survey might be how data was obtained from the participants. The reading experiment by Grondelaers et al. (2009) is closer to the online language processing group described by Lewis and Phillips, while the preference survey is closer to the offline grammatical theories group. Although Lewis and Phillips argue that most evidence is consistent with the one-system view, the difference between the reading experiment by Grondelaers et al. (2009) and the preference survey might provide some proof that there are two separate cognitive systems.

Finally, the survey conducted in this thesis showed us that speakers of Netherlandic Dutch do not like to use 'er' when they have the choice not, which is likely because the use of 'er' in such a way is considered ungrammatical even though it does have a beneficial effect for Dutch readers (Grondelaers, 2020). Furthermore, the survey showed no significant influence for the predictability of the sentence subject on the preference of using 'er'. However the preference of 'er' might be influenced by the use of verbs within the sentence, as well as the plausibility of the sentence subject.

References

Bech, Gunnar. 1952. *Über das niederländische Adverbialpronomen er. Traveaux du Cercle Linguistique de Copenhague* 8, 5–32.

Bennis, Hans. 1986. *Gaps and dummies*. Dordrecht: Foris.

Brysbaert, M. & Keuleers, E. (2010, 17 June). Hoe vaak worden Nederlandse woorden gebruikt? Geraadpleegd van http://crr.ugent.be/archives/287

Donaldson, B. (2008). *Dutch: A comprehensive grammar*. Routledge.

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*, 153-160.

Grondelaers, S. (2020). Verwacht het onverwachte: Over hoe Nederlanders er anders gebruiken dan Vlamingen. *Nederlandse Taalkunde*, *25*(2-3), 205-211.

Jones, S. (2020) Multifunctional Dutch er. Proceedings of the LFG'20 Conference, On-Line, 148–168.

Klooster, Wim G. 2014. Reer. In Freek Van de Velde, Hans Smessaert, Frank Van Eynde and Sarah Verbrugge (eds.), *Patroon en Argument: een dubbelfeestbundel bij het emeritaat van William van Belle en Joop van der Horst,* pages 255–270, Universitaire Pers Leuven.

Lewis, S., & Phillips, C. (2015). Aligning grammatical theories and language processing models. *Journal of Psycholinguistic Research*, *44*(1), 27–46. https://doi.org/10.1007/s10936-014-9329-z

Neeleman, Ad and van de Koot, Hans. 2006. Syntactic haplology. In Martin Everaert and Henk van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, volume 1, Chapter 60, pages 687–710, Oxford: Blackwell.

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

van Alst, B. (2016, 17 Juni). Het woordje 'er': wanneer mag het worden gebruikt? Retrieved from https://www.scribbr.nl/taalregels-schrijftips/het-woordje-er/

van Riemsdijk, Henk. 1978. *A case study in syntactic markedness*. Dordrecht: Foris.

Webelhuth, Gert and Bonami, Olivier. 2019. Syntactic haplology and the Dutch proform "er". In Stefan Müller and Petya Osenova (eds.), *Proceedings of the 26th International Conference on Head-Driven Phrase Structure Grammar*, pages 100–119, Stanford, CA.: CSLI Publications.

A. Appendix

The sentences used in the preference survey and their variants are as follows:

Q2.31: Onder de zee zwemt (er) een haai op jacht naar prooidieren.
Q3.31: Onder de zee zwemt (er) een jager op jacht naar prooidieren.
Q2.32: Onder de brug in de stad slaapt (er) een zwerver.
Q3.32: Onder de brug in de stad slaapt (er) een maniak.
Q2.33: In de hoogste toren van het kasteel verblijft (er) een prinses.
Q3.33: In de hoogste toren van het kasteel verblijft (er) een senator.
Q2.34: In de brandende flat dooft (er) een brandweerman het vuur.
Q3.34: In de brandende flat dooft (er) een monteur het vuur.
Q2.35: Op de stoel van mijn vader slaapt (er) een zwarte kat.
Q3.35: Op de stoel van mijn vader slaapt (er) een zwarte kip.
Q2.36: Aan de andere kant van de vijver zit (er) een groene spin.
Q3.36: Aan de andere kant van de vijver zit (er) een groene kikker.
Q2.37: Aan de andere kant van het speelbord wordt (er) een gele tovenaar verplaatst.
Q3.37: Aan de andere kant van het speelbord wordt (er) een gele pion verplaatst.
Q2.38: In de Atlantische Oceaan duikt (er) een man op zoek naar de Titanic.
Q3.38: In de Atlantische Oceaan duikt (er) een onderzeeër op zoek naar de Titanic.
Q2.39: In de donkere blauwe lucht zweeft (er) een grijze parachute.
Q3.39: In de donkere blauwe lucht zweeft (er) een grijze wolk.
Q2.40: Op het weiland staat (er) een vlag met zwarte en witte vlekken.
Q3.40: Op het weiland staat (er) een koe met zwarte en witte vlekken.