



THE INFLUENCE OF VERB LOCATION ON DUTCH SENTENCES USING ‘ER’

Bachelor’s Project Thesis

Niels Burgler, s3665828, n.burgler@student.rug.nl,

Supervisor: Dr S. Jones

Abstract: In this project, Dutch sentences using the word ‘er’ are examined. The Dutch word ‘er’ is a pronoun that can have multiple functions at once. For every function that ‘er’ takes, some information is removed from the sentence. This lack of information makes people dislike the sentences. In some sentences where ‘er’ takes multiple functions, the verb still holds a lot of information. This raises questions about the influence of the verb on sentence acceptability. A syntactic acceptability survey was conducted, where participants were asked to rate certain sentences. The sentences varied in verb location, pronominal function of er, and grammatical functions of er. A logistic regression of the data revealed that the verb location only has an effect on acceptability in certain types of sentences. This effect seems to be caused by the words that end these sentences. This brings us one step closer to understanding how usage of ‘er’ interacts with sentence structure.

1 Introduction

Er is a Dutch pronoun that can be used in sentences in four distinct ways. These different pronominal functions of *er* can be combined and used multiple times in a single occurrence of *er*. No proper substitute of *er* exists in English. At times, *er* can be translated as ‘there’, at other times, *er* is not translated at all. There are multiple accounts of which pronominal functions *er* has. In this research, four different pronominal functions will be used (Odiijk, 1993; Haeseryn et al., 1997; Donaldson, 2008; Bennis, 2011; Jones, 2020). These are referred to as existential er_X , locative er_L , prepositional er_P , and quantitative er_Q .

Many different analyses of *er* have been done (De Schutter, 1992; Odiijk, 1993; Haeseryn et al., 1997; Donaldson, 2008; Grondelaers et al., 2009; Bennis, 2011; Grondelaers, 2020; Jones, 2020). Despite this, much remains unknown about how *er* is actually used. This research project aims to do an exploratory analysis of *er* for different pronominal and grammatical functions of *er*, with a focus on the location of the verb in the sentence, in order to gain further understanding about *er*.

1.1 Existential er_X

Existential er_X is used in combination with an indefinite subject (1) or in passive sentences (2). Some authors categorize these as two separate pronominal functions of *er* (Voortman, 2005). Existential er_X is optional in certain sentences, such as sentences with a time clause at the start (3) (Haeseryn et al., 1997).

- (1) *Er_X loopt een vrouw op straat.*
ER walks a woman on street.
“There’s a woman walking down the street.”
- (2) *Er_X wordt gefeest.*
ER becomes partied.
“There is partying.”
- (3) *Morgen ruimt (er_X) een ouder*
Tomorrow clears (ER) a parent
alles op.
everything up.
“Tomorrow a parent will clean up everything.”

Newspapers tend to leave out meaningless *ers* such as in (3), which can result in ungrammatical

sentences. Existential er_X can sometimes be translated with ‘there’ or ‘hier’ ((1), (2)), but not always (3).

1.2 Locative er_L

Locative er_L refers to a location and is a less specific counterpart to *daar/hier* (there/here). Locative er_L is always translated with ‘there’ or ‘here’.

- (4) *De vrouw loopt er_L.*
The woman walks ER.
“The woman walks there.”

1.3 Prepositional er_P

Prepositional er_P is used in combination with a preposition when the preposition refers to an object. Prepositional er_P can be combined with the preposition to form a single word (5), but er and the preposition can be split if necessary (6). In English prepositional er_P will often be translated with ‘it’ ((5), (6)), but can also be translated with words like ‘thereon’. Prepositional er_P can be replaced by *daar/hier* (7).

- (5) *Ik zit erop.*
I sit ER.on.
“I sit on it.”
- (6) *Ik zit er_P vaak op.*
I sit ER often on.
“I often sit on it.”
- (7) *Ik zit daar vaak op.*
I sit there often on.
“I often sit on that.”

1.4 Quantitative er_Q

Quantitative er_Q is used in combination with a quantity. Quantitative er_Q is not translated (8) and cannot be replaced by *daar/hier* (9). Usage of quantitative er_Q varies heavily per individual, and is mostly determined by geographical location (De Schutter, 1992; Haeseryn et al., 1997).

- (8) *Ik heb er_Q twee.*
I have ER two.
“I have two (of it).”

- (9) **Ik heb daar twee.*
I have there two.
(intended) “I have two (of it) over there.”

1.5 Multifunctional er

As mentioned before, er can have multiple pronominal functions at once in a single sentence (10). In those cases, only one er occurs in the sentence (11). The only exception to this rule occurs with the combination of the existential and quantitative er when the existential er_X is at the start of the sentence (12). All pairwise combinations of pronominal functions are possible, and er can also hold one pronominal function multiple times (13).

- (10) *Er_XL wonen mensen.*
ER live people.
“People live there.”
- (11) **Er_X wonen er_L mensen.*
ER live ER people.
(intended) “People live there.”
- (12) *Er_X lopen er_Q twee op straat.*
ER walk ER two on street.
“Two are walking down the street.”
- (13) *Hij haalt er_PP een haak mee uit.*
He takes ER a hook with out.
“He takes a hook out of it with it.”

1.6 Er and information

It is important to realise that a sentence with er always has less information than a sentence without er . This can clearly be seen when comparing sentence (5) with (14). The sentences are identical, except for the addition of a prepositional er_P in sentence (5). While it is clear in sentence (14) that the person is sitting on a chair, this is impossible to know from sentence (5). The addition of an er to a sentence thus leads to a loss of information.

- (14) *Ik zit op een stoel.*
I sit on a chair.
“I sit on a chair.”

In small sentences, like (5) and (14), people often do not mind this loss of information. After all, the

necessary information is often implied from context. With larger sentences (15), however, the lack of information becomes too big to overcome and people start disliking the sentence (Barf, 2021).

- (15) *Ik zag dat er_{XQPPP} vijf mee voor
I saw that ER five with for
naartoe zijn gereden.
to are driven.
“I saw that five have driven with it for it to
it.”*

1.7 *Er* and verbs

In sentences where *er* refers to both the object and the subject, the verb can give a lot of information. This can be seen in sentences (16) and (17), which are identical except for the verb that is used. This change not only completely changes the meaning of the sentence, but the possible antecedents of *er* as well. The antecedents of *er* are the words or word groups *er* refers to (Crystal, 2011). Pronouns typically occur after the antecedent. Every pronoun has an antecedent, and whenever a pronoun is encountered in a sentence the corresponding antecedent is necessary to correctly interpret the pronoun.

- (16) *Ik zag dat er_{XQQ} vijf drie kochten.
I saw that ER five three bought.
“I saw that five bought three.”*
- (17) *Ik zag dat er_{XQQ} vijf drie voerden.
I saw that ER five three fed.
“I saw that five fed three.”*

In sentences (16) and (17), *er* has two antecedents. It is unknown what these antecedents are, but we can assume they originate from some unknown context. Despite the exact antecedents being unknown, we can still determine from the information given by the verb what the antecedents could be.

For sentence (16), the first antecedent is someone or something that is capable of buying things. This could be a grandfathers, a child, a customer, or even a robot. While the possibilities seem endless, there are also many things the first antecedent cannot be, such as a plant or a closet, since plants and closets are incapable of buying things. The second antecedent is something or someone that can be

bought. This could be a car, a plant, furniture, or even a person in some contexts.

For sentence (17), the first antecedent is someone or something that can feed another someone or something, This could be a person, a farmer, a robot, basically the same things as the first antecedent in sentence (16). For the second antecedent, however, the possibilities are different compared to sentence (16). The second antecedent in this sentence can be a person or a plant, but not a car or furniture such as in (16). This illustrates how the verb influences the possible antecedents of *er*, and consequently the amount of information in the sentence.

When a pronoun without a clear antecedent is encountered, it becomes hard to interpret the sentence. By using the information in the sentence, the antecedent can sometimes be deduced. In sentences (16) and (17) possible antecedents could be deduced using the verb. The verb in these sentences is, however, located at the end. Since antecedents are expected to be before the pronoun, placing the verb before the pronoun could be beneficial to determining the possible antecedents.

It is also possible for *er* to only refer to the object (18) or the subject (19) in such sentences. In sentences where *er* only refers to the object or the subject, the verb influences the information in the sentence less, but still adds information about the remaining antecedent of *er*.

- (18) *Ik zag dat er_{XQ} vijf mensen drie
I saw that ER five people three
voerden.
fed.
“I saw that five people fed three.”*
- (19) *Ik zag dat er_{XQ} vijf drie dieren
I saw that ER five three animals
voerden.
fed.
“I saw that five fed three animals.”*

1.8 Research Question(s)

The goal of this research is to examine how the verb influences sentence acceptability in Dutch sentences using *er*. It seems that *er* influences the amount of information in a sentence, which in turn influences the acceptability of the sentence. On the

other hand, the verb also influences the information in the sentence, but it is unclear whether this also influences the sentence acceptability.

There are several ways to examine the verb in Dutch *er*-sentences. One could manipulate the specificity or tense of the verb. In this research, the location of the verb will be examined. The research question that will be (attempted to be) answered is the following:

Does the verb location influence acceptability? The verb location will be examined for different grammatical functions of *er* and different pronominal functions of *er*. This results in the following sub-questions:

Is this effect different for different grammatical functions of *er*? and **Is this effect different for different pronominal functions of *er*?** By examining the different grammatical and pronominal functions of *er*, the following questions can also be examined:

Is there a preference regarding the grammatical functions of *er*?, **Is this preference different for different pronominal functions of *er*?**, and **Is there a preference regarding the pronominal functions of *er*?**

1.9 Expectations

First, while the verb influences the amount of information in the sentence, the verb location does not. The verb location only influences where in the sentence the information will be presented. My expectation is that it is preferred to have the verb as early as possible in the sentence, so that as much information as possible is known before encountering the *er* in the sentence. This should make it easier to find a possible antecedent of *er* if necessary. Such an effect would be bigger when *er* takes more functions in a sentence, since the verb would then give more information.

Second, between the different grammatical functions of *er*, it would be preferred that *er* takes as little functions as possible. If a decision had to be made between *er* referring to the object or *er* referring to the subject, it is likely that *er* referring to the object would be preferred, since the object often plays a less important role in the sentence. According to Haeseryn et al. (1997), a sentence consists of a subject and (at least) a predicate. The predicate contains at least the verbs, but can also contain

the object if necessary and can additionally contain other clauses. This structure shows how subjects are always necessary in Dutch sentences, while objects are not.

Third, between the different pronominal functions of *er*, there is probably no preference. All pronominal functions of *er* are grammatical and should not influence the sentence acceptability differently.

2 Methods

2.1 Participants

81 people participated in the experiment. Their age and sex were not recorded. The only requirement for participating in the experiment was having Dutch as a first language. The participants were all volunteers who were asked to participate through a message that was spread via social media. Participants were encouraged to spread the message to their friends and family. No compensation was offered for participating in the experiment. Every participant was assigned to the same group, since only one group was used in the experiment.

2.2 Stimuli

Dutch sentences using *er* were used. These sentences were made to be as distinct as possible from each other, while still following the required structure. Sentences could be categorized by the pronominal functions *er* can take in the sentence, the location of the verb, and the grammatical functions of *er*.

In every sentence, an existential er_X was present. In addition, *er* could either take up to two quantitative functions, or up to one quantitative and one prepositional function. This leads to two sentence types, referred to as XQQ and XQP respectively. This variation enables us to see the influence a prepositional er_P can have on sentence acceptability.

- (20) *Soms schoppen er_{XQQ} twee één.*
 Sometimes kick ER two one
 “Sometimes, two kick one.”

- (21) *Vorig jaar vloegen er_{XQP} twee tegenaan.*
 Last year flew ER two against
 “Last year, two flew into it.”

As can be seen from the pronominal functions of *er*, (20) is an XQQ sentence, while (21) is an XQP sentence. Sentence (20) will still be called an XQQ sentence, even if there is only an *er_X* or an *er_{XQ}* in the sentence. The most important difference between these sentences is the prepositional functions *er* can take (or in this case, takes), of which one changes from a quantity to a preposition.

The main verb can be located in the front or at the end of the sentences. The main verb refers to the verb that is relevant to the meaning of *er*. This variation in verb location enables us to examine what influence the verb location has on sentence acceptability.

- (22) *Over een jaar krijgen er_{XQQ} veertig twee.*
 Over a year get ER forty
 two
 “In a year, forty will get two.”

- (23) *Wij hopen dat er_{XQQ} veertig twee krijgen.*
 We hope that ER forty two
 get
 “We hope that forty will get two.”

In sentences (22) and (23) the variations in verb location can be seen. Sentences with the main verb (**krijgen**, in this case) in the front, such as (22), will be referred to as FRONT sentences, while sentence with the main verb at the end, such as (23), will be referred to as END sentences. To accommodate for the change in verb location, the rest of the sentence has to change as well. FRONT sentences have time words at the start of the sentence, while END sentences have a pronoun and a verb with a that-clause at the start.

Finally, the grammatical functions of *er* in the sentence can change. The grammatical functions *er* can take in the sentences are the object and the subject. By adding this variation the influence of grammatical function of *er* on sentence acceptability can be examined.

- (24) *Soms schoppen er_X twee jongens één muur.*
 Sometimes kick ER two boys
 one wall
 “Sometimes, two boys kick one wall.”

- (25) *Soms schoppen er_{XQ} twee jongens één.*
 Sometimes kick ER two boys
 one
 “Sometimes, two boys kick one.”

- (26) *Soms schoppen er_{XQ} twee één muur.*
 Sometimes kick ER two one
 wall
 “Sometimes, two kick one wall.”

In sentence (24), both the subject, *jongens*, and the object, *muur*, can be seen since *er* refers to neither. In sentence (25), *er* refers to the object, but not to the subject. In sentence (26), the opposite is the case. *Er* can also refer to both the object and the subject, which can be seen in sentences (20) to (23). The grammatical functions of *er* will be referred to with *SUBJ* \pm and *OBJ* \pm . For example, sentence (25) is *SUBJ* $-$ and *OBJ* $+$

Since the variables are mutually exclusive, a total of 16 different sentence types exist. The sentences were often grouped as XQQ FRONT, XQQ END, XQP FRONT, and XQP END. Each of these groups had 4 sentences that were distinct in meaning. These 4 sentences each had 4 ‘variants’, in which the grammatical function of *er* differed. Sentences (20), (24), (25), and (26) form one of those ‘sets’. This means that, for every one of the 16 combinations, 4 sentences exist. This leads to a total of 64 experiment sentences. In addition to these experiment sentences, 64 filler sentences were used.

The filler sentences can be split into 3 groups. The first group, GOOD, consisted of sentences that are grammatically correct and semantically common. The second group, BAD, consisted of sentences that were translated word for word from English to Dutch and are thus grammatically incorrect. The third group, WEIRD, consisted of sentences that were grammatically correct, but semantically impossible. The GOOD and BAD filler sentences could be used to determine whether participants gave random answers or not. The WEIRD

sentences were mainly use to determine whether participants followed the instructions correctly and ignored sentence meaning. The filler sentences were also meant to mask the goal of the questionnaire, so that people would answer based on their intuitions and not on the grammar they learned in school.

2.3 Procedure

The sentences together formed a syntactic acceptability survey. This survey was made using Qualtrics. Participants could use their own computer or phone to complete the survey. First, instructions about the task were presented. A short explanation of the task was given, and participants were instructed that they should try to look at the meaning of the sentence as little as possible. After reading the instructions, 3 example sentences were shown. These example sentences consisted of filler sentences, one for each category (GOOD, WEIRD, and BAD). They aimed to further explain to the participant that they should not look at the meaning of the sentence. After reading the example sentences a message was shown to tell the user the actual experiment was about to start. The 128 sentences were presented to the participants 1 at a time, and in a random order. This order was different for each participant. When a sentence was presented, participants were asked how good they found the sentence. The possible answers were on a four-point Likert scale, which contained the options ‘Heel goed’, ‘Goed’, ‘Slecht’, and ‘Heel slecht’, or ‘Very good’, ‘Good’, ‘Bad’, and ‘Very bad’. Only one answer could be chosen for each question. Participants had an unlimited amount of time to answer each question, and they could quit and resume the experiment at any moment. An answer was required for every question, and participants could not go back to a previously answered question. At the end of the survey, an opportunity to give feedback or other remarks was presented.

2.4 Design

The variables that were mentioned thus far were kept in mind when designing the experiment. Later, during the data analysis stage, other variables were defined. Some of these variables described the sentences. These were the Scenario, which referred to the semantic content of the sentence, Q and P,

which were the amount of quantitative and prepositional *ers* in the sentence, respectively, the Tense, which was the tense of the main verb, the DistanceVE, which was the amount of words between the verb and *er*, and the WeirdEnd, which referred to whether the sentence had a ‘weird’ ending. This was the case in FRONT *OBJ+* sentences. Other datapoints that were collected were Participant, which was the participant number, and Sentence, which was the sentence number. Additionally, the amount of errors the participant made in the filler sentences were measured, as well as whether the participant made less than ten errors in the WEIRD filler sentences. These measurements corresponded to the variables Errors and Weird, respectively. The ratings for each sentence were counted to gain the frequencies of each rating for every sentence. This way, the sentence ratings could be compared to each other.

3 Results

In total, 43 participants completed the survey. Of these participants, 1 was removed from the data based on their answers on the filler sentences and the time they took to complete the survey. The results can be seen in Figures 3.1, 3.2, 3.3, and 3.4. The answers for each sentence type were added up, which means every sentence type has 164 ratings in total.

By looking at the distribution of colors in the bars, the ratings of the sentence types can be seen and compared. By comparing Figures 3.1 and 3.2 with Figures 3.3 and 3.4, it becomes clear that XQP sentences have higher acceptability than XQQ sentences in all cases. Other than that, it seems that the SUB- OBJ- sentences are always liked best, while the SUB+ OBJ+ sentences are liked least. There is variation regarding the preference between SUB- OBJ+ and SUB+ OBJ- sentences, but SUB- OBJ+ sentences are preferred the most often. Finally, and most importantly, the differences between FRONT and END sentences seem very minimal. The biggest difference between FRONT and END is in the XQQ sentences, where the acceptability of the SUB- OBJ+ sentences drops from 75 positive ratings for XQQ END to 55 positive ratings for XQQ FRONT.

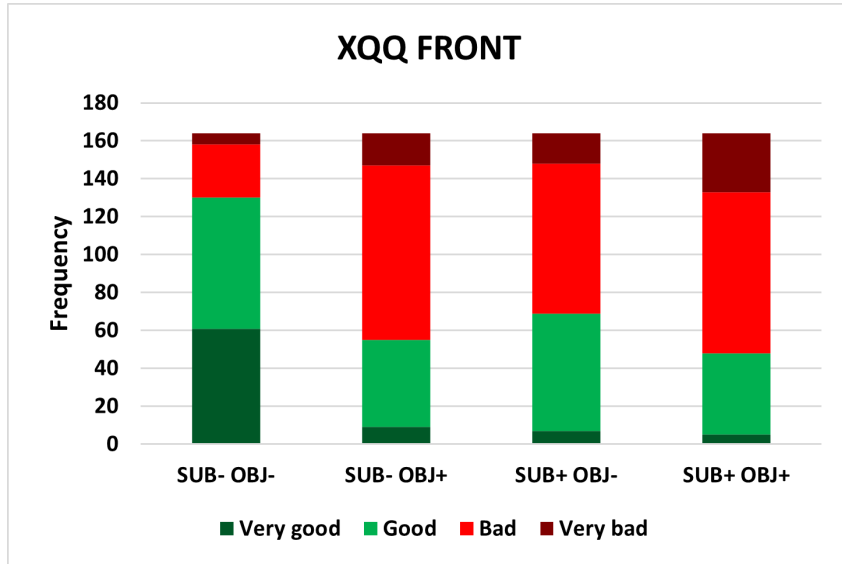


Figure 3.1: Results for XQQ FRONT sentences.

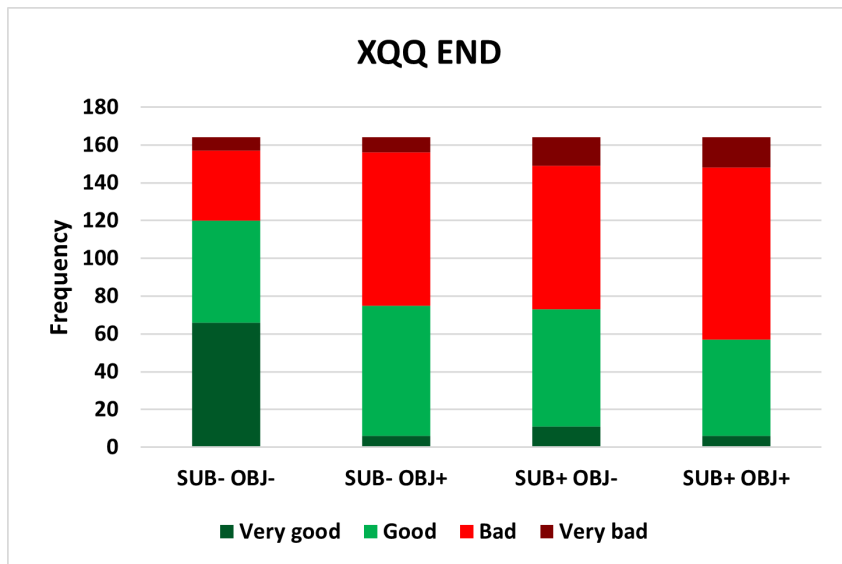


Figure 3.2: Results for XQQ END sentences.

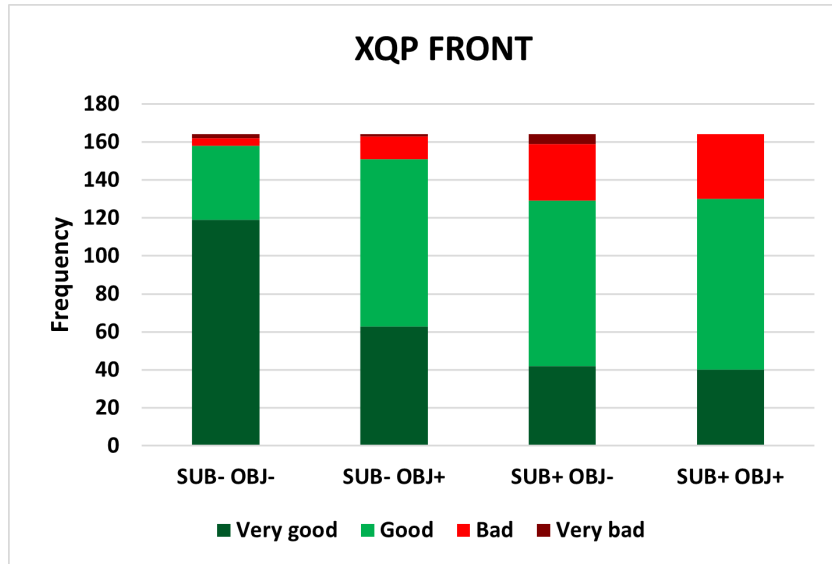


Figure 3.3: Results for XQP FRONT sentences.

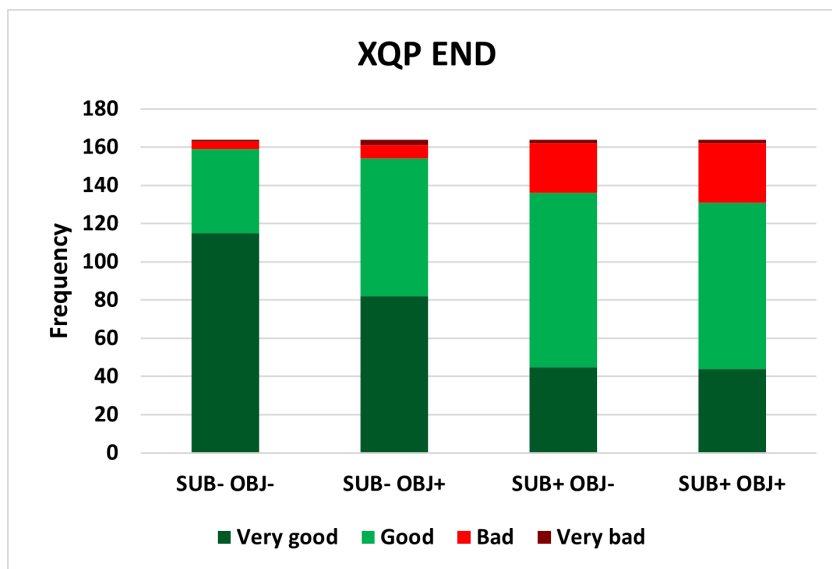


Figure 3.4: Results for XQP END sentences.

3.1 Data Analysis

In order to analyse the influence of the independent variables, binary logistic regression was used. Binary logistic regression was used instead of ordinal logistic regression because of the limited sample size. Especially in certain conditions where no sentences were given the rating ‘Very bad’, ordinal logistic regression proved unfit for the task. Using binary logistic regression meant that the response variables had to be put in two groups, which meant the ‘Very good’ and ‘Good’ ratings were combined, as well as the ‘Very bad’ and ‘Bad’ ratings.

The binary logistic regression was applied to all data at once, but also to the XQQ and XQP data separately. This was done to inspect a possible difference in the influence of verb location for the different pronominal functions of *er*. First, an appropriate model was found by using the stepAIC function from the MASS package in R, a stepwise regression method. The model was selected based on its AIC score, but also based on its predictive value. Variables such as Type and Tense were preferred over a variable such as Scenario. For all models, the dependent variable was the Rating. Then, binary logistic regression was applied on these models. The results of the binary logistic regression were converted to odds ratios, to make them easier to interpret.

3.1.1 All Data

The selected model had the following predictor variables: Participant, Type, Q, Tense, WeirdEnd, and Errors. The results of the binary logistic regression can be seen in Table 3.1. The 95% confidence intervals indicate that the effects of Tense and Errors are not significant. For the remaining variables, the odds ratios can be interpreted in the following way:

- For a 1 unit increase in participant number, the odds of getting a good rating decrease by a factor of 0.99.
- For an XQP sentence, the odds of having a good rating are 5.48 times that of XQQ sentences.
- For every additional quantitative er_Q in a sentence, the odds of getting a good rating decrease by a factor of 0.18.

- For a sentence without a weird ending, the odds of having a good rating are 1.33 times that of sentences with a weird ending.

3.1.2 XQQ

The selected model had the following predictor variables: Participant, Subject, Object, Tense, WeirdEnd, and Errors. The results of the binary logistic regression can be seen in Table 3.2. The 95% confidence intervals indicate that the effect of Participant is not significant. For the remaining variables, the odds ratios can be interpreted in the following way:

- For an XQQ sentence where *er* does not take the role of the subject, the odds of getting a good rating are 2.44 times that of an XQQ sentence where *er* does take the role of the subject.
- For an XQQ sentence where *er* does not take the role of the object, the odds of getting a good rating are 2.22 times that of an XQQ sentence where *er* does take the role of the object.
- For an XQQ sentence where the main verb is in the present tense, the odds of getting a good rating are 1.37 times that of an XQQ sentence where the main verb is in the past tense.
- For an XQQ sentence without a weird ending, the odds of getting a good rating are 1.62 times that of an XQQ sentence with a weird ending.
- For every additional error the participant makes in the filler sentences, the odds of giving a good rating to an XQQ sentence increases with a factor of 1.04.

3.1.3 XQP

The selected model had the following predictor variables: Participant, Subject, and Scenario. The results of the binary logistic regression can be seen in Table 3.3. The 95% confidence intervals indicate that the effects of ScenarioF and ScenarioH are not significant. For the remaining variables, the odds ratios can be interpreted in the following way:

Table 3.1: Binary logistic regression results for the entire data set.

Predictor	Log Odds	SE Log Odds	Odds Ratio	2.5%	97.5%
Intercept	2.93444	0.16095	18.8110084	13.7875868	25.9169316
Participant	-0.01367	0.00657	0.9864226	0.9737631	0.9991775
TypeXQQ	-1.70157	0.11074	0.1823974	0.1464785	0.2261468
Q	-0.99760	0.08323	0.3687644	0.3126687	0.4333660
TensePast	-0.18963	0.10093	0.8272644	0.6782577	1.0076034
WeirdEndYes	-0.28511	0.11540	0.7519323	0.5999488	0.9433482
Errors	0.01814	0.01203	1.0183021	0.9946031	1.0426491

Table 3.2: Binary logistic regression results for XQQ sentences only.

Predictor	Log Odds	SE Log Odds	Odds Ratio	2.5%	97.5%
Intercept	1.062163	0.165526	2.8926215	2.0970254	4.0141378
Participant	-0.013050	0.007917	0.9870344	0.9718042	1.0024533
SUB+	-0.891852	0.118149	0.4098960	0.3247487	0.5161275
OBJ+	-0.799253	0.142835	0.4496649	0.3393276	0.5941786
TensePast	-0.313447	0.123336	0.7309228	0.5735639	0.9303269
WeirdEndYes	-0.485230	0.171318	0.6155558	0.4393162	0.8602550
Errors	0.038174	0.014610	1.0389119	1.0096450	1.0691863

- For every 1 unit in Participant, the odds of getting a good rating decrease by a factor of 0.97.
- For a sentence where *er* does not take the role of the subject, the odds of getting a good rating are 4.63 times that of a sentence where *er* takes the role of the subject.
- For a sentence where the scenario is E, the odds of getting a good rating are 1.62 times that of a sentence where the Scenario is G.

4 Discussion

4.1 Goals of research

The goal of the research was to find whether the location of the verb has an influence on sentence acceptability in Dutch *er*-sentences. This was examined for sentences where *er* has different grammatical functions, and sentences where *er* has different pronominal functions. By examining these different grammatical and pronominal functions as well as the verb location, their influence on acceptability could also be seen. The ultimate purpose

of this was to gain a better understanding of how people use *er*.

4.2 Brief summary of results

On first glance, the differences in acceptability between FRONT and END sentences are very small. The biggest differences can be seen in XQQ OBJ+ sentences, where the acceptability in FRONT sentences is lower than in END sentences. Applying binary logistic regression on the XQQ sentences revealed that this is likely caused by FRONT OBJECT and FRONT BOTH sentences having a ‘weird ending’. This happens in sentences where *er* refers to the word that was otherwise at the end of the sentence. This can be seen in sentences (20), (21), (22) and (25) in section 2.2. People dislike sentences with a weird ending compared to sentences without a weird ending, but only in XQQ sentences. This is likely caused by the fact that, for some XQP sentences, the word that would be at the end of the sentence changes. This can be seen in sentence (21) in section 2.2, where ‘tegen’ changes to ‘tegenaan’.

For the different grammatical functions of *er*, sentences where *er* referred to neither the subject nor the object got the highest ratings, while sentences where *er* referred to both got the lowest rat-

Table 3.3: Binary logistic regression results for XQP sentences only.

Predictor	Log Odds	SE Log Odds	Odds Ratio	2.5%	97.5%
Intercept	3.590974	0.297627	36.2693718	20.6126170	66.2760730
Participant	-0.025857	0.007442	0.9744742	0.9602213	0.9886823
SUB+	-1.532809	0.203051	0.2159283	0.1430335	0.3177799
ScenarioF	0.064536	0.254089	1.0666643	0.6477633	1.7593123
ScenarioG	-0.484841	0.234548	0.6157950	0.3866320	0.9718126
ScenarioH	0.132103	0.257228	1.1412255	0.6893618	1.8957927

ings, which was expected. For the SUB- OBJ+ and SUB+ OBJ- conditions, the binary logistic regression revealed that *er* referring to the object is preferred over *er* referring to the subject. This is especially the case for XQP sentences, where OBJ± has no significant effect on acceptability. For XQQ sentences, people do dislike *er* referring to the object, but not as much as *er* referring to the subject. It is unclear why people prefer *er* referring to the object over *er* referring to the subject, but it is likely caused by the fact that the objects are optional in the experiment sentences.

The different pronominal functions of *er* greatly influence sentence acceptability. XQP sentences were rated higher than XQQ sentences in all conditions. This result is likely not caused by *er*, but by other parts of the sentence. This is based upon the fact that the difference in acceptability is also present in the sentences with only an existential er_X , the SUB- OBJ- sentences. The only change in these sentences, other than semantics, is the change between a quantity and a preposition at the end of the sentence. This could suggest several things, such as that two quantities next to each other are not liked, or that existential er_X is preferred when a preposition is in the sentence.

4.3 Conclusion with respect to Research Question

When looking back to the main research question(s), the following can be concluded:

- **Does the verb location influence acceptability?** No, in general, there is no difference in acceptability between sentences where the verb is in the front and sentences where the verb is at the end.
- **Is this effect different for different gram-**

matical functions of *er*? Yes, for sentences with *er* referring to the object, having the verb in the end of the sentence is preferred over having the verb in the front of the sentence.

- **Is this effect different for different pronominal functions of *er*?** No, at least for quantitative and prepositional *er*, there is no difference regarding the influence of verb location.

The additional research questions can be answered as follows:

- **Is there a preference regarding the grammatical functions of *er*?** Yes, *er* is preferred to refer to as little as possible in the sentence. In general, if *er* has to refer to either the object or the subject, it is preferred that *er* refers to the object.
- **Is this preference different for different pronominal functions of *er*?** Yes, for sentences with an er_{XQQ} and the verb in the front, *er* referring to the subject is preferred.

And finally, with regard to the pronominal functions of *er*, the question can be answered the following way:

- **Is there a preference regarding the pronominal functions of *er*?** Yes, sentences with an er_{XQP} are preferred over sentences with an er_{XQQ} .

4.4 Discussion of unexpected results

There were a few unexpected results. First off, some results of the binary logistic regression indicate that an increase in participant number decrease the odds

of a sentence getting a good rating. This is an artifact caused by ordering of the data with the participants that answered more than ten WEIRD filler questions wrong at the end. This seems to hint at a small interaction between sentence meaning and acceptability, where looking at the sentence meaning leads to lower acceptability in general, especially for XQP sentences.

A second unexpected result is the influence of the amount of errors the participant made on sentence rating. According to the results of the binary logistic regression, an increase in error number leads to an increase in rating. This could have to do with the fact that some people are taught that meaningless er_X is ungrammatical and should be left out of a sentence. If someone makes more errors in the filler sentences, it could be the case that they do not know grammar rules well, which would mean they do not think the meaningless er_X is bad. Another, more likely explanation is the disparity in sample size for error numbers. Most participants made none or only a few errors in the filler sentences, while only a few participants made many errors. This difference in amount of data has likely lead to the logistic regression finding an effect that does not actually exist.

A third unexpected result is the supposed influence of verb tense on sentence acceptability. The results of the logistic regression indicated that an XQQ sentence with the main verb in the present tense has a higher chance of getting a good rating than an XQQ sentence with the main verb in the past tense. This result is unreliable because of the low sample size, and the different sample size for verb tense in XQQ sentences. There are twenty XQQ sentences with the main verb in the present tense, but only twelve with the verb in the past tense.

The final unexpected result was that the verb location essentially has no influence on sentence ratings. It was speculated that sentences with the verb in the front would be preferred over sentences with the verb in the end, but the results do not support this. In certain cases, having the verb at the end of the sentence is even preferred over having er at the front of the sentence.

4.5 Conclusion with respect to related theories and bigger picture

According to linguists such as Odijk (1993) and Bennis (2011), the existential er_X that is used in the SUB- OBJ- sentences is grammatical, even though many schools teach, and newspapers say, that such a meaningless er should be left out of sentences. The ratings of XQP sentences show that the participants did not mind the use of er in these sentences, but the XQQ sentences show noticeably less positive ratings in this condition. According to the ANS (Haeseryn et al., 1997) existential er_X is often optional, and usage of existential er_X can differ per person. These results seem to indicate that usage of er_X is quite consistent when a preposition is present in the sentence, but not in sentences without a preposition.

The results also reveal that Odijk (1993) his intuitions regarding sentences with two quantites adjacent were accurate. He thinks that such sentences are not that good. Apparently, he is not the only one with this intuition. In the XQQ SUB+ OBJ+ sentences, where two quantities are adjacent, less than half of the participants gave a positive rating. Odijk (1993) attributes this to a focus clash. This focus clash could also be the reason for the generally lower ratings for XQQ sentences compared to XQP sentences.

In the end, the results of this study do give more insight into the usage of er . It was revealed that verb location only has an indirect influence on sentence acceptability. The additional analysis of grammatical and pronominal functions of er resulted in interesting insights as well, such as the big difference in results for the XQQ and XQP sentences. These insights bring us one step closer to understanding how er is used.

4.6 Potential problems (limitations)

There were some problems with this study. First off, the use of social media, as well as the message being spread to friends and family, likely resulted in a relatively young population. Furthermore, it is unknown in what part of the Netherlands the participants originated from. This is important because usage of er is often geographical (Haeseryn et al., 1997), (Grondelaers, 2020).

Second, the small sample size led to inability to

use ordinal logistic regression, which could have given more detailed results. In some conditions, such as SUB- OBJ+ for XQP sentences, the influence verb location has on acceptability is only visible when taking all response values in account. The change to only two response values removes such details from the results, which impacted the logistic regression.

Third, as was already shortly discussed in the Unexpected Results section, some questionable decisions were made when designing the stimuli and when preparing the data for analysis. These include the unnecessary variation in verb tense, a variation in verb tense for sentences of the same scenario, the lack of variation in quantities that were used (*twee/two* was used a lot), and an unnecessary change in order for the data. In addition to the unexpected results that this lead to, it might have influenced the other results of the logistic regression as well.

4.7 Suggestions for future research

In the future, it might be interesting to further examine the verb in *er*-sentences. It has already been pointed out by Grondelaers et al. (2009) that sentence context influences the processing of sentences using *er* with unexpected antecedents. This suggests that the main verb, which is the context in the sentences used in this study, does have some influence on processing. Introducing unexpected antecedents to the sentences might reveal this influence. Changing the main verb from more to less specific might also reveal this. An analysis of verb location for locative *er_L* might also be interesting. The results of this study already suggest a difference in influence of verb location for quantitative *er_Q* and prepositional *er_P*, but locative *er_L* was not examined.

References

- Barf, I. (2021). *Acceptability of sentences in which the Dutch word 'er' holds multiple functions*. Bachelor's thesis, Rijksuniversiteit Groningen. Retrieved from <https://fse.studenttheses.ub.rug.nl/24116>.
- Bennis, H. (2011). *Gaps and Dummies*. New York: De Gruyter Mouton.
- Crystal, D. (2011). *A dictionary of linguistics and phonetics*, volume 30. John Wiley & Sons.
- De Schutter, G. (1992). Partitief of kwantitatief 'er', of over de verklaring van syntactische variatie. *Taal & Tongval*, 44:15–26.
- Donaldson, B. C. (2008). *Dutch: a comprehensive grammar*. Abingdon: Routledge, 2nd edition.
- Grondelaers, S. (2020). Verwacht het onverwachte: Over hoe nederlanders er anders gebruiken dan vlamingen. *Nederlandse Taalkunde*, 25(2-3):205–211.
- Grondelaers, S., Speelman, D., Drieghe, D., Brysbaert, M., and 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.
- Haeseryn, W. J. M., Romijn, K., Geerts, G., De Rooij, J., and Van den Toorn, M. C. (1997). *Algemene Nederlandse spraakkunst*. Groningen, Wolters-Noordhoff.
- Jones, S. M. (2020). Multifunctional dutch 'er'. In Butt, M. and Toivonen, I., editors, *Proceedings of the LFG'20 Conference, On-Line*, pages 148–168, Stanford, CA. CSLI Publications.
- Odiijk, J. E. J. M. (1993). *Compositionality and syntactic generalizations*. PhD thesis, Katholieke Universiteit Brabant.
- Voortman, W. (2005). The use of er.

A Appendix

The sentences that were used in the questionnaire. The experiment sentences all follow the same structure. For every set of 4, which can be recognized by being identical except for the words *er* refers to, the order is the following: SUB- OBJ-, SUB- OBJ+, SUB+ OBJ-, SUB+ OBJ+.

EXPERIMENT SENTENCES

XQQ FRONT

1. Soms schoppen er twee jongens één muur.
2. Soms schoppen er twee jongens één.
3. Soms schoppen er twee één muur.
4. Soms schoppen er twee één.
5. Af en toe maakt er één kok twee desserts.
6. Af en toe maakt er één kok twee.
7. Af en toe maakt er één twee desserts.
8. Af en toe maakt er één twee.
9. Gisteren ontvoerden er vijf boeven drie kinderen.
10. Gisteren ontvoerden er vijf boeven drie.
11. Gisteren ontvoerden er vijf drie kinderen.
12. Gisteren ontvoerden er vijf drie.
13. Over een jaar krijgen er veertig studenten twee diploma's.
14. Over een jaar krijgen er veertig studenten twee.
15. Over een jaar krijgen er veertig twee diploma's.
16. Over een jaar krijgen er veertig twee.

XQQ END

17. Ik zag dat er twee jongens één muur schopten.
18. Ik zag dat er twee jongens één schopten.
19. Ik zag dat er twee één muur schopten.
20. Ik zag dat er twee één schopten.

21. Hij vermoedt dat er één kok twee desserts maakt.
22. Hij vermoedt dat er één kok twee maakt.
23. Hij vermoedt dat er één twee desserts maakt.
24. Hij vermoedt dat er één twee maakt.
25. Ze dachten dat er vijf boeven drie kinderen ontvoerden.
26. Ze dachten dat er vijf boeven drie ontvoerden.
27. Ze dachten dat er vijf drie kinderen ontvoerden.
28. Ze dachten dat er vijf drie ontvoerden.
29. Wij hopen dat er veertig studenten twee diploma's krijgen.
30. Wij hopen dat er veertig studenten twee krijgen.
31. Wij hopen dat er veertig twee diploma's krijgen.
32. Wij hopen dat er veertig twee krijgen.

XQP FRONT

33. Morgen gaan er twee agenten naar het huis.
34. Morgen gaan er twee agenten naartoe.
35. Morgen gaan er twee naar het huis.
36. Morgen gaan er twee naartoe.
37. Plotseling sprongen er vijf meisjes in het meer.
38. Plotseling sprongen er vijf meisjes in.
39. Plotseling sprongen er vijf in het meer.
40. Plotseling sprongen er vijf in.
41. Afgelopen week keken er twaalf studenten naar de maan.
42. Afgelopen week keken er twaalf studenten naar.
43. Afgelopen week keken er twaalf naar de maan.
44. Afgelopen week keken er twaalf naar.

45. Vorig jaar vlogen er twee vogels tegen het raam.
46. Vorig jaar vlogen er twee vogels tegenaan.
47. Vorig jaar vlogen er twee tegen het raam.
48. Vorig jaar vlogen er twee tegenaan.

XQP END

49. Zij hoopt dat er twee agenten naar het huis gaan.
50. Zij hoopt dat er twee agenten naartoe gaan.
51. Zij hoopt dat er twee naar het huis gaan.
52. Zij hoopt dat er twee naartoe gaan.
53. Wij zagen dat er vijf meisjes in het meer sprongen.
54. Wij zagen dat er vijf meisjes in sprongen.
55. Wij zagen dat er vijf in het meer sprongen.
56. Wij zagen dat er vijf in sprongen.
57. Jullie weten dat er twaalf studenten naar de maan kijken.
58. Jullie weten dat er twaalf studenten naar kijken.
59. Jullie weten dat er twaalf naar de maan kijken.
60. Jullie weten dat er twaalf naar kijken.
61. Ik denk dat er twee vogels tegen het raam vlogen.
62. Ik denk dat er twee vogels tegenaan vlogen.
63. Ik denk dat er twee tegen het raam vlogen.
64. Ik denk dat er twee tegenaan vlogen.

FILLER SENTENCES

GOOD

65. Die eend is groen.
66. De man loopt naar huis.
67. Ik zit op een stoel.
68. Hij had nog nooit zoiets gekst gezien.

69. Het feest is voorbij.
70. Wij stellen het erg op prijs.
71. De man laat zijn hond uit.
72. Na drie jaar ben ik dan toch eindelijk geslaagd voor het examen.
73. De helft van de vragen werd fout beantwoord.
74. Één persoon slaapt in één tweepersoonsbed.
75. Wil je een koekje voor mij pakken?
76. Hoe luid snurkt jouw buurman?
77. Jullie werden vorige keer behoorlijk vaak aangesproken.
78. Elke ochtend aten zij drieëndertig beschuitjes.
79. Waarom liep die man langs die huizen?
80. Vorig jaar hadden ze mij voor het eerst gezien.
81. Zij zijn hier nog niet eerder geweest.
82. Jij komt hier niet zo vaak, toch?
83. Zij is allergisch voor pinda's.
84. Wie zit er in dat vliegtuig?
85. Lees je zulke spannende boeken vaker?

WEIRD

86. Hoe heet jouw kaas?
87. Wanneer valt die piloot nou van zijn computer?
88. Normaal gesproken hinkel ik met mijn handen.
89. Zij zet haar koffie altijd op haar raam neer.
90. Die eend is een fiets.
91. Iedere dinsdag snuit ik mijn middelste neusgat.
92. Hij is geboren op 31 februari.
93. Mijn lievelingskleur is frikandel.
94. Willen jullie morgen op mijn auto zwemmen?
95. Zij zei dat ze deze aarde het mooiste vond.

96. De grootmeester had eindelijk gewonnen van zijn printer.
97. Ik heb al drie stukken taart gelezen.
98. Jullie typen altijd de mooiste thee.
99. De televisie vindt wortels het lekkerst.
100. Hun water smaakt paars.
101. Het toetsenbord houdt erg van wandelen.
102. Zij hebben een hek gebouwd zodat hun walnoot niet weg loopt.
103. Ik neem elke dag een woning mee naar school.
104. Jij stapelt mijn auto soms in tweeëntwintig stukjes.
105. Waarom zit jouw been in mijn hoofd?
106. Die groep viert ieder weekend oud en nieuw.
107. Zijn tandpasta heeft last van puistjes.
122. Doet u willen ik te komen en u op te pakken?
123. Wat tijd zijn jullie gaan naar het busstation?
124. Hoe kan ik gaan naar het centrum?
125. Zijn de boeken gestopt in de tas?
126. Iemand die weet het antwoord zou moeten opsteken hun vingers.
127. Na het examen, studenten zullen hebben twaalf dagen van vakantie.
128. De voorstelling is over te beginnen.

BAD

108. Wees voorzichtig rijdend.
109. Kan jij meenemen dit voor mij?
110. Ik ben geweest hier voor twee dagen.
111. Wanneer ik was acht jaar oud, ik rende weg van huis.
112. Ik was over te verlaten het restaurant wanneer mijn vrienden arriveerden.
113. Wilt u nemen een berichtje alstublieft?
114. Hij nooit geeft mij iets.
115. Zij is gaan naar Amerika volgend jaar.
116. Jij nog steeds moet poetsen je tanden en nemen een douche.
117. Zij is gaat mee morgen.
118. Wat doe jij denk van deze schoenen?
119. Hij denkt wij doen niet willen gaan.
120. Ik heb nodig dit om te zijn daar door morgen.
121. Alsjeblieft vertel hem dat hij is weg.