Master thesis Logopediewetenschap

Information-theoretical approach to the omission of functional categories in Broca's aphasia

Geertje Zilverschoon, 0203971

Supervisor: Sergey Avrutin, PhD

Universiteit Utrecht, June 2009

Abstract

This study is focusing on the omissions of two functional categories in the speech of Broca's aphasics; articles and prepositions. For each selected functional element, the complexity can be calculated by using the Information Theory. The aim of this study is to find a connection between the complexity of the elements and their omission rate. Are more complex elements omitted the most? For the articles a very precarious tendency is found for 3 of the 12 subjects: More complex articles were omitted the most. For the prepositions only one subject showed that more complex prepositions were omitted more.

Keywords: omission functional categories, Broca's aphasia, complexity, Information Theory.

1.0 Introduction

Speech and language therapists in a rehabilitation centre often treat clients with aphasia. Aphasia is a language disorder which results from brain damage. Several disorders can cause brain damage, but the main cause is a cerebral vascular accident. Aphasia may affect all language modalities: writing, reading, speaking and comprehension (Goodglass & Kaplan, 1972). There are different types of aphasia. Goodglass & Kaplan (1972) distinguished 7 types. The type of aphasia is focused on in this study is Broca's aphasia. Broca's aphasia is characterized by slow and effortful speech, difficulties in articulation and intonation, and sentences are short and displaying a simplified syntactic structure. This simplification contains omissions and/or substitutions of functional categories such as complementizers, determiners, pronouns, and tense (Avrutin, 2001).

This study is looking to some aspects of the speech of Dutch Broca's aphasics. Because of the time limit this study is only focusing on the omission pattern of two functional categories: articles and prepositions. For each selected functional element, the complexity can be calculated by using the Information Theory. The aim of this study is to find a connection between the complexity of the elements and their omission rate by putting the pieces together.

This chapter will discuss the selected functional categories, the omissions of the selected functional categories in Broca's aphasics and the Information Theory.

1.1 Functional categories

Functional words are articles, pronouns, auxiliary verbs, complementizers and some prepositions, and a few other elements. They determine aspects of syntactic structure, such as the introduction of relative clauses and questions. Function words can also introduce different types of abstract semantic information, for example possession. These words are frequently found in languages and they have a fixed number of elements; they form a closed class (Kean, 1977). As mentioned above this study is only focusing on two types of functional categories. In the first subsection the articles will be discussed. The second subsection will discuss the prepositions.

1.1.1 The articles

In Dutch there are only three articles; 'de', 'het' and 'een'. The article 'de' is a definite article. It can function as singular common (i.e. both masculine and feminine), plural common, plural

neuter, diminutive plural common and diminutive plural neuter. The article ‘het’ is also a definite article which can function as singular neuter, diminutive neuter and diminutive common. The article ‘een’ is an indefinite article. It can function as singular common, singular neuter, diminutive common and diminutive neuter. (Algemene Nederlandse Spraakkunst, 2002).

1.1.2 The prepositions

In Dutch there are about eighty prepositions. Actually there is no consensus in the literature about whether the prepositions side with the lexical categories nouns, verbs and adjectives or whether they are a functional category like determiners, complementizers or conjunction elements (Zwarts, 1997). Zwarts (1995) distinguished two types of prepositions in Dutch, real prepositions (type A), as in (1a) and a periphery of words that behave like prepositions in certain respects (type B), as in (1b).

(1) a: aan (at), achter (behind), bij (near), binnen (inside) , boven (above), buiten (outside), door (trough), in (in), langs (along), met (with), na (after), naar (to), naast (beside), om (around), onder (under), op (on), over (over), tegen (against), tot (as far as, until), tussen (between), uit (from), van (of), voor (for, in front of).
b: aangaande (concerning), inzake (on the subject of), inclusief (including), overeenkomstig (in correspondence with), richting (in the direction of), sedert (since), zonder (without).

Zwarts (1995) argued that the type B prepositions should not be treated as genuine instances of the category prepositions. When we focus only on the real prepositions (type A) they turn out to be functional in some respects (they form a closed class of small words) and lexical in other respects (they can be used intransitively and in morphological processes and they can be stranded) (Zwarts, 1997).

Froud (2000) also mentioned this precarious position of the prepositions. In her study she examined data from an aphasic patient who demonstrated an extremely robust dissociation between lexical and functional categories. Froud (2000) described this patient as extremely poor at reading functional categories but relatively unimpaired at reading substantives. Froud provided this patient as a testing ground to help decide whether prepositions are underlyingly represented as lexical or functional. The results of Froud’s study showed that this aphasic patient treated the prepositions exactly as if they were functional.

On the basis of the findings of Froud (2000) I like to include prepositions in this present study. Because Zwarts (1995) argued that type B prepositions should not be treated as genuine instances of category prepositions I will only use type A prepositions.

1.2 Omission of functional categories

As mentioned in the introduction, Broca's aphasics often omit functional categories. It is argued that the damage to the Broca's region does not result in a reduction of morphosyntactical knowledge, nor that some parts of the linguistic structure are missing (Avrutin, 2002). But if not, why do Broca's aphasics omit functional categories? How different is their speech from normal speech? To give a possible answer to these questions I would like to outline the syntax-context model of Avrutin (1999, 2002, 2004).

1.2.1 Avrutin's syntax-context model (1999, 2002, 2004)

In some cases, omission of functional categories is allowed. For example (1) (Avrutin, 2004):

1) Leuk huisje heb je!
Nice house have you

This expression with the determinerless NP 'huisje' is fully acceptable and productive in Dutch, but it does require specific contextual conditions. Determinerless NP's in Dutch are acceptable only in specific contexts where there is a presupposition with regard to the referent of the NP. This means that the context sometimes can take over the functional category. But if such conditions are not satisfied, functional elements must be provided in order to make an utterance interpretable (Avrutin, 2002).

During the language acquisition children often omit functional categories even if such conditions are not satisfied. For example (2) (Avrutin, 2004):

2) Daar komt trein (Niek, 3;00;09)
there comes train.

Similar observations have been made for Dutch Broca's aphasics. For example (3) (Baauw et al, 2002) 'Rekening' has no determiner.

3) Rekening is voldaan (patient HB)
Bill is paid

Why do children and Broca's aphasics omit functional categories?
Avrutin (1999) introduced a model about the interaction between the narrow syntax, information structure and context, based on the speech of unimpaired adults. He assumes that narrow syntax is a computational system that is isolated and encapsulated with respect to meaning. It puts lexical items together in some specific order that is allowed in a given language. Eventually the output of this system must be interpretable. However, the meaning of lexical items by themselves is not always sufficient. The interpretation of an element could depend on the information in the information structure.
In the information structure the information about topic, focus, specificity and pronominal anaphora is encoded. Avrutin (2004) describes the information structure as a system that is constructed in the course of a given conversation and that operates by rules that go beyond a sentence level. He describes context as a non-linguistic system that can be modified by different means.

Avrutin (2004) illustrates the relationship between the narrow syntax, information structure and context as follows:

Figure 1
Normal way of introducing an individual information unit (Avrutin, 2004).

narrow syntax	information structure	context
D: → all features NP: information → Normal way →	Frame Heading	?

The units of information in the information structure must be separated from each other and they must be identifiable. The information structure depends on the input from the narrow syntax and is constructed on its basis. The output of the narrow syntax must provide for building independent, interpretable units of information. This building exists of two different types of elements, functional and lexical elements (Avrutin, 2004).

The information structure can be seen as a computational system that operates on information units. One information unit consists as seen in the figure above of a frame and a heading. The task of a frame is to ensure that information units are separated from each other. The task of a heading is to provide the information that's necessary for interpretation. In figure 2 an example for 'a dog' is presented.

Figure 2

Normal way of introducing an individual information unit for DP 'a dog' (Avrutin, 2004).

narrow syntax	information structure	context
D: → all features NP: information → dog Normal way →	Frame Heading: dog	?

The frame unit in this example is supplied by determiner 'a'. The heading is supplied by the noun 'dog'.

The frame of the information structure can also be introduced by the context. Let us go back to example (1): 'Leuk huisje heb je!'
The response 'huisje' has a determinerless expression, thus the narrow syntax does not supply the material necessary for introducing an individual frame. Specific contextual circumstances can however introduce a frame as well through strong presupposition, as in fig 2.

Figure 3

Non-syntactic way of introducing an individual frame (Avrutin, 2004).

narrow syntax	information structure	context
D: → No features in narrow syntax NP: information → huisje	Frame Heading: huisje	← (Special registers individual frame introduced through presupposition)

Avrutin (2004) hypothesizes that in normal adult speakers narrow syntax is the cheapest, most economical way of building information structures.

During the development of children, when the neural substrates supporting syntactic computations are not fully developed, it is possible that the narrow syntax operations are more resource consuming than for unimpaired adult speakers. In this case they may sometimes rely on alternative means of introducing information units, like context. Avrutin (2004) speaks of a competition between the narrow syntax and the context for building an information unit. In child speech the competition sometimes results in a victory for the narrow syntax route, and sometimes for the context route. So a child sometimes produces a functional category and sometimes the functional category is omitted.

In the course of brain maturation, the narrow syntax becomes more powerful and turns to be the most economic way of building the information structure. This will result in the production of more functional categories and less omissions (Avrutin, 2004).

Avrutin (2004) points out, that omissions of functional categories are also present in the speech of patients with Broca's aphasia. Actually the omission of functional categories is considered to be one of the most prominent features of the speech of Broca's aphasics.

Avrutin (2002) describes that as a result of damage to Broca's region, the processing resources are limited. Therefore the power necessary for conducting operations involving narrow syntax is decreased. This reduced power will sometimes result in the situation where alternative systems become more powerful, and therefore are used to build the information structure.

So if we return to the questions of section 1.1.2. Why do Broca's aphasics omit functional categories? How different is their speech from normal speech? Avrutin (2002) explains the omission of functional categories in Broca's aphasics through a limitation of processing resources. The limited resources result in what can be described as a reduction of processing capacity of the narrow syntax route, which leads, in turn, to the fact that some elements are omitted. Besides the noticeable omissions of the functional categories, the speech of Broca's aphasics is not very different from normal speech because there seem to be no reduction of morphosyntactical knowledge, nor are some parts of the linguistic structure missing.

1.3 Information Theory

In the previous section two routes to build the information structure were discussed; the route of the narrow syntax and the route of the context. Avrutin (2004) hypothesizes that in normal adult speakers narrow syntax is the cheapest, most economical way of building information structures. But as a result of damage to Broca's region there will be more competition between the narrow syntax and the context for building an information unit. This competition will sometimes results in a victory for the narrow syntax route, and sometimes for the context route. So the Broca's aphasic will sometimes produce a functional category and sometimes omit the functional category. It's assumable that more complex functional elements need more processing resources, and therefore they might be often omitted, comparing to less complex functional elements. To measure the complexity of functional elements the Information Theory can be used.

The Information Theory finds its roots in Shannon and Weaver (1949) with the purpose to study and solve the problem of the transmission of signals over communication channels. In the early forties of the twentieth century it was thought that the increase of the transmission rate of the information over a communication channel would increase the probability of error. Shannon proved that not the transmission rate of the information but the complexity level of the encoded information established the probabilities of errors. He named the complexity level of the encoded information the *entropy* (H) and proposed a model to calculate this complexity level.

De Lange (2008) outlines the idea that we can think of the human language processing brain in terms of channels transmitting information in order to communicate. If we do, we can hypothesize that there can also be errors in the processing of human language (e.g. omissions of articles and prepositions) when the entropy of the encoded information exceeds the channel capacity.

In her dissertation De Lange (2008) used a formula introduced by Kostić (2004) based on an application of Shannon's and Weaver's (1949) Information Theory, to measure the complexity of processing articles. Every functional element has its own complexity which can be calculated by using this formula; the information load, presented as formula 1. This formula will be explained with an example of the article 'de'.

Formula 1

Information load.

$$I_m = \left[-\log_2 \left(\frac{\frac{F_m}{R_m}}{\sum \frac{F_{m_j}}{R_{m_j}}} \right) \right]$$

F_m : indicates the frequency of the specific functional element expressed as a percentage of the total number of all articles. De Lange (2008) calculated that the frequency of the article 'de' as a percentage of the total number of articles in Dutch is 47.90 %. Thus, $F_m = 47.90$.

R_m: the number of functions and meanings the functional element can appear in. For the article 'de' there can be distinguished the following number of functions and meanings:

1) singular / common / definite
2) plural / common /definite
3) plural / neuter / definite
4) diminutive / plural / common / definite
5) diminutive / plural / neuter / definite

Thus the article 'de' has 5 functions and meanings, $R_m = 5$.

F_m / R_m: indicates what the frequency is with which the specific functional element appears in a certain meaning. The average frequency with which the article 'de' appears in one of its 5 functions/meanings is 47.90 / 5 = 9.58. However this value by itself does not give us all necessary information with respect to the processing cost of this specific form. It's necessary to know how this value depends on the F/R values of the other elements, in this case the articles.

Σ F_{mj} / R_{mj}: The F/R rates for the other articles must be calculated, and added up. In this example the F/R rates for the articles 'het' and 'een' are respectively 6.07 and 8.47. Thus the sum of the three F/R rates is 9.58 + 6.07 + 8.47 = 24.12.

Next the F/R rate of the article ‘de’ must be divided by the sum, to know how this value depends on the F/R values of all the articles:

$((F_m / R_m) / \Sigma F_{mj} / R_{mj})$: 9.58 / 24.12 = 0.397

Of this value the logarithm must be calculated to find the Information Value of the individual element. Thus in the case of the article ‘de’ we take the logarithm of 0,397 = 1,332.

The information load of the article ‘de’ is: I = 1,332.

The bigger the information load, the more complex an element is.

2.0 Research question

Given that Broca’s aphasics sometimes omit functional categories (such as articles and prepositions), and given that Broca’s aphasics have limited processing resources, how can we connect these limited resources with the omissions of functional categories, and are some specific categories omitted more than others?

To answer this question it’s necessary to put up a hypothesis. Avrutin (2002) describes in his theory that in some cases, when there is not enough power for the narrow syntax to build the information structure, some other systems become more powerful and build the information structure. But what decides whether there is or isn’t enough power to build the information structure by using the route of the narrow syntax? It is assumable to think that the complexity of the functional category decides if there is enough power or not, and thus if it is produced or omitted. Every functional category has its own complexity. This complexity can be calculated using the Information Theory, which finds its roots in Shannon and Weaver (1949).

Based on the model of Avrutin (2002) it is reasonable to think that a more complex element needs more power to take the narrow syntax route, and is therefore more often omitted in the speech of Broca’s aphasics because of the limited processing resources. This assuming will lead to the hypothesis:

H_0: Omissions of functional categories (such as articles and prepositions) are dependent on the complexity, as measured by the information load.

Prediction: More complex elements are omitted the most.

To test the H_0 hypotheses three steps must be taken. First the pattern of omission of various functional categories in speech of Dutch Broca’s aphasics must be studied. In this study the

following functional categories will be studied: the articles ‘de’, ‘een’ and ‘het’ and type A prepositions.
Second, the complexity of the selected functional elements will be calculated by using the information load formula, introduced by Kostić (2004). With the information of the different calculations a complexity hierarchy of the functional elements can be built up.
Third, on the base of the previous two steps the hypothesis can be tested: Are more complex elements omitted the most?

3.0 Part 1: Pattern of omission of articles and type A prepositions in speech of Dutch Broca’s aphasics

3.1 Research questions

To discover the pattern of omission of articles and prepositions, two questions must be answered.

Question 1)
Is there a significant difference in omitting the articles ‘de’, ‘een’ and ‘het’ in the speech of Dutch Broca’s aphasics?

In her dissertation De Lange (2008) investigates the difference in omitting the articles ‘de’ and ‘het’ in Dutch child speech. In her study she found a significant difference in omitting between these two articles. The article ‘het’ is significantly more often omitted by Dutch children than the article ‘de’. In this article comparison De Lange (2008) leaves the article ‘een’ out of consideration. It turns out to be problematic to make a proper comparison of the use and omission of definite and indefinite articles in child speech because it is not always possible to see if a definite or indefinite article has been omitted in spontaneous speech (De Lange, 2008).
When we look at the age of first use of the three articles there is also a difference in the age of first use of ‘de’, ‘een’ and ‘het’. All children start to use ‘de’ earlier than ‘het’. Schaerlaekens and Gillis (1987) describe that there is no real difference in age of first use of ‘de’ and ‘een’. Given these facts and given the fact that both Broca’s aphasics and children have limited processing resources it’s reasonable to put up a hypothesis that Broca’s aphasics also omit the article ‘het’ more than the articles ‘de’ and ‘een’.

H_0: Dutch Broca's aphasics omit 'het' significantly more often than 'de' and 'een'.
Prediction: The speech samples show significantly more omissions of the article 'het' compared to the articles 'de' and 'een'.

Question 2)
Is there a significant difference in omitting type A prepositions in the speech of Dutch Broca's aphasics?

Several studies have been done to the prepositional use in Broca's aphasia. Friederici (1982) distinguished two types of prepositions; semantic and syntactic prepositions. In the German language the same form of a preposition can occur both in a semantic and syntactic way only by varying its functional role. For example, a preposition can be used as a lexical preposition, which bears at least some semantic information, as in (4).

(4) Peter steht *auf* dem Stuhl (Peter stands *on* the chair).

The same form of the preposition can also be used as an obligatory preposition that is lexically dependent on the preceding verb, as in (5).

(5) Peter hofft *auf* den Sommer (Peter hopes *for* the summer).

In an experiment she investigated whether there was a difference in production of Broca's aphasics for the two types of prepositions. The analysis of the experiment revealed that the Broca's aphasics produce a preposition more easily when it bears some semantic information. The production is dramatically reduced when it is primarily syntactically based.
Bennis et al. (1983) put also up a similar experiment, but they distinguished three types of prepositions; syntactic prepositions (6), subcategorized prepositions (7) and lexical prepositions (8).

(6) The book is *for* Mary.
(7) Mary is waiting *for* John.
(8) John put the book *on* the table.

Bennis et al. (1983) found that syntactic prepositions were significantly difficult to produce compared to the subcategorized and the lexical prepositions.

Both experiments found significant differences in omission rate between the different functional types of prepositions. In my study I'm not interested in the differences in production of different functional types of prepositions. I'm rather more interested in differences in production of the prepositions on the basis of their individual form. Frequency could be an important factor in the production and omission of the prepositions. Table 1 gives an overview of all Dutch type A prepositions and their frequencies which were distinguished by using the Corpus of Spoken Dutch (CGN) (2004). Experimental studies utilizing speech error analysis in people with aphasia have often reported word frequency effects on lexical performances, indicating that higher frequency words were named more accurately than lower frequency words (Bose et al. 2006). It seems assumable for Broca's aphasics that lower frequency words will be omitted significantly more often than higher frequency words. So it can be predicted that for instance 'van', a higher frequency preposition, will be easier to produce than 'langs', a lower frequency preposition, and therefore omitted less often.

Table 1

Dutch prepositions and their frequencies.

prepositions	**frequency**	**prepositions**	**frequency**
naast	1551	uit	18268
langs	1921	over	22597
boven	2127	bij	31556
buiten	2353	naar	32691
achter	3248	om	34389
binnen	3961	aan	38680
tussen	4375	voor	50992
onder	4625	op	59166
na	5818	met	61134
tegen	8319	in	118629
tot	10321	van	132758
door	14455		

H_0: Dutch Broca's aphasics omit lower frequency prepositions significant more than higher frequency prepositions.

Prediction: The speech samples show significantly more omissions of lower frequency prepositions comparing to higher frequency prepositions.

3.2 Materials and Methods

For this part of the research I looked at the speech of twelve right-handed Dutch Broca's aphasics, with Akense Afasie Test (AAT) scores of 100% aphasia and 85-100% Broca's aphasia. Seven data files were obtained from the department language research of rehabilitation centre De Hoogstraat in Utrecht. Six data files were given by Dr. E. Ruigendijk from the University of Oldenburg. Two data files were obtained from a published case study by Kolk et al. (1990).

Table 2 gives a summary of the data used in the study, including information about the gender, the age, the occupation, the cause of the aphasia, the handedness, the time post onset (TPO), the percentage of aphasia, the percentage of Broca's aphasia, utterances with article-requiring nouns, and utterances with prepositions- requiring word groups, in which the articles and prepositions were either used or omitted.

Eliminated from the utterances in the files were unclear sentences, immediate repetitions of the same sentence, clear imitations of the researcher, idiomatic expressions, rhymes and songs, and routine sentences (based on de Lange, 2008).
The relevant utterances were analyzed by hand by two speech and language therapists.

Table 2

Broca's aphasic speech data used in the study.

	Gender	**Age**	**Occupation**	**Cause**	**handedness**	**TPO**	**% aphasia**	**% Broca-aphasia**	**utterances with article-requiring nouns**	**utterances with prepositions-requiring word groups**
AN	M	73	chief administration	?	right	20y	?	?	28	27
IH (1)	F	57	receptionist	CVA-?	right	5m	100	95	12	11
IH (2)	F		receptionist	CVA-?	right	1y7m	?	?	9	10
JW	M	41	senior consultant	CVA-l	right	2y6m	100	98,7	19	30
AK	M	82	school director	CVA-l	right	2y	100	99,9	19	15
MK	F	63	saleswoman	CVA-l	right	1y6m	100	100	11	12
Barn	M	40	motor mechanic	CVA-l	right	?	?	?	62	48
Heck	M	44	construction draftsman	aneurysm	right	?	?	?	86	70
SH	F	34	manager	CVA-l	right	4,5m	100	85,3	18	21
RB (1)	M	49	service station attendant	CVA-l	right	1,5m	100	100	12	8
RB (2)	M	50	service station attendant	CVA-l	right	5m	100	98.5	11	22
ME	M	30	road worker	CVA-l ACM	right	2m	100	91,7	8	8
KK (1)	F	24	teacher primary school	CVA-l	right	2m	100	100	3	11
KK (2)	F	25	teacher primary school	CVA-l	right	6,5m	100	99	3	8

The most desirable statistical method for this research is the chi-square. Two assumptions must be made. For the test to be meaningful it is imperative that each person, item or entity contributes to only one cell of the contingency table. The expected frequencies should be greater than 5. If expected frequencies are below 5 the result is a loss of statistical power (Field, 2005).

Before the experiment was conducted, the number of utterances required to obtain the desired power of .8 for both articles and prepositions was estimated.
The desirable power is .8, the degrees of freedom contain 2, and an alpha level of .05 is used. To find a significant difference between the omission of different elements (for example 'de', 'het' en 'een') Chi^2 must be 6 ($df=2$, $p < 0.05$). A fictive pilot study to find a significant differences (at alpha <.05) was calculated for different values of Chi^2 (Lenth, 2006-9).

Table 3

Number of utterances needed to obtain the desired power of 0.8 (at alpha=.05) for different values of the Chi^2.

power	**α**	**Chi^2**	**Number of utterances**
0.8	0.05	2	96
0.8	0.05	3	64
0.8	0.05	4	48
0.8	0.05	5	39
0.8	0.05	6	32

This pilot study shows how many utterances of each functional category are needed to find a Chi^2 with $p < 0.05$.

Unfortunately, as seen in table 2, of all subjects only Barn and Heck meet the established number of utterances; for the articles respectively 62 and 86, and for the prepositions respectively 48 and 70. All other subjects do not meet the minimum of 32 utterances per word class for a Chi^2 of 6. Because the low number of utterances it is better to use as statistical analysis the Likelihood ratio. This statistic will be roughly the same as Pearson's chi-square, but is preferred when samples are small (Field, 2005).

3.3 Results and Analysis

In this section the data of the study on article and preposition omission in Broca's aphasic's speech are presented. In the first subsection the data on the articles will be discussed. In the second subsection the data on the prepositions will be explained.

3.3.1 Article omission

I analyzed whether there were differences in article omission for the whole group of Broca's aphasics. First the following calculations were made for each individual subject.

- The total number of article 'de'-requiring nouns used without 'de' divided by the total number of 'de'-requiring nouns.
- The total number of article 'het'-requiring nouns used without 'het' divided by the total number of 'het'-requiring nouns.
- The total number of article 'een'-requiring nouns used without 'een' divided by the total number of 'een'-requiring nouns.

Second the mean omission rates for all subjects together were calculated. Table 4 shows the mean omission rates and the total number of required articles of the three Dutch articles for all the Broca's aphasics together. The article 'de' is more often used than 'een' and 'het'. But no significant difference in omission was found between the three types of articles (F= 1.70, p=.193).

Table 4

Total number required and overall omission rate: the mean.

articles	**total number required**	**overall omission rate: M**
de	156	27.02
het	49	50.53
een	113	46.79

Table 5 reports the omission and use of 'de', 'het' and 'een' in obligatory contexts in verbal utterances for each individual subject. It also represents the omission rate in percentage, the Likelihood ratio and the corresponding significance level. The significant results are shaded in grey.

Tabel 5

Omission and use of 'de', 'het' and 'een' in obligatory contexts in verbal utterances.

DP is the number of article requiring contexts in which an article was used.
NP is the number of article requiring contexts in which an article was omitted.

	AK			**AN**			**IH (1)***			**IH (2)****		
	de	**het**	**een**	**de**	**het**	**een**	**de**	**het**	**een**	**de**	**het**	**een**
DP	2	2	9	5	2	8	8	0	3	4	1	4
NP	1	2	3	6	2	5	0	0	1	0	0	0
omission rate (%)	33.3	50	25	54.5	50	38.5	0	0	25	0	0	0
	$L\chi^2$= .839			$L\chi^2$=.647			$L\chi^2$=2.385					
	p= .658			p=.724			p=.122					

***** The article 'het' was not concluded in the statistical analysis.
** No statistics are computed because the NP value was a constant.

	JW			**MK***			**Barn**			**Heck**		
	de	**het**	**een**	**de**	**het**	**een**	**de**	**het**	**een**	**de**	**het**	**een**
DP	8	0	3	3	0	0	4	0	1	42	8	9
NP	2	2	4	6	2	0	22	13	22	9	11	7
omission rate (%)	20	100	57.1	66.7	100	0	84.6	100	95.7	11.8	57.9	43.8
	$L\chi^2$=6.295			$L\chi^2$=1.434			$L\chi^2$=4.211			$L\chi^2$=11.697		
	p=.043			p=.231			p=.122			p=.003		

***** The article 'een' was not concluded in the statistical analysis..

	HZ			**KK (1)***			**KK (2)**			**ME****		
	de	**het**	**een**	**de**	**het**	**een**	**de**	**het**	**een**	**de**	**het**	**een**
DP	9	0	2	2	0	0	1	1	0	1	0	1
NP	0	1	2	1	0	0	0	0	1	0	0	7
omission rate (%)	0	100	50	33.3	0	0	0	0	100	0	0	87.5
	$L\chi^2$=9.003						$L\chi^2$=3.819			$L\chi^2$=3.506		
	p=.011						p=.148			p=.061		

***** No statistics are computed because the article value was a constant.
** The article 'het' was not concluded in the statistical analysis.

	RB (1)			**RB (2)**			**SH***		
	de	**het**	**een**	**de**	**het**	**een**	**de**	**het**	**een**
DP	3	0	0	3	0	3	8	0	7
NP	3	1	5	2	1	4	1	0	2
omission rate (%)	50	100	100	40	100	57.1	11.1	0	22.2
	$L\chi$2=5.178			$L\chi^2$=1.654			$L\chi^2$=.407		
	p=.075			p=.437			p=.524		

***** The article 'het' was not concluded in the statistical analysis.

Looking at the level of Broca's aphasics individually in table 5, it can be seen that there is a relation between articles and omission statistically significant for 3 of the subjects: JW, Heck and HZ. All three of them omit 'de' significantly less than the other two articles. For 3 subjects, IH (1), ME and SH, the article 'het', and for one subject MK, the article 'een' was

not present in their samples and therefore not concluded in the statistical analysis (represented as striped cells in table 5). For two subjects, IH (2) and KK (1), no statistics were computed because one of the variable values was a constant (represented as black cells in table 5).
So for three subjects there is found a statistically significant value. Unfortunately, as mentioned in section 3.2 only subject Heck is meeting the established number of utterances, and therefore the power is strong enough to support the significant difference between the omissions of the different articles.

3.3.2 Preposition omission

I analyzed whether there were differences in preposition omission for the whole group of Broca's aphasics. First the following calculations were made for each individual subject.

- The total number of 'aan'-requiring word groups used without a preposition divided by the total number of 'aan'-requiring word groups.
- The total number of 'achter'-requiring word groups used without a preposition divided by the total number of 'achter'-requiring word groups.
- The total number of 'bij'-requiring word groups used without a preposition divided by the total number of 'bij'-requiring word groups.
- The total number of 'binnen'-requiring word groups used without a preposition divided by the total number of 'binnen'-requiring word groups.
- The total number of 'boven'-requiring word groups used without a preposition divided by the total number of 'boven'-requiring word groups.
- The total number of 'buiten'-requiring word groups used without a preposition divided by the total number of 'buiten'-requiring word groups.
- The total number of 'door'-requiring word groups used without a preposition divided by the total number of 'door'-requiring word groups.
- The total number of 'in'-requiring word groups used without a preposition divided by the total number of 'in'-requiring word groups.
- The total number of 'langs'-requiring word groups used without a preposition divided by the total number of 'langs'-requiring word groups.
- The total number of 'met'-requiring word groups used without a preposition divided by the total number of 'met'-requiring word groups.
- The total number of 'na'-requiring word groups used without a preposition divided by the total number of 'na'-requiring word groups.
- The total number of 'naar'-requiring word groups used without a preposition divided by the total number of 'naar'-requiring word groups.

- The total number of ‘naast’-requiring word groups used without a preposition divided by the total number of ‘naast’-requiring word groups.
- The total number of ‘om’-requiring word groups used without a preposition divided by the total number of ‘om’-requiring word groups.
- The total number of ‘onder’-requiring word groups used without a preposition divided by the total number of ‘onder’-requiring word groups.
- The total number of ‘op’-requiring word groups used without a preposition divided by the total number of ‘op’-requiring word groups.
- The total number of ‘over’-requiring word groups used without a preposition divided by the total number of ‘over’-requiring word groups.
- The total number of ‘tegen’-requiring word groups used without a preposition divided by the total number of ‘tegen’-requiring word groups.
- The total number of ‘tot’-requiring word groups used without a preposition divided by the total number of ‘tot’-requiring word groups.
- The total number of ‘tussen’-requiring word groups used without a preposition divided by the total number of ‘tussen’-requiring word groups.
- The total number of ‘uit’-requiring word groups used without a preposition divided by the total number of ‘uit’-requiring word groups.
- The total number of ‘van’-requiring word groups used without a preposition divided by the total number of ‘van’-requiring word groups.
- The total number of ‘voor’-requiring word groups used without a preposition divided by the total number of ‘voor’-requiring word groups.

Second the mean omission rates for all subjects together were calculated. Table 6 shows the mean omission rates and the total number of required prepositions of the twenty-three Dutch type A prepositions for all the Broca’s aphasics together. No significant difference in omission was found between the different types of prepositions (F= 1.102, p=.362).

Table 6

Total number required and overall omission rate: the mean.

prepositions	**total number required**	**overall omission rate: M**
aan	16	45.24
achter	1	0
bij	22	66.67
binnen	7	0
boven	2	0
buiten	5	0
door	11	41.67
in	71	31.28
langs	0	0
met	23	50.60
na	0	0

naar	30	37.47
naast	1	0
om	12	57.14
onder	0	0
op	38	40.18
over	5	60
tegen	0	0
tot	7	25
tussen	3	50
uit	9	20
van	28	25.14
voor	7	42.86

Table 7 reports the omission and use of the different prepositions in obligatory contexts in utterances for each individual subject. It also represents the omission rate in percentage, the Likelihood ratio and the corresponding significance level. The significant results are shaded in grey.

Tabel 7

Omission and use of the type A prepositions in obligatory contexts in verbal utterances.

PP is the number of preposition requiring contexts in which a preposition was used.
DP is the number of preposition requiring contexts in which a preposition was omitted.

	AK							**AN**									**IH (1)**					
	in	van	aan	met	uit	voor	om	in	op	tot	naar	bij	met	over	uit	buiten	naar	van	op	in	bij	uit
PP	0	6	1	2	2	1	1	12	1	0	2	0	0	0	1	1	2	1	0	4	0	0
DP	1	1	0	1	0	0	0	1	3	1	2	1	1	1	0	0	0	0	1	1	1	2
omission rate (%)	100	14.3	0	33.3	0	0	0	7.7	75	100	50	100	100	100	0	0	0	0	100	20	100	100
	$L\chi^2$=5.882							$L\chi^2$=18.499									$L\chi^2$=11.297					
	p=.437							p=.018									p=.046					

	IH (2)						**JW**										**MK**					
	op	met	naar	in	om	aan	naar	om	van	in	met	tot	bij	op	over	buiten	op	door	naar	in	voor	bij
PP	2	1	1	3	0	1	7	1	1	4	2	3	0	0	0	1	2	0	1	2	1	0
DP	0	0	0	0	2	0	2	0	2	4	0	0	1	1	1	0	0	1	0	3	0	0
omission rate (%)	0	0	0	0	100	0	22.2	0	66.7	50	0	0	100	100	100	0	0	100	0	60	0	100
	$L\chi^2$=10.008						$L\chi^2$=14.985										$L\chi^2$=6.730					
	p=.075						p=.091										p=.151					

	Barn																	**ME**			
	van	voor	om	naar	op	aan	tussen	met	in	door	boven	bij	uit	naast	over	binnen	buiten	in	van	over	aan
PP	2	0	1	0	2	2	0	0	1	2	2	1	3	1	1	2	1	1	2	0	0
DP	1	1	2	4	3	4	1	1	2	1	0	0	0	0	0	0	0	0	3	1	1
omission rate (%)	33.3	100	66.7	100	60	66.7	100	100	33.3	50	0	0	0	0	0	0	0	0	60	100	100
	$L\chi^2$=27.169																	$L\chi^2$=3.855			
	p=.040																	p=.278			

	Heck														**SH**						
	aan	in	bij	voor	met	van	op	naar	door	om	uit	binnen	buiten	over	met	bij	in	tot	door	naar	op
PP	0	17	0	0	1	2	7	3	2	2	1	4	2	1	2	4	2	1	1	1	1
DP	5	4	1	1	1	2	2	2	0	1	0	0	0	0	0	2	4	0	0	0	0
omission rate (%)	100	19	100	100	50	50	22.2	40	0	33.3	0	0	0	0	0	33.3	66.7	0	0	0	0
	$L\chi^2$=26.822														$L\chi^2$=9.158						
	p=.013														p=.241						

	HZ						**KK (1)**						**KK (2)**						**RB (1)**					
	op	van	naar	aan	met	bij	op	door	met	in	voor	bij	van	op	voor	in	bij	naar	met	in	voor	bij	om	met
PP	6	2	1	1	0	2	2	2	0	2	1	0	2	0	1	1	2	0	2	1	0	0	0	2
DP	2	0	0	0	1	0	0	0	2	0	0	2	0	1	0	0	0	1	1	1	1	1	1	1
omission rate (%)	25	0	0	0	100	0	0	0	100	0	0	100	0	100	0	0	0	100	33.3	50	100	100	100	33.3
	$L\chi^2$=6.015						$L\chi^2$=14.421						$L\chi^2$=8.997						$L\chi^2$=3.993					
	p=.305						p=.013						p=.109						p=.407					

	RB (2)										
	met	van	op	door	bij	om	aan	tot	naar	achter	binnen
PP	3	2	2	0	1	0	1	2	0	1	1
DP	2	0	0	1	2	2	1	0	1	0	0
omission rate (%)	40	0	0	100	66.7	100	50	0	100	0	0
	$L\chi^2$=16.445										
	p=.088										

Looking at the level of Broca's aphasics individually in table 7, it can be seen that there is a statistically significant relation between prepositions and the omission rate for five subjects: AN, IH (1), Barn, Heck and KK (1). All five of them omit some prepositions significantly more than other prepositions. When we focus on the prediction that lower frequency prepositions are omitted more than higher frequency prepositions, only subject IH (1) turns out to meet this prediction. For the other four subjects there is no similar pattern noticeable. Unfortunately, as mentioned in section 3.2 only subjects Heck and Barn are meeting the established number of utterances. So for subject IH (1) the power is not strong enough to support the significant difference between the omissions of the different prepositions.

3.4 Conclusion

Although the article 'de' is more often used for all the subjects together than the articles 'het' and 'een', as seen in table 4, there was no significant difference found between the three types (F= 1.70, p=.193). In the calculations for each individual subject significant differences were found between the three articles for 3 of the subjects; JW, Heck and HZ, as seen in table 5. All three of them omit 'de' significantly less than the other two articles. Only subject Heck is meeting the established number of utterances, and therefore the power is strong enough to support the significant difference between the omissions of the different articles

For the prepositions there was also no significant difference found in omission between the different types for all the subjects together (F= 1.102, p=.362). In the calculations for each individual subject significant differences were found between the prepositions for 5 of the subjects; AN, IH (1), Barn, Heck and KK (1), as seen in table 7. All five of them omit some prepositions significantly more than other prepositions. Only subject IH (1) turns out to meet the prediction that lower frequency prepositions are omitted more often than higher frequency prepositions Unfortunately IH (1) is not meeting the established number of utterances and therefore the power is not strong enough to support the significant difference between the omissions of the different prepositions.

4.0 Part 2: Calculating the complexity of the selected functional elements by using the Information Theory

4.1 Research goal

The aim of this part of research is to put up a hierarchy of complexity for the articles and a hierarchy of complexity for the selected prepositions. To put up hierarchies the information loads for each individual element must be calculated (by using formula 1). For the articles the number of functions and meanings described by De Lange (2008) will be used. To distinguish all possible functions and meanings for the prepositions a dictionary will be used. The number of frequencies for both articles and prepositions will be collected using the CGN (2004).

Formula 1

Information Load.

$$I_m = \left[-\log_2 \left(\frac{\frac{F_m}{R_m}}{\sum \frac{F_{m_j}}{R_{m_j}}} \right) \right]$$

4.2 The implementation of the formula by De Lange (2008)

De Lange (2008) calculated the information load of the three Dutch articles. First she distinguished the following number of meanings and functions:

‘de’ 1) singular / common / definite
2) plural / common /definite
3) plural / neuter / definite
4) diminutive / plural / common / definite
5) diminutive / plural / neuter / definite

‘het’ 1) singular / neuter / definite
2) diminutive / neuter / definite
3) diminutive / common / definite

‘een’ 1) singular / common / indefinite

2) singular / neuter / indefinite

3) diminutive / common / indefinite

4) diminutive / neuter / indefinite

Based on these functions and meanings De Lange (2008) calculated the information load of the three Dutch articles, table 8.

Table 8

Calculation information load individual articles, based on the numbers of functions and meanings.

	Frequency in % (F) (corpus ‘Gesproken Nederlands’)	**number of functions / meanings (R)**	**F / R**	**(F/R) / (sum F/R) (p)**	**I = - log $_2$ p**
‘de’	47,90 %	5	9,58	0,397096	1,332439
‘het’	18,22 %	3	6,07	0,251775	1,989796
‘een’	33,88 %	4	8,47	0,351129	1,509927

The article ‘het’ has a higher information load than the articles ‘de’ and ‘een’, so it’s a more complex element. Based on this data a hierarchy of complexity can be put up: ‘de’-‘een’-‘het’.

4.3 Results and Analysis

In this section the outcomes of the calculations of the information loads for both article and preposition are presented. In the first subsection the data on the articles will be described and in the second subsection the data on the prepositions.

4.3.1 Information load of the articles

To calculate the information loads for the articles the frequencies of the articles were distinguished by using CGN. The number of functions and meanings were taken over from De Lange (2008).

Table 9

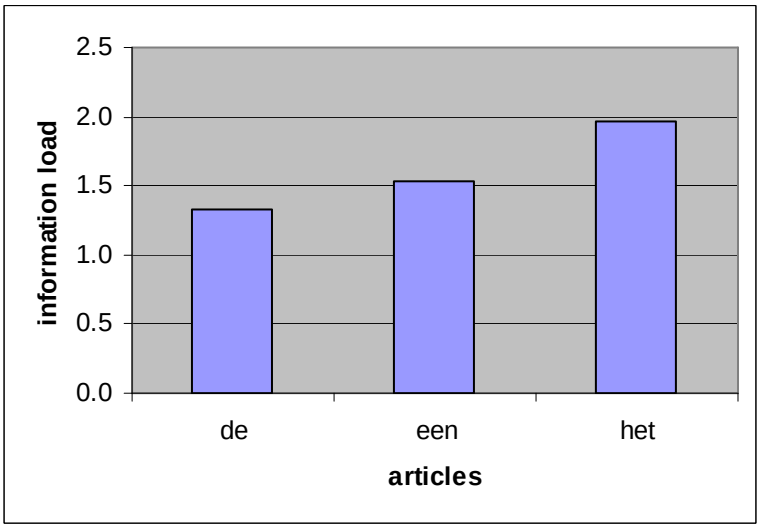
Calculation informative load individual articles.

Articles	Frequencies	R	F in %	F/R	(F/R)/(ΣF/R)	Info load
de	249303	5	48.13	9.63	0.399	1.326415
het	96264	3	18.59	6.2	0.257	1.961680
een	172363	4	33.28	8.32	0.345	1.537368
total	**517930**		**Σ F/R →**	**24.15**		

Based on this data a hierarchy of complexity can be put up (figure 4).

Figure 4

Information load of the articles.



4.3.2 Information load of the prepositions

To calculate the information loads for the prepositions the frequencies of the articles were distinguished by using CGN. The numbers of functions and meanings were distinguished using the dictionary (van Dale, 1992).

Table 10

Calculation informative load individual prepositions.

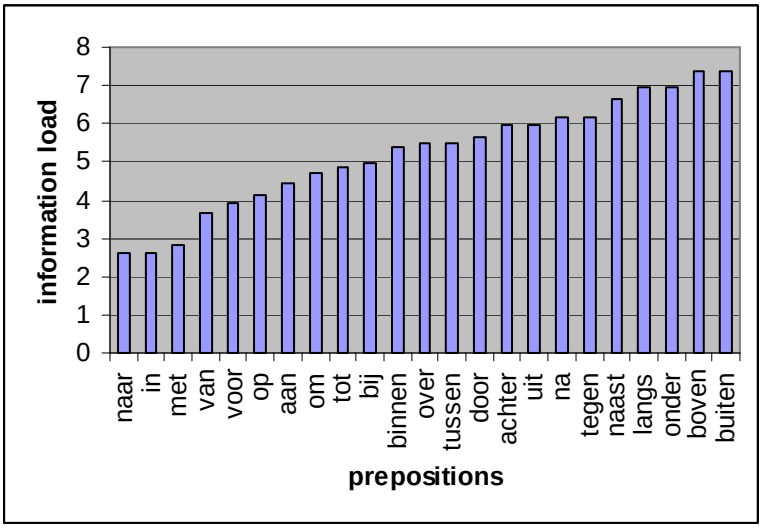
Prepositions	Frequencies	R	F in %	F/R	(F/R)/(ΣF/R)	info load
aan	38680	25	5.83	0.23	0.046	4.442222
achter	3248	6	0.49	0.08	0.016	5.965784
bij	31556	29	4.75	0.16	0.032	4.965784
binnen	3961	5	0.6	0.12	0.024	5.380822
boven	2127	12	0.32	0.03	0.006	7.380822
buiten	2353	11	0.35	0.03	0.006	7.380822
door	14455	22	2.18	0.1	0.02	5.643856

in	118629	22	17.87	0.81	0.161	2.634867
langs	1921	7	0.29	0.04	0.008	6.965784
met	61134	13	9.21	0.71	0.141	2.826233
na	5818	13	0.88	0.07	0.014	6.158429
naar	32691	6	4.92	0.82	0.163	2.617056
naast	1551	5	0.23	0.05	0.01	6.643856
om	34389	27	5.18	0.19	0.038	4.717857
onder	4625	18	0.7	0.04	0.008	6.965784
op	59166	31	8.91	0.29	0.058	4.107803
over	22597	31	3.4	0.11	0.022	5.506353
tegen	8319	18	1.25	0.07	0.014	6.158429
tot	10321	9	1.55	0.17	0.034	4.878321
tussen	4375	6	0.66	0.11	0.022	5.506353
uit	18268	33	2.75	0.08	0.016	5.965784
van	132758	50	20	0.4	0.079	3.662004
voor	50992	23	7.68	0.33	0.065	3.943416
total	**663934**		**Σ F/R →**	**5.04**		

Based on this data a hierarchy of complexity can be put up (figure 5).

Figure 5

Information load of the prepositions.



4.4 Conclusion

By using formula 1, the complexity of every single element was calculated. For both articles and prepositions a complexity hierarchy is put up, respectively figure 4 and figure 5. In part 3 of this research, these hierarchies will be correlated to the omission rates of part 1.

5.0 Part 3: Putting the parts together

In this chapter there will be an attempt to answer the research question *"Are more complex elements omitted the most?"*. First of all it must be noticed that there can only be calculations for the subjects that have shown significant differences in omission for the different types of articles and/or prepositions. Because there were found no significant differences in omission for both articles and prepositions for all the subjects together there cannot be a calculation to look if more complex elements are omitted the most. So in this part the research question can only be answered for a couple of subjects.
In the first section of this chapter the research question for the articles will be discussed and this will be done for the prepositions in the second section.

5.1 Are more complex articles omitted the most?

Individual calculations were only made for the subjects for which a significant difference in omission between the three types of articles has been found; JW, Heck and HZ (established in part 1). The correlation was calculated between this omission rate and de information load (established in part 2).

Table 11

Pearson's correlations between omission rate and information load for articles.

Subject	correlation	significance
JW	.989	.095
Heck	.916	.262
HZ	.982	.122

Although the correlation values of all three are very close to 1 none is significant. This is caused by the fact that there are only three data points; one for 'de', one for 'het' and one for 'een'.

5.2 Are more complex prepositions omitted the most?

Individual calculations were only made for the subjects for whom was found a significant difference in omission between the different types of prepositions; AN, IH (1), Barn, Heck and KK (1).

Table 12

Pearson's correlations between omission rate and information load for prepositions.

Subject	correlation	significance
AN	-.237	.539
IH (1)	.820	.046
Barn	-.644	.005
Heck	-.404	.152
KK (1)	-.082	.877

As mentioned in de section 3.4, only IH (1) is showing in part 1 that lower frequency prepositions were omitted more than higher frequency prepositions. That's why IH (1) now is the only subject who shows a positive correlation between the omission rate and the information load. The others didn't show that lower frequency prepositions were omitted the most and therefore we find negative correlations. So in this part only the outcome of IH (1) is interesting. Subject IH (1) remains with a positive correlation and a significance of $p < .05$.

5.3 Conclusion

In this part of the research, part 1 and part 2 were combined for answering the overall research question. Only the significant outcomes of part 1 could be matched with the outcomes of part 2. For the articles therefore only three subjects were used to correlate their omission rates to the calculated information loads. None of the subjects showed a significant correlation, although their outcomes were very close to 1. So for these three subjects it can be said very precarious that there is a tendency that more complex articles (with a higher information load) are omitted more. For the prepositions, five subjects showed a significant difference in omitting different kinds of prepositions. Only one of them showed that lower frequency prepositions were omitted the most. For this one subject the correlation between its omission rate and information load was significant and therefore it can be said that (in this single case) more complex prepositions (with a higher information load) are omitted more. Because this subject was not meeting the established number of utterances (in part 1), the power was not strong enough to support the significant difference between the omissions of the different prepositions. So the significant correlation found in this part, is unfortunately based on an omission rate with a weak power and therefore it does not count as strong evidence in answering the question whether more complex prepositions are omitted the most.

6.0 Discussion

Although the minimum number of utterances was established to obtain the desired power of .8 for both articles and prepositions previous to the execution of this research, none of the subjects, except two, fulfilled this condition. The cause of this serious failing lies in the fact that this study was dependent on already existing samples. Unfortunately most of these samples were very short, thereby resulted in not making the established minimum number of utterances. This failure has a very negative impact on the whole study.

In part one, an analysis has been made on the basis of percentages of omissions for all subjects together. Because these percentages are calculated on so few utterances per subject, the analysis is distorted. In the individual analyses the calculations seem to turn out better, because the calculations are based on the rough data. Nevertheless there are still not enough utterances to make a good comparison with the desired power of .8 between the different articles and prepositions.

In part one, it is determined which subjects are suitable for the calculations in part three. Because of the lack of utterances in part one, only the data of a few subjects are suitable to correlate in part three with the information load, which is calculated in part two.

The research question whether omissions of functional categories (such as articles and prepositions) are depending on the complexity as measured by the information load cannot be answered for all subjects, based on these insufficient data.

In further research I recommend the usage of self-recorded samples. Existing samples, as part of the AAT, available in rehabilitation centers and hospitals are often used only for diagnosing and therefore very short. By using your own recordings you can stop recording as soon as you have enough utterances. Only when the desirable number of utterances is reached, the power will be high enough to give a reliable answer to the research question.

Besides spontaneous speech I would also recommend to record the speech of the Amsterdam Nijmegen Test voor Alledaagse Taalvaardigheden (ANTAT) and picture description tests. Spontaneous speech only elicits what the patient is able to do, and will not always directly show what he or she cannot do. In this study for instance, the subjects might avoid prepositional constructions. In case of the ANTAT, subjects still have to formulate their own sentences, but in restricted contexts. In a picture description task, the constructions of interest can be provoked.

As to the calculation of the information load of the articles, information of the dissertation from De Lange (2008) is used to establish the number of functions and meanings. To establish the number of functions and meanings for the prepositions the dictionary (Van Dale, 1992) is used. Although this seems not the most objective manner, there is no present information in the literature about the number of functions and meanings of every single preposition in Dutch. It could be possible that the established numbers in this study do not correspond with the actual numbers. Kostić (2004) already mentioned concerning for the definition of functions and meanings in Serbian, that the definition of the number functions and meanings is an arbitrary issue as it depends on one's theoretical viewpoint. Kostić (2004) underlined that it is however not important to know the absolute number of functions and meanings, but the proportion of a specific element, relative to other elements. In further research, as for the prepositions, it is necessary to do better linguistic analysis to establish the number of functions and meanings for the prepositions.

Acknowledgements

First of all I would like to thank Sergey Avrutin, for supervising me in a very pleasant and motivating manner. I also like to thank the department language research of rehabilitation centre De Hoogstraat in Utrecht and Dr. E. Ruigendijk from the University of Oldenburg for providing me speech samples. I'm grateful to Lizet van Ewijk, speech and language therapist, for analyzing the speech samples with me. I extend special thanks to all reviewers for their very helpful comments on the text.

References

Avrutin, S. (1999). *Development of the syntax-discourse interface*, Dordrecht: Kluwer Academic Publishers. In Lange, J. de (2008). *Article omission in headlines and child language: Processing approach.* Dissertation, Utrecht University.

Avrutin, S. (2001). Linguistics and agrammatism**,** *Glot Internationa*l *5.3*, 1–5.

Avrutin, S. (2002). 'Weak Syntax,' In Grodzinsky, Y. & Amunts, K. (eds.) *Broca's region.* New York: Oxford Press, 49-62.

Avrutin, S. (2004). Optionally in child and aphasic speech, *Lingue e Linguaggi, 3, (1),* 65-95.

Baauw, S., Roo, E. de & Avrutin, S. (2002). *Determiner omission in language acquisition and language impairment: syntactic and discourse factors,* "Proceedings of the Boston University conference on language development" Boston: Cascadilla Press, 23-35.

Bennis, H., Prins, R., & Vermeulen, J. (1983). Lexical-semantic versus syntactic disorders in aphasia: The processing of prepositions. *Publikaties van het instituut voor algemene taalwetenschap, 40,* Amsterdam University.

Bose, A., Lieshout, P. van. & Square, P.A. (2006). Word frequency and bigram frequency effects on linguistic processing and speech motor performance in individuals with aphasia and normal speakers, *Journal of Neurolinguistics, 20,* 65-88.

Field, A. (2005). *Discovering Statistics Using SPSS.* London: Sage.

Froud, K. (2000). Prepositions and the lexical/functional divide: Aphasic evidence. *Lingua, 111,* 1-28.

Dale, van (1992). Groot woordenboek der Nederlandse taal. 12de druk. Utrecht: Van Dale Lexicografie.

Friederici, A.D. (1982). Syntactic and semantic processes in aphasic deficits: The availability of prepositions. *Brain and language, 15,* 249-258.

Goodglass, H. & Kaplan, E. (1972). *The assessment of aphasia and related disorders.* Philadelphia: Lea & Febiger. In Roo, E. de (1999). *Agrammatic grammar,* Doctoral dissertation, LOT 24, Leiden University.

Kean, M.-L. (1977). The linguistic interpretation of aphasic syndromes: agrammatism in Broca's aphasia, an example. *Cognition 5,* 9-46. In Roo, E. de (1999). *Agrammatic Grammar,* Doctoral dissertation, LOT 24, Leiden University.

Kolk, H., Heling, G. & Keyser, A. (1990). Agrammatism in Dutch: Two case studies. In Menn, L., Obler, L.K. & Miceli, G. (1990). *Agrammatic aphasia: a cross-language narrative sourcebook.* Amsterdam: Benjamins.

Kostić,A. (2004). The effects of the amount of information on processing of inflected morphology, ms.University of Belgrade. In Lange, J. de (2008). *Article omission in headlines and child language: Processing approach.* Dissertation, Utrecht University.

Lange, J. de (2008). *Article omission in headlines and child language: Processing approach.* Dissertation, Utrecht University.

Schaerlaekens, A.M. & Gillis, S. (1987). *De taalverwerving van het kind,* Groningen: Wolters-Noordhoff.

Shannon, C. & Weaver, W. (1949). *The mathematical theory of communication,* London: University of Illinois Press. In Lange, J. de (2008). *Article omission in headlines and child language: Processing approach.* Dissertation, Utrecht University.

Zwarts, J. (1995). Complex prepositions and P-stranding in Dutch. *Linguistics 35* (1997), 1091-1112.

Zwarts, J. (1997). Lexical an functional properties of prepositions. In: Haumann, D. & Schierholz S.J. (1997) *Lexikalische und grammatische Eigenschaften präpositionaler Elemente.* Tübingen: Niemeyer.

Websites:

- Algemene Nederlandse Spraakkunst (2002). Faculteit der Letteren van de KU Nijmegen & Nederlandse Taalunie. http://www.let.ru.nl/ans/e-ans/index.html
- Corpus Gesproken Nederlands, (2004). Nederlandse Taalunie. http://lands.let.kun.nl/cgn/
- Lenth, R. V. (2006-9). Java Applets for Power and Sample Size [Computer software]. http://www.stat.uiowa.edu/~rlenth/Power