

**Article Omission in Headlines
and Child Language:
A Processing Approach**

Published by
LOT
Janskerkhof 13
3512 BL Utrecht
The Netherlands

phone: +31 30 253 6006
fax: +31 30 253 6406
e-mail: lot@let.uu.nl
<http://www.lotschool.nl>

Cover illustration: Peter J. van Dongen.

ISBN 978-90-78328-63-6
NUR 616

Copyright © 2008: Joke de Lange. All rights reserved.

Article Omission in Headlines and Child Language: A Processing Approach

Het Weglaten van Lidwoorden in Krantenkoppen
en Kindertaal: Een Benadering vanuit Taalverwerking
(met een samenvatting in het Nederlands)

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. J.C. Stoof,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op vrijdag 5 september 2008
des ochtends te 10.30 uur

door

Johanna de Lange

geboren op 13 augustus 1961 te Gouda

Promotor: Prof. dr. P.H.A. Coopmans
Co-promotor: Dr. S. Avrutin

Contents

Chapter 1 Introduction	1
1.1 General Introduction.....	1
1.2 Article Paradigm and Article Use in Dutch and Italian.....	4
Chapter 2 Psycholinguistic Background	11
2.1 Introduction	11
2.2 Levelt's (1989) model of language production	12
2.3 Avrutin's (1999, 2004a, 2004b) syntax-context model	17
2.4 Levelt's model and Avrutin's model combined.....	25
2.5 Language production of open class versus closed class words	27
2.5.1 Introduction	27
2.5.2 Determiner selection models	28
2.6 Caramazza's determiner selection model.....	33
2.6.1 Introduction	33
2.6.2 Primed Unitized Activation Model, a description and basic assumptions.....	34
2.6.3 Crosslinguistic differences: early selection languages versus late selection languages.....	37
2.6.4 An apparent contradiction between data from acquisition studies and data from processing studies	39
2.7 Conclusion: Model for article selection	41
2.7.1 Introduction	41
2.7.2 The components.....	43
2.7.3 The process of article selection: a general overview.....	44
2.7.4 The process in detail: The model's prediction for different speakers/conditions.....	49
2.7.4.1 Normal speakers, normal discourse situation	49
2.7.4.2 Speakers with limited processing resources	55
2.7.4.3 Normal adults in specific contexts (telegram style, colloquial speech, headlines)	58
2.7.5 Why more omissions in Dutch than in Italian?	59
2.7.6 Concluding remark: separate storage of articles.....	62
2.7.7 Summary	63
Chapter 3 Article omission in Headlines and Child Speech	65
3.1 Introduction	65

3.2 Article omission in headlines.....	66
3.2.1 Introduction	66
3.2.2 Previous studies on headlines	67
3.2.3 Headlines: the language and the functions	69
3.2.4 False beliefs about omissions in headlines.....	73
3.2.5 Why do headline writers omit functional categories?.....	78
3.3 Database of Headlines in Italian and Dutch	80
3.3.1 Introduction: database set-up	81
3.3.2 Effect of finiteness and preposition	84
3.3.2.1 Introduction.....	84
3.3.2.2 Results.....	89
3.3.2.3 Discussion.....	90
3.3.3 Effect of position	93
3.3.4 Noun Phrases in isolation	97
3.3.5 Hanging Topics.....	100
3.3.6 Summary	104
3.4 The experiment.....	105
3.4.1 Introduction	105
3.4.2 Set-up of the experiment.....	105
3.4.3 The conditions and the results	108
3.4.3.1 Effect of finiteness.....	108
3.4.3.2 Effect of position (1): sentence-initial and sentence-internal	113
3.4.3.3 Effect of position (2): order of omission	119
3.4.3.4 Omissions in Nouns in Isolation and Hanging Topic Constructions	123
3.4.3. Summary results.....	126
3.5 Article omission in Child Speech.....	129
3.5.1 Introduction	129
3.5.2 Method.....	130
3.5.3 Results and Analysis.....	136
3.5.3.1 Introduction.....	136
3.5.3.2 Relation with finiteness	138
3.5.3.3 Division based on sentence position of the noun.....	144
3.5.3.4 Division based on type of article	150
3.5.4 Results child speech summarized and compared to headlines	154
3.6 Conclusion	156

Chapter 4 An information-theoretical approach to omission of articles	157
4.1 Introduction	157
4.2 Aspects of the Information Theory and its application to language processing	159
4.2.1 Introduction	159
4.2.2 Earlier applications of Information Theory to Language Processing.....	160
4.2.3 A brief introduction to Information Theory.....	161
4.2.4 From theory to practice.....	169
4.2.5 Application of Information Theory to Processing of Inflected Morphology in Serbian.....	179
4.3 Application of information theory to processing of articles.....	189
4.4 Putting the pieces together	195
4.4.1 Model of article selection	195
4.4.1.1 Article selection in child speech	198
4.4.1.2 Relative entropy in children's files: additional evidence for the relation between omission in child speech and relative entropy	200
4.4.1.3 Article selection in headlines	208
4.4.1.4 Conclusion.....	210
4.4.2 How to account for finiteness and position effect?	211
4.4.3 Differences between article types.....	223
4.4.3.1 The difference between 'de' and 'het' in Dutch	223
4.4.3.2 The difference between definite and indefinite articles in Italian headlines	230
4.5 Conclusion: Why we do not always omit the articles.	234
 Chapter 5 Application of the model to another language: German	237
5.1 Introduction	237
5.2 Headlines database	243
5.3 Headlines Experiment	248
5.3.1 Finiteness effect.....	248
5.3.2 Effect of position(1): omissions from sentence-initial and sentence-internal position and definiteness effect	250
5.3.3 Effect of position(2): order of omission	252
5.4 German Child Speech.....	253
5.5 An information-theoretical account of crosslinguistic findings	256

5.6 Conclusion and further predictions.....	263
CONCLUSION	267
Bibliography.....	271
Appendix-A: Test-Headlines	283
Appendix B.....	297
Samenvatting in het Nederlands	299
Curriculum Vitae.....	305

Acknowledgements

Some people think that writing a dissertation is a lonely job, but they are wrong. It is impossible to do it all by yourself. There are many, many people who supported me. Every single one of them has, in his or her own special way, contributed to the writing of this dissertation, and therefore their names should have been on the cover as well, not only mine.

My first thanks go to my supervisors, Sergey Avrutin and Peter Coopmans. From the first time I met Sergey I was charmed by his exciting new ideas and by his catching enthusiasm for doing science. I'll never forget our inspiring talks on the 'balcony' of ADD, introduced by the famous words: 'Joke, do you want a sigarette?'. I want to thank Peter Coopmans for introducing me to Sergey, and making it possible for me to spend a year in Milan, working on my MA-thesis on a study which made part of the NWO Comparative Psycholinguistics Project. This study formed the starting point for my work with Sergey. But I also want to thank Peter for reading and commenting my work, over and over again, and for asking exactly the right, critical questions at the right moment. Thank you Peter, your help has been indispensable.

A very special 'grazie mille' goes out to Denis Delfitto. When I started studying Italian in 1999 I never thought that I would have liked linguistics, but one class of 'Sintassi' was enough to make me change my mind completely. I am not exaggerating when I say that without his unforgettable enthusiastic and inspiring courses this dissertation would never have been written, for the simple reason that I never would have discovered how fascinating doing linguistics can be.

I would like to thank the UiL OTS for having given me the opportunity to work as an AiO, it will always remain a memorable period of my life. I also thank them for the financial support which made it possible to travel and to present my work at various conferences and workshops, giving me the possibility to discuss my ideas with linguistic colleagues.

I received much support from other linguistic colleagues. I want to express my gratitude to Anna Asbury, Sergio Baauw, Natalie Boll, Elise de Bree, Ivana Brasileiro, Jakub Dotlacil, Jacqueline van Kampen, Annemarie Kerkhoff, Arnout Koornneef, Iris Mulders, Jacomien

Nortier, Maren Pannemann, Hugo Quené, Esther Ruigendijk, Oren Sadeh-Leicht, Rianne Schippers, Marieke Schouwstra, Natalia Slioussar, Sharon Unsworth, Willemijn Vermaat, Nina Versteeg, Frank Wijnen, Shalom Zuckermann, Arjen Zondervan, and many, many others. A special thanks to Nada Vasić, it has been a pleasure to have you as my roommate in ADD and as a colleague in the Comparative Psycholinguistics Project, and to my Italian colleagues in Utrecht and Milan, Flavia Adani, Claudia Caprin, Francesca Foppolo, Nino Grillo, Manuela Pinto and Roberta Tedeschi for our nice Italian discussions.

Special thanks go to Maria Teresa Guasti for her sincere interest in my work, for her indispensable help in the Italian experiment and in the headline database and child speech data collection, and to Aleksandar Kostić, for our fruitful discussions on Information Theory, and on many other interesting things in life, in Utrecht and also in his lab in Belgrade. I want to thank both Maria Teresa Guasti and Aleksandar Kostić for their hospitality when I visited them in Milan, Cambridge and Belgrade.

I would like to thank all the people who participated in the headlines experiments. Without their help this study would not have been possible.

I want to thank all of my family and friends who have always been interested in my work, and have always encouraged me to go on. I mention Elena for really always being there for me when I need her, and for knowing exactly, by some for me completely unknown reason, when her help is needed. I mention Mohamed for helping me to survive a computer crash, Jacqueline for all the ‘gezellige lunches’ we enjoyed in ‘Het Zuiden’. I want to thank Peter, for his unique solution for finding participants for the Dutch experiment, for the reliable copy services, for the cover photo of this book, but in particular for his moral support, and, for the very important ‘klavertje 4 (four-leaf clover)’. And I want to thank Kareen for cheering me up in the difficult last months of writing my dissertation and for reminding me that ‘een dag niet gelachen, is een dag niet geleefd’.

And, last but definitely not the least, my mother:

Mama, jij was er altijd van overtuigd dat het me zou lukken, en je hebt gelijk gekregen. Dank je wel voor je aanmoedigingen, voor het altijd met me meelevende en voor het me ervan overtuigen dat ik altijd moet blijven geloven in mijzelf.

Chapter 1 Introduction

1.1 General Introduction

Children omit articles in their earlier productions. This happens in spite of the fact that in the input children receive articles are among the most frequently used elements in a language. In the past years several accounts have been offered for these omissions of articles in child speech. Most of these attempt to explain the data in terms of deficient linguistic knowledge. For example some propose that the child's grammar is in a state where functional categories have not yet been acquired (or are not yet available) (Radford, 1990, among others), others argue that certain morphosyntactic features have not been acquired (Roeper et al., 2001). However, some phenomena related to the omission of articles in child speech cannot straightforwardly be accounted for by these structural approaches. First, we find crosslinguistic differences in article omission in child speech. Children acquiring a Romance language omit articles less frequently and for a shorter period of time than children acquiring a Germanic language (Gerken, 1991; Leo and Demuth, 1999; Van der Velde, 2004, Rozendaal, 2006, Guasti et al., 2007; among others). A second and more serious problem for these approaches is the fact that none of them can explain the optionality we find in article omission in child speech. Sometimes the child omits the article, sometimes she uses the article. This optionality is not compatible with the above-mentioned type of accounts.

Moreover, children are not the only category of speakers who optionally omit articles. We find a similar pattern of omission in speakers with a specific brain damage, e.g. agrammatic Broca's aphasics (Menn and Obler, 1990, among others). And what is even more interesting and what has so far been largely ignored in the language acquisition literature is the fact that we find omissions of articles in the speech of normal adults as well. We find these omissions in so-called 'special registers' used by adults. In this study I will present data on article omission in newspaper headlines and child speech that show that we find the same crosslinguistic differences between the omission of articles in Dutch and Italian adult speech as we find in child speech (e.g. significantly more omissions in Dutch). We will also observe interesting *similarities* between

the omission patterns of adults and children in Dutch and Italian, like, for example, the fact that there are more omissions from sentence-initial position than from sentence-internal position. These parallels between child speech and adult speech strongly suggest that an account that explains omission of articles in terms of developing linguistic knowledge cannot be the right approach.

More recent approaches have focused on the role articles play in connecting syntactic structure with information structure and suggest that the most problematic area for children (and aphasic patients) is related to the integration of grammatical and discourse-related knowledge, or, in current terms, to the interface between these domains (Avrutin, 2004). In Avrutin's account the optionality, that is the fact that one and the same child sometimes produces and sometimes omits a functional category, is related to a competition between various systems of encoding information, syntax and context. For one and the same subject this competition process may have different outcomes in different circumstances. So far what exactly determines the outcome of this competition process has been an open question. For this reason processing accounts have often been criticized for being 'too vague', 'too intuitive' or 'not scientific enough'.

In this study I will go beyond arguing that syntax may not always be the cheaper option, by explaining why and when it is not. Specifically, I will answer the question *what* determines the outcome of the competition process between syntax and context. I will show *why* the use of an article is sometimes a too 'costly' operation, in terms of required processing resources for children and, under specific circumstances, also for adults (in the so-called 'special registers'). I will show *why* introducing an information unit by syntactic means (e.g. by using an article) is cheaper in Italian than in Dutch. I will give an explanation for these '*why*'s' in terms of processing cost that are required for the selection of an article from the article set. I will show that there are crosslinguistic differences in these processing cost and that these can actually be calculated. I will present a method which enables us to make an exact calculation of the cost of introducing an information unit by syntactic means, e.g. by the use of an article. We will then no longer have to *assume* that the processing cost of articles are too high under certain circumstances, but we will actually be able to *calculate* that they are. This calculation of the processing cost of article selection is based on applications of information-theoretical notions and provides us with a measure of the complexity level of the article selection process in a given language. The

higher the complexity level of the process, the more processing resources are required to select an article. I will show that these differences in complexity level of the article selection process, and the related differences in processing resources that are necessary to retrieve the articles from the article set can explain the differences and similarities in omission patterns we find between Dutch and Italian, in both child and adult speech.

Summarized, the research questions of this study are:

1. Why do children omit articles, in spite of the high frequency of articles in the input?
2. Why do we find crosslinguistic differences in omission patterns of language acquiring children?
3. Why do adults omit articles in special registers, like for example headlines?
4. Why do we find similar crosslinguistic differences in adults' special registers and child speech?
5. Is it possible to devise a model for child language development that accounts for omission and crosslinguistic differences in these omission patterns?
6. Can this model help us in explaining why we find omissions and crosslinguistic differences in adults' special registers?

The structure of this dissertation is as follows. In the second chapter I will present the psycholinguistic background, and propose a base model for article selection using a combination of Levelt's (1989) speech production model, some of the ideas expressed in Avrutin (1999, 2004a, 2004b) and the model for article selection proposed by Caramazza and colleagues (Alario and Caramazza, 2002; Caramazza et al., 2001; Schiller and Caramazza, 2002, 2003). This base model describes the process of article selection under normal and specific circumstances, with normal speakers and speakers with limited processing resources.

The third chapter is devoted to the presentation of my findings on article omission by Dutch and Italian adults in newspaper headlines and by Dutch and Italian language acquiring children. Comparing article omission in both categories of speakers (adults and children) we will see that there are crosslinguistic differences in the patterns of omissions in Dutch and Italian child speech and headlines. But we will also see that there are intriguing similarities in the omission patterns, for example with respect to differences in omissions of articles from sentence-initial and

sentence-internal position and the relation between omissions of articles and the use of a finite verb in the sentence. I will argue that these differences and similarities cannot be explained by the traditional accounts of article omission.

In chapter 4 I will propose a processing account for article omission, using applications coming from the field of Information Theory, which I will incorporate into the base model of article selection I propose in the second chapter. I will argue that, although the base model of article selection is the same in all languages, there are crosslinguistic differences in the complexity level of the article selection process, caused by differences in the probability distribution of the articles in the article sets. The higher the complexity level of the set, the more processing resources are necessary to retrieve the articles from the set. I will show how the differences in complexity level can be calculated by using a measure introduced in Information Theory (Shannon and Weaver, 1949): the relative entropy. I will show how the differences in relative entropy value of the Dutch and Italian article sets can account for the findings on article omission presented in the third chapter.

In the fifth chapter I will show the results of a replication of my study on article omission in child speech and in newspaper headlines in another language, German. The data on article omission in German will show that the crosslinguistic differences we find in the omission patterns of articles in Dutch and Italian are not related to global properties of Germanic and Romance languages. I will show that the differences in article omission in German, Dutch and Italian are related to differences in the processing cost that are required for the selection of an article and that are reflected in the relative entropy value of the article sets. Thus, this chapter will provide further evidence for the strength of the proposed measure of the complexity level of the article set, relative entropy, by showing that it makes the right predictions with respect to crosslinguistic differences in article omission in special contexts or by 'special' speakers.

1.2 Article Paradigm and Article Use in Dutch and Italian

The focus of my dissertation will be on Dutch and Italian. In this section I will give a brief overview of the morphosyntactic properties of the

article system in both languages, and I will discuss differences in the adult system of article-use in both languages.

I will start with a simple but important observation concerning the article paradigm in both languages. The Italian article set contains far more elements than the Dutch article set. The Dutch article set contains 3 elements: the Italian article set contains 18 elements. It will come as no surprise to the reader that this difference in number of elements will lead to a difference in specification of morphosyntactic form-function combinations between the articles in both languages. It is evident that the great variety within the Italian article paradigm provides the possibility of a more clear-cut division of tasks between the elements in the system. Tables 1 and 2 give a complete overview of the different morphosyntactic forms of the articles in both languages:

Table 1 *Morphosyntactic forms of Dutch articles.*

	Definite		Indefinite	
	Common gender	Neuter gender	Common gender	Neuter gender
Singular	de	het	een	een
Plural	de	de	-	-

Table 2 *Morphosyntactic forms of Italian articles*

	Definite		Indefinite		Partitive	
	Mas-culine	Femi-nine	Mas-culine	Femini-ne	Mas-culine	Femi-nine
Singu-lar	il lo l'	la l'	un uno	una un'	del dello dell'	della dell'
Plural	i gli	le	dei degli	delle	dei degli	delle

As these tables show, Italian has a full paradigm of articles: definite, indefinite and partitive articles varying in gender and number, Dutch only has definite and indefinite articles. In Italian, there are also some allophonic variants:

- the masculine articles (lo, gli, uno, dello, degli) have to be used for masculine nouns beginning with
 - s + consonant (lo scoglio (the cliff), gli scogli, uno scoglio)

- z (lo zio (the uncle), gli zii, uno zio)
 - the definite singular feminine and masculine articles (il, la) and the indefinite feminine article (una) can be reduced to l' or un' in front of vowels (il libro di Veronesi, l' ultimo libro di Veronesi (the latest book of Veronesi); la velocità, l'alta velocità (the high speed).

In a rich morphological system each grammatical feature can be reflected in a specific article form, in a poor system this unambiguity can not be achieved. Therefore, in Dutch, articles necessarily have to perform multiple tasks, there is an overlap of functions on a single article. For example the Dutch article form '*de*' performs the functions of singular definite as well as plural definite article:

- | | | |
|-----|-----------------------|-----------------------|
| (1) | de jongen (the boy) | (Italian: il ragazzo) |
| | de jongens (the boys) | (Italian: i ragazzi) |

The Dutch article form '*een*' functions as indefinite article before neuter as well as common gender singular nouns:

- | | |
|-----|--|
| (2) | een huis (singular, neuter gender, indefinite) |
| | a house |
| | (Italian: una casa) |
| |
 |
| | een jongen (singular, common gender, indefinite) |
| | a boy |
| | (Italian: un ragazzo) |

The Italian examples between brackets show that in Italian there is a specific form for each gender. Dutch article forms, coming from a 'poor' paradigm, are more ambiguous in their grammatical functions than the Italian article forms.

If we look at article *use* as it is described in Italian grammar books used by students we find similarities between the use of articles in Italian in Dutch. Very broadly speaking, both in Italian and Dutch the definite article is used with nouns denoting entities or classes of entities which are part of the shared knowledge of both speaker and hearer. This includes unique entities (or unique sets of entities) which are simply part of common human experience, as illustrated in the following examples:¹

¹ All examples are from: Maiden & Robustelli, 2000

- (3) *Il sole è una stella come le altre.*
De zon is een ster, gelijk aan de andere sterren.
The sun is just another star.
- (4) *Adoro il mare*
 Ik hou van *de* zee.
I love the sea.

Both in Italian and Dutch the definite article is used to denote entities which have become part of interlocutors' shared knowledge, for example because they have previously been mentioned. The indefinite article is used to denote entities which do not make part of interlocutors' shared knowledge.

- (5) *Sappiamo che c'è una donna che porta cocaina dal Brasile. E sei tu, vero?*
 No! *La* donna che porta cocaina non sono io.

We weten dat er *een* vrouw is die cocaïne smokkelt uit Brazilië?
 En dat ben jij, toch?
 Nee! Ik ik ben niet *de* vrouw die cocaïne smokkelt.

We know that there's a woman bringing in cocaine from Brazil. It's you, isn't it?
No! I'm not the woman bringing in cocaine.

But, especially with respect to the use of definite articles we also find differences between Dutch and Italian. As the following examples illustrate Italian frequently requires an article (usually the definite article) where Dutch requires none, or a definite article where Dutch requires an indefinite article or a possessive pronoun. We find differences for example:

- in the use the definite article with nouns which have 'generic', 'universal' reference:

- (6) *Il* vino fa male alla salute. (Italian: definite article)
 Wijn is slecht voor de gezondheid. (Dutch no article)
Wine is bad for your health

- in the use of the definite article with names of body parts and other 'inherent' attributes:

- (7) (a) Ha *i* capelli bianchi. (Italian: definite article)
 Hij heeft grijs haar. (Dutch: no article)
He has some white hairs
- (b) Hai *la* macchina? (Italian: definite article)²
 Heb je *een* auto? (Dutch: indefinite article)
Do you have a car?
- (c) Maria si dipinge *le* unghie. (Italian: definite article)
 Maria verft *haar* nagels. (Dutch: possessive)
Maria paints her nails.
- (d) Ha *la* febbre. (Italian: definite article)
 Hij heeft koorts. (Dutch: no article)
He has a/the fever

- in the use of the definite article as an expletive article, used with proper names (of places, countries, persons, etc.):

- (8) (a) Devo parlare con *la* Pisani. (Italian: definite article)
 Ik moet met Pisani spreken. (Dutch: no article)
I have to speak to Pisani.
- (b) *Il* Brasile attraversa un periodo di crisi. (Italian: definite article)
 Brazilië gaat door een ernstige crisis. (Dutch: no article)
Brazil is going through a period of crisis

Differently from Dutch, the Italian article paradigm has partitive articles. In combination with singular nouns the partitive article is used to indicate an unspecified quantity or part of the whole denoted by a noun. In Dutch no article is used in these contexts.

² In Italian the sentence with an indefinite article *Hai una macchina?* can be used if speaker is addressing a car rental office, inquiring whether there is a car available.

- (9) (a) C'è dell'acqua dentro la bottiglia.
Er zit water in de fles.
There is (some) water in the bottle.
- (b) Mangio del pane.
Ik eet brood.
I am eating (some) bread.

In the plural the partitive article serves as a plural form of the indefinite article and it serves to indicate an unspecified quantity or part of the whole denoted by the plural noun. In Dutch no article is used in these contexts.

- (10) (a) Ci sono delle mosche dentro la bottiglia.
Er zitten vliegen in de fles.
There are (some) flies in the bottle.
- (b) Mangio delle ciliegie.
Ik eet kersen.
I am eating (some) cherries.

Chapter 2 Psycholinguistic Background

2.1 Introduction

Although this study is largely concerned with ‘abnormal’ patterns of speech production, it is important to outline first the model of the speech production process formulated for unimpaired adult speakers in normal contexts. Therefore I will present a psycholinguistic background in this chapter, giving a description of Levelt’s (1989) model of language production in normal adults. I will then turn to ‘abnormal’ patterns of speech production and discuss Avrutin’s (1999, 2004a, 2004b) syntax-discourse model. Avrutin proposes that language production and comprehension are the result of a competition process between two different information encoding modules: narrow syntax and discourse. This competition process leads to different outcomes depending on the (discourse) circumstances and processing capacities of the participants in the language process.

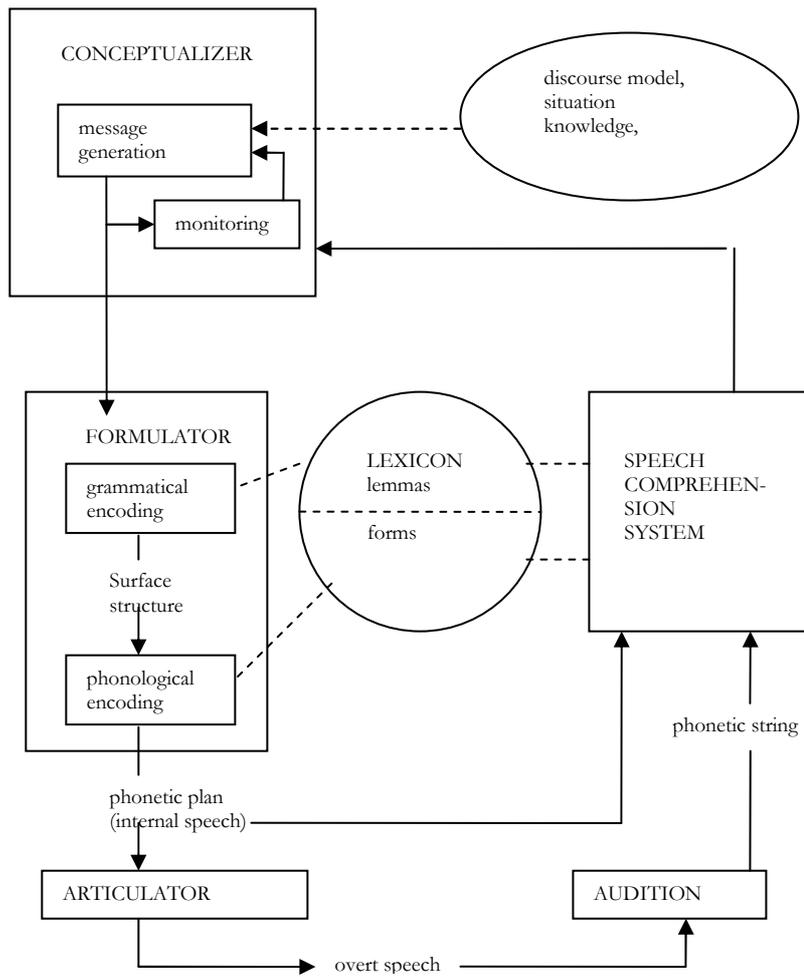
This study focuses on the use of articles. To the best of my knowledge, the only speech production model that focuses on the determiner selection process and on crosslinguistic differences in this process is the model proposed by Caramazza and colleagues (Caramazza et al., 2001, Alario and Caramazza, 2002; Schiller and Caramazza, 2002, 2003; Jansen and Caramazza, 2003). I will discuss this model in section 2.6.

In the last section of this chapter I will propose a model for article selection based on a combination of the three previously mentioned models. This model should explain the process of article selection under normal and specific circumstances, with normal speakers and speakers with limited processing resources. In chapter 4 I will use the experimental results of this study to show that a combination of this model with an information-theoretical model based on complexity of article selection can account for the crosslinguistic differences in article omission in child speech and special registers.

2.2 Levelt's (1989) model of language production

Levelt's model aims at describing the normal, spontaneous speech production of adults. The model consists of a number of processing components. The flow of information within and between these components is depicted in Figure 1. In this figure boxes represent processing components, while circles and ellipses represent knowledge stores.

Figure 1 Partitioning of the various processes involved in the generation of fluent speech (from Levelt, 1989)



In this subsection I will describe the processes that take place within and between the different components. As this study focuses on the stages of the speech production process that take place up to the level of grammatical encoding and lemma selection, I will pay extra attention to the description of these stages.

Conceptualizer

Each speech act starts with a communicative intention. The speaker wants to achieve a goal by means of saying something, and he wants the addressee to recognize this goal from what is said. When a communicative intention has been conceived, the speaker will select information that according to him will be appropriate to reach his communicative goal, and construct a 'message'. The generation of this message in abstract, non-linguistic form takes place at the level of the conceptualizer. Important in the speaker's choice of the informational content of the message is the amount of knowledge shared between him and the interlocutors and his discourse record of what has been said before in the conversation. What has been said earlier or what belongs to shared knowledge among the participants either does not have to be expressed again, or, when expressed has to be shaped in a different form than completely new information. Another aspect of the processes taking place at the conceptualizer level concerns the speaker's choice about the utterance form he has to use for his message in order to realize his communicative goals. So, for example, if the speaker wants to know something he has to use a 'question'. It is at this stage that the level of formality, politeness and directness that are required by the context of the discourse are fixed. A further important process taking place in the conceptualizer is the ordering of information along the level of relative importance. Here, for example, the speaker assigns 'topic-' or 'focus-' hood to referents. Important for the selection of articles is the fact that at this level an index of the accessibility status is assigned to each referent in the message. This accessibility index informs the listener where the referent can be found: in the store with shared knowledge, in the store with general knowledge or somewhere else. This index is an important determinant of the linguistic shape which the next module in the speech production process, the formulator, will compute for the referent. Let me illustrate this with an example from English. Suppose the speaker wants to inform the hearer about the car he bought, and that his thoughts are the non-linguistic counterpart of (1) or (2)

- (1) I bought *a* yellow car
 (2) I bought *the* yellow car

The final form of the utterance produced (with definite or indefinite article) depends on the amount of shared knowledge between participants. Therefore, in the preverbal message not only information about the meaning of the content word ‘*car*’ has to be encoded, but also information about the state of shared knowledge between the interlocutors. In the first example the speaker assumes that the car he bought constitutes completely new information for the hearer. The preverbal message will contain information that indicates that there is no shared knowledge between speaker and hearer with respect to the object ‘*car*’. Therefore an accessibility index will be assigned that indicates that the referent is not accessible in a shared or general knowledge store, but that a new referent has to be stored in the discourse model. In the second example the speaker assumes that the hearer is familiar with the car and the speaker will therefore assign an accessibility index to ‘*car*’ that instructs the hearer to search the intended referent in the shared knowledge store.

Note that the accessibility index forms part of the preverbal message of the noun. In other words, there is no separate preverbal message that will by itself trigger the production of the article. The selection of the article depends on the information on the accessibility status of the referent that is contained in the preverbal message of the noun.

Formulator: Grammatical and Phonological Encoding

The output of the conceptualizer is the input to the formulator. At this level the conceptual structure of the preverbal message is translated into a linguistic structure. This highly automatized process takes place in two steps: the grammatical encoding and the phonological encoding. On the basis of the semantic information in the preverbal message the grammatical encoder will activate a matching lemma in the *mental lexicon*, the store of information about the words in one’s language. This lemma information consists of the lexical item’s meaning and syntax.³ The lemma will activate the grammatical encoder to generate a syntactic structure that is appropriate for the type of lemma (noun, verb, etc.). In

³ For example the lemma *car* is categorized as a count noun, the verb *give* is categorized as a verb which can take a subject expressing the actor, X, a direct object expressing the possession Y and an indirect object expressing the recipient Z.

order to build this structure a categorial procedure is initiated, which is stored in our mental lexicon. For example, a noun-lemma calls upon a categorial-syntactic-noun-environment-building procedure from the mental lexicon, which triggers syntactic processes like, for example determiner selection. Thus, in the case of a DP like *'the car'* the first element to be selected on the basis of the meaning expressed in the preverbal message will be the noun, *'car'*. The *'car'*-noun lemma will call, from the lexicon, a categorial noun procedure containing a building instruction for the syntactic environment of the noun. This environment consists of complements and specifiers which, depending on the concept that has to be expressed according to the preverbal message, will accompany the noun. Therefore the categorial procedure will inspect the preverbal message for modifying or specifying information attached to the concept *'car'*. It will, for example, inspect the concept for number, and find, in this case, the value *'singular'*. It will also check the accessibility status of *'car'* and, in this case, find the value *'+ accessible'*. Then the categorial procedure will start functional procedures that will handle the realization of the complements and specifiers. In the example *'the car'* the functional determiner procedure DET will be activated in order to generate an appropriate determiner. The functional DET procedure not only needs conceptual information on the accessibility status of the referent but it will also need information inherent to the noun, for example number information. The values thus found will be inserted in a list of parameters of the article lemma.⁴ On the basis of the information in the article lemma the functional DET procedure will generate the definite article *'the'*, and deliver it to the categorial noun procedure, which will insert the material received from the functional procedures in the slots of the syntactic environment of the noun. The surface structure thus created, is provided to the phonological encoder, where a phonetic plan is built for the utterance.

⁴ According to Levelt, the exact number of parameters depends on language specific properties: English, for example, does not have a gender distinction for articles, hence the functional DET procedure for English does not have to inspect the gender information. In Dutch DET will have to inspect both gender and number information and in German the functional procedure DET needs gender, number and case information, since the word form of the article depends on the grammatical function of the noun.

Articulator

The articulator converts the phonetic plan into actual speech. In order to do so the output of the formulator is processed and temporarily stored in such a way that the phonetic plan can be fed back to the speech-comprehension system and the speech plan can then be produced as normal speech.

Speech-comprehension system

The speech-comprehension system is connected with an auditory system which plays a role in the two ways in which feedback takes place within the model: the phonetic plan as well as the overt speech are guided to the speech-comprehension system to find any mistakes that might have crept in.

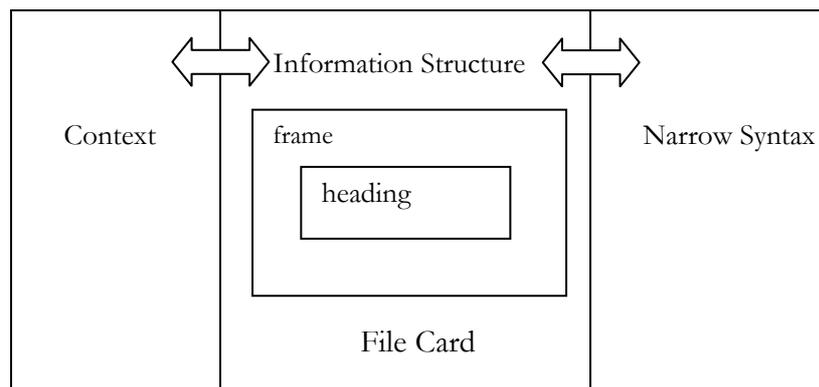
2.3 Avrutin's (1999, 2004a, 2004b) syntax-context model

Levelt's model is a very useful, well-argued description of the modules involved in the process of language production. It focuses on the speech act, that is, the formulation and articulation of the message, but it is based on the 'normal' speaker in the 'normal' discourse situation. In these conditions retrieval of the lemmas will create a 'normal' syntactic environment for the lexical elements in the sentence. Because of the highly automatized nature of the process, the outcome, given a certain input, will always be the same. The produced sentence is a 'normal' grammatical sentence, with all syntactic elements present. Levelt's model therefore cannot answer the question why we find differences in language production between normal adults on the one hand and, for example, language acquiring children or agrammatic patients on the other hand. A model that does answer this question is the one proposed by Avrutin (1999, 2004a, 2004b). Avrutin shows that differences in production and comprehension between normal adults and agrammatic patients (and children) are related to a specific partitioning of the language production and comprehension processes: the level of Information Structure, where we find interaction between the linguistic modules syntax and context.

Avrutin takes Heim's (1982) File Card Semantics as a starting point of his model. In Heim's discourse model, in each conversation the information is stored as in a library file catalogue: each DP introduces a file card with certain information about the corresponding entity. Each

file card contains a heading and a number that allows speakers to keep track of the information and to update discourse entities. Avrutin develops this model further and proposes that not only individuals but also events form independent discourse entities, each with its own file card. We thus have *individual file cards* for individuals, formed by DPs, and *event file cards* for events, formed by VPs. Avrutin suggests that each file card should be seen as an information unit that is processed at the level of Information Structure⁵. Information Structure is part of the computational system involved in language. It is the intermediate level between narrow syntax and our system of thought where the output of narrow syntax is ‘translated’ into information that is interpretable to our system of thought.

Figure 2 *The modules in Avrutin’s syntax-context model*



Let us now focus on this level of Information Structure and examine the processes that take place at this level more closely, by raising the following questions:

- a. What are the basic elements of the units of information that are processed at the level of Information Structure? To put it more concretely: what are the basic elements of the File Cards?
- b. How do these basic elements become available to the level of Information Structure?

⁵ Avrutin speculates that this level is largely analogous to Chomsky's (1995) Conceptual Interface.

c. Which operations take place at the level of Information Structure. Alternatively: what happens with the File Cards at the level of Information Structure?

a. What are the basic elements of the File Cards?

The units of information that are processed at the level of Information Structure are the File Cards. Each File Card consists of a *heading* and a *frame*. The heading of a File Card provides information about the referential content of the information package, such as person, object, place, event, etc. Therefore the heading is usually formed by a content word (lexical category). The frame of the File Card has the role of anchoring the information package in the right way into the discourse. It contains instructions, one might say, about the way in which the lexical content word has to be related to the context. In ‘normal’ adult speech the frame is provided by a functional category.

In a DP like ‘*a/ the car*’ the noun ‘*car*’ provides the lexical content, the ‘heading’. The article ‘*a*’ or ‘*the*’ gives instructions about how the ‘*car*’ has to be related to the context, whether it is a newly introduced referent or a referent that is considered to be known to the interlocutors.⁶

b. How do the basic elements become available to the level of Information Structure?

There are two ways, or in Avrutin’s terminology two channels, through which the basic elements can become available to the level of Information Structure: *narrow syntax* and *context*. I will first describe these two channels. How are they constructed, how do they function and what does their output look like?

Narrow syntax can be seen as an autonomous computational system that selects lexical and functional categories from the mental lexicon and manipulates them (putting them in a language-specific order, coding agreement relations etc.) (Chomsky 1995).

Context is formed by the ‘surroundings’ of the utterance in a very broad sense. It contains information about the linguistic surroundings, such as knowledge on what has been said before and on what is said during the discourse. It further contains information about the non-linguistic surroundings such as shared knowledge between speaker/hearer, gestures, presence of the denoted individual(s) or

⁶ The function of the File Cards in Avrutin’s model is similar to the function of the lemmas in the model proposed by Levelt. Both, the File Cards and the lemmas have the role of ‘mediator’ between context and narrow syntax/grammatical encoding.

event(s) in the context.⁷ Differently from the output of narrow syntax, the output of context to the level of Information Structure does not have an overt verbal form. Avrutin suggests that in principle all information (information contained in functional words, but also information contained in lexical, content words) can be supplied by the context, in which case no overt output will be provided to the Information Structure Level. He further speculates that if information is provided by context, some kind of non-verbal signs (e.g. gestures, facial expressions, body movements) or sounds (for example ‘mmm’, ‘ah’,) may possibly occur in the act of communication, as an indication or an ‘overt’ expression of the fact that information has been provided by the context.⁸

Let me illustrate Avrutin’s proposal with two examples, which illustrate the introduction of an individual unit of information for a DP in ‘normal’ adult speech and in a situation in a ‘specific context’.

⁷ This aspect of the Context channel in Avrutin’s model is largely analogous to what in Levelt’s model is called ‘Discourse model, situational and encyclopedial knowledge’

⁸ And in fact, a number of recent studies (Goldin-Meadow, 2006; Joutteau 2006) have shown that gestures that are produced along with verbal speech form a fully integrated system with verbal speech and that non-verbal signs can convey substantive information that is not (or even cannot be) expressed by verbal speech. In a study on gestured functional heads of the left periphery in Breton Joutteau (2006) showed that use of non-verbal signs can make an unacceptable utterance acceptable. Sentence (i) without an overt preverbal subject and without a gesture is ungrammatical. However, if an appropriate gesture (for example movement of the eyebrow or of the upper body) is used, the sentence is fully acceptable.

(context: *Someone is looking desperately for something, but it is late...*)

(i) *trouvera ça une autre fois
will find that another time

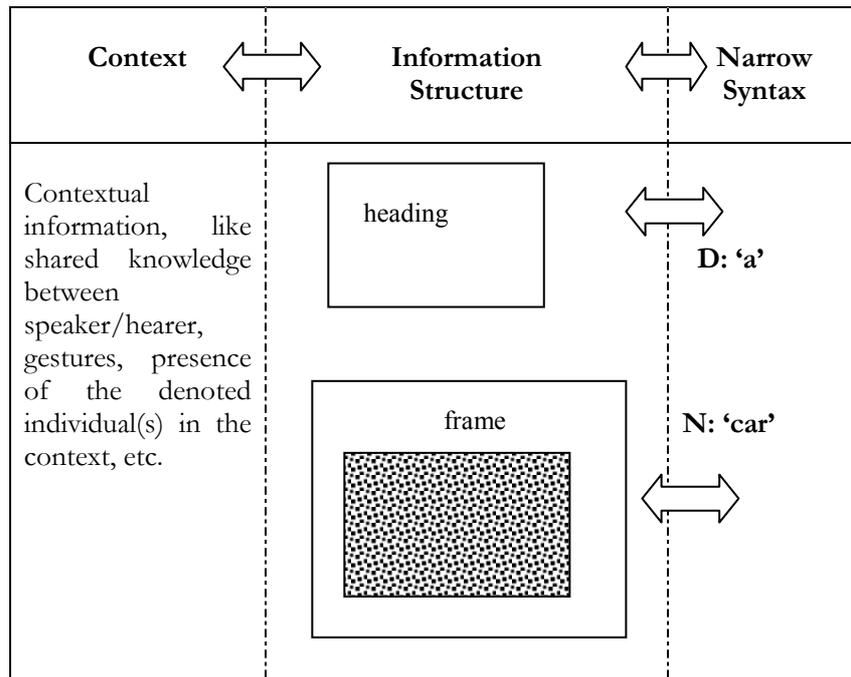
(ii) [gesture] trouvera ça une autre fois

Whether or not Breton is a V-second language is a debated issue in the literature (Borsley and Kathol, 2000), however the fact that in colloquial speech gestures can save the acceptability of the omission of the subject is worthy of notice.

DP introduction in normal speech

In 'normal speech' both the heading as well as the frame will be provided to the level of Information Structure by Narrow Syntax. This process is illustrated for the DP 'a car' in Figure 3.

Figure 3 *Frame and Heading provided to Information Structure Level by narrow syntax.*



It is important to note that in this situation too both channels, narrow syntax *and* context, are available to the level of Information Structure. But, here the necessary elements, frame and heading, are both provided by narrow syntax. This is, however, not always the case.

DP introduction in specific contexts

Avrutin observes that sometimes functional categories are missing in adult speech. This happens in the so-called 'special registers' like, for example, question-answer pairs in colloquial speech, with a strong presupposition by the listener.

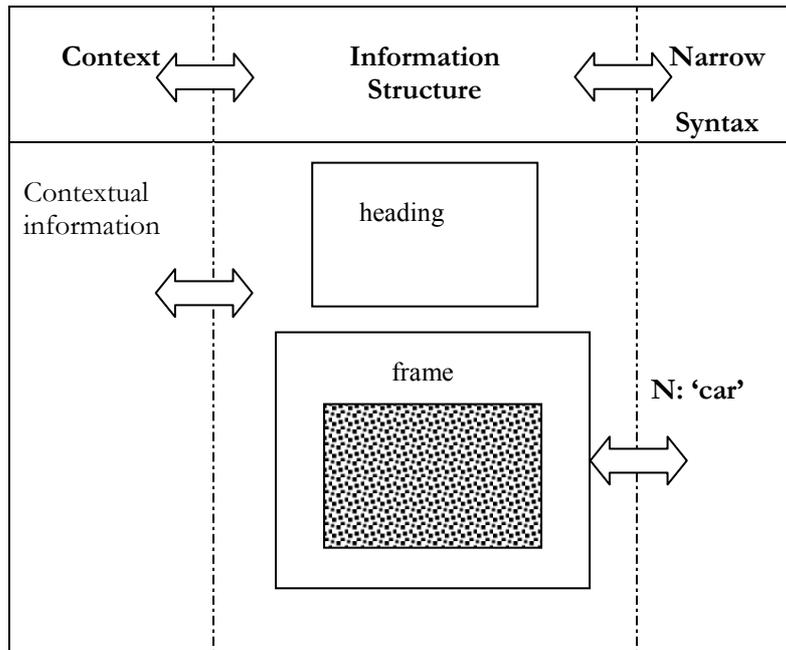
- (3) Q: Hebben jullie dat parket zelf gelegd?
Did you place this floor all by yourselves?
A: Ja, gigantisch karwei!
Yes, gigantic job!
- (4) Q: Wat is er gebeurd?
What happened?
A: Oh, ik zie het al! Gebroken nek!
Oh, I can see it now! Broken neck!

Not only in question answer-pairs, but also in other colloquial conversations in which a considerable amount of shared knowledge is present between the interlocutors, do we find omissions of functional categories in adult speech. Interestingly, as some of the following examples show, we sometimes even find several omissions in one utterance. The following examples come from conversations *between adults* that were recorded in the Chiles files.

- (5) O, mijn God, nieuwe rage, beest op zijn kop neerzetten!
Oh my god, new rage, putting animal upside down!
(note: two omissions of articles)
(File Abel 20729, frase 1318)
- (6) Laatste weekenden wat vrienden op bezoek gehad.
Last weekends some friends (=object) visiting us.
(note that in this example also the subject and the auxiliary are omitted)
(File Tom 11011, frase 231)

Translated in the terms of the File Card Model this would mean that we have a heading without a frame, which normally speaking is not allowed. Still the sentences are fully acceptable, but *only* when they are used in a particular context. This leads Avrutin to the conclusion that in these cases the role of the functional category is taken over by the context, and that it is in fact context that provides the frame for the heading in these cases.

Figure 4 Frame provided to Information Structure Level by Context, and Heading by narrow syntax.



In this situation both available channels provide elements to the level of Information Structure. The heading 'karwei' ('job') is a lexical element provided by narrow syntax, the frame, however, is provided by the context, through a strong presupposition by the speaker that the listener is able to interpret his utterance relying on contextual information. In this case context provides the frame and makes the utterance acceptable. The fact that in some cases (see example 6) we find several omissions of functional categories within the same utterance provides extra evidence for the fact that context is particularly salient in these utterances. It is important to note that context plays a role here in the *production* of the utterance, by supplying the frame that is necessary to make the utterance interpretable.

Let us now proceed to the last question:

c. Which operations take place at the level of Information Structure?

At the level of Information Structure the basic elements of the message have to be made fully interpretable for our cognitive system. Therefore

at this level the heading and the frame will be bundled together into a unitized information package, the File Card. This file card represents a *unit of information*. To be fully interpretable the units of information structure must necessarily contain both parts: there can be no frame without a heading, nor can there be a heading without a frame. It has to be a unitized bundle of heading and frame⁹. A further observation is that a unit of information consists of only one frame and only one heading. We have seen that the elements can become available through two channels. This means that it is possible that at the level of Information Structure a competition arises between elements that can in some (or, arguably, all) cases become available through two channels, narrow syntax and context. But only one frame is possible.¹⁰ How is this situation resolved?

Avrutin argues that the outcome of this competition process is based on economy considerations.¹¹ In the case of two equally suitable options the option that can become available at the lower amount of processing cost will be chosen. Thus, the channel that will win the competition will be the channel through which the required element becomes available with less processing effort, hence the channel with the lower resistance. In a fully developed unimpaired language system the morphosyntactic encoding of information is the more economical route, which is due to the automatized nature of language. It may be instructive, he argues, to view the two systems, narrow syntax and context, as competing with each other for the ‘right to encode information’. The cheaper option

⁹ Avrutin (2004a,b) argues that these requirements follow from rules of information packaging and transmission that apply to message transmission in any communication system.

¹⁰ A good illustration of the fact that only one frame is possible can be found in the study of Jouisseau (2006) on gestured functional heads of the left peripher, see footnote 6. She found that, although gestures in colloquial speech in Breton could take over the function of the preverbal subject pronoun, a sentence with both a gesture and an overt preverbal subject pronoun is pragmatically marked, in the sense that it changed the meaning of the sentence. So either context or syntax had to be chosen to provide the subject pronoun. If both are chosen the utterance expresses another meaning.

¹¹ The notion of economy has played a major role in recent theoretical work (e.g. Chomsky 1995). There have been various ways of implementing this idea. One of them is Avrutin, who takes the notion of economy one step further and implements it in an approach to language processing. He suggests that it reflects the amount of resources utilized by the brain when performing a specific computation. Avrutin claims, based on experimental evidence (Avrutin, 2000) that for normal adult speakers the narrow syntax operations are the cheapest.

wins. This explains our normal reliance on morphosyntax rather than context and this is why under normal circumstances speakers encode information using the morphosyntactic system. In other words, this explains why we talk rather than presuppose. Taking this idea somewhat further, Avrutin suggests that the morphosyntactic route is not fully automatized in young children (most likely due to their limited processing resources available for lexical retrieval) and that it is ‘weakened’ in the case of Broca’s aphasia too, as a consequence of the limitation of resources for lexical retrieval.¹² The relevant evidence comes from a variety of studies, such as gap filling experiments (Zurif et al 1993), lexical access studies (Swinney et al. 1989) and lexical decision (Piñango and Burkhardt 2001). In both cases, child and aphasic speech, the use of morphosyntax, which in normal speakers is the cheaper route, becomes less efficient, or formulated differently: more ‘expensive’ . Therefore, the omission of certain morphosyntactic elements we find in the output of child and aphasic speech can be viewed as a defeat of the (no-longer-the-cheaper) morphosyntactic system by the competing contextual route.

2.4 Levelt’s model and Avrutin’s model combined

The two models of language production described in the preceding sections have different objectives. Levelt describes a model of language production of a normal speaker in normal conditions. In these conditions retrieval of the lemmas will create a ‘normal’ syntactic environment for the lexical elements in the sentence. Because of the highly automatized nature of the process, the outcome, given a certain input, will always be the same. The produced sentence is a ‘normal’ grammatical sentence, with all syntactic elements present. Given the importance of the input in his model, Levelt pays considerable attention

¹² Broca’s aphasia is a language disorder that is caused by brain damage, such as a cerebro-vascular accident or a traumatic brain injury. In the present literature Broca’s aphasia is generally associated with a lesion in the left frontal lobe of the brain. Characteristic of the speech of Broca’s aphasia patients is a non-fluent, halting and telegraphic speech output. Such patients rely on the simplest possible structures of their language. Their utterances are reduced to mainly content words, function words are frequently omitted (Goodglass & Kaplan, 1972, Linebarger et al., 1983, Prins, 1987, Menn & Obler, 1990, among others).

to the stage of the conception of the message: the preparation of the preverbal message that triggers the construction of the syntactic environment.

Avrutin's model aims at explaining why in 'non-normal' circumstances, like for example conditions in which specific contextual requirements are met or in the case of speakers with limited processing resources, we find different outcomes from what we find in the case of a normal speaker in normal conditions.

Despite different terminology, the models have important characteristics in common and to a large extent the models agree on the role of the modules they have in common. The role of narrow syntax in Avrutin's model is comparable to what in Levelt's model is called 'grammatical encoding' at the level of the formulator. The role of the conceptualizer in Levelt's model shares important properties with the role of the information structure level in Avrutin's model. Both are intermediate levels between context and the level where the grammatical encoding takes place. In addition, I suggest that where the models clearly diverge, they do so in a non-contradictory fashion. The aspects on which they differ are aspects that play an important role in one model, but are not taken into consideration in the other model. Let me explain this in more detail.

- Avrutin discusses extensively the competition process that takes place at the Information Structure Level, a competition process that in the case of abnormal language processing situations may lead to the creation of a non-normal syntactic environment. Levelt does not take into consideration this competition process, since he focuses on normal language processing situations, in which a normal syntactic environment is created.

- In Levelt's model differences in output can be explained only by differences in input. Therefore he pays considerable attention to the construction of the preverbal message, in contrast to Avrutin. However Levelt's proposals on the formulation of the preverbal message are very much compatible with Avrutin's model, and can be seen as an analysis of one of the processes taking place in our cognitive system.

I therefore suggest that the two models are in a complementary relation to each other. I propose that a combination of the models provides us with a more comprehensive view on language processing than each of the models viewed separately. In the last section of this chapter I will present my proposal for a model that is based on a combination of Levelt's and Avrutin's model, but first I will examine an

important aspect that is lacking in both models: a detailed view on which factors are decisive in the determiner selection process. To the best of my knowledge, the only model that focuses on the determiner selection process in particular, and on crosslinguistic differences in this process, is the model proposed by Caramazza and his colleagues (Alario et.al., 2002). I will discuss this proposal in section 2.6. Before turning to it, I will discuss in the next section some crucial differences between the production of open class versus closed class words in general, and different models that can be proposed for the production of closed class words, focusing on possible models for determiner selection.

2.5 Language production of open class versus closed class words

2.5.1 Introduction

A major distinction that can be made in words used in natural language is the distinction between open-class words (content words, such as nouns, verbs and adjectives) and closed-class words (function words, such as auxiliaries, determiners and prepositions). An important difference between these two wordclasses is the information that is used to select the words:

- selection of open class words depends primarily on their individual meanings.
- selection of closed class words depends partly on properties of other words in the sentence (number, gender, phonological context). Besides these 'inherited' properties closed class words can also have an individual semantic meaning, independent of the other words in the sentence (a typical example is the distinction definite/indefinite for articles).

Therefore, if we look at the architecture of language production models proposed by different researchers (e.g. Levelt, Roelofs & Meyer, 1999, Dell, 1986), we find differences between the representation of open class and closed class words. For open-class lexical nodes the only input comes from the semantic system. Processes involved in closed class items require a different architecture. If we focus, for example, on the article, the information necessary for retrieval of the article can come

from different sources: semantic (definite/indefinite), grammatical (gender/number) and phonological (if the article form depends on the phonological properties of the following word, as for example in Romance languages). This means that in a language production model an article node will receive activation from different systems. To make things even more complicated, these different types of information do not become available simultaneously. Most language processing studies assume that semantic properties are activated before syntactic properties (see for production: Levelt, 1999, for production *and* comprehension: Schmitt et al. 2001, Müller and Hagoort, 2006), followed by phonological properties (Levelt, 1999, van Turenout et al. 1998). This raises the question whether the different types of information are used individually, each by themselves contributing to the activation of an element, or whether they are first integrated in a kind of configuration, which is then used to activate and retrieve the element.¹³ In addition, the information necessary for article retrieval depends on language specific requirements. English, for example, has no gender agreement rules and Dutch has no phonological context rules. Such observations raise the question whether and how production of closed-class words can fit in a language-universal processing model.

2.5.2 Determiner selection models

The architecture of the determiner production process can be modelled in various ways (Janssen & Caramazza, 2003)

- hierarchical activation model
- frame activation model:
 - o with cascaded processing
 - o with unitized activation

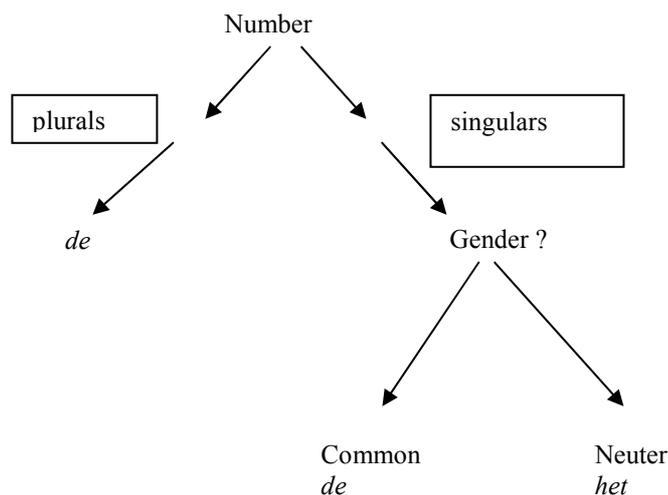
Examining these different models gives us good insight in the processes involved in determiner selection and in the structure of these processes. For this reason I will discuss these different models in some detail in this section.

¹³ A combination of these two possible options is also possible.

Hierarchical activation model

In a hierarchical activation model *only* the sufficient conditions for determiner selection are considered. When a relevant feature for the determiner becomes available, it activates its associated determiner form(s). Selection takes place as soon as sufficient information is available for selection of the appropriate determiner. For example, according to this hypothesis, selection of plural determiners in Dutch can take place as soon as the feature *plural* becomes available. Let me explain this prediction. Previous studies have shown that information about grammatical number becomes available before information about a noun's gender (see Vincenzi & De Domenico, 1999). Since in Dutch plural definite NPs the feature 'plural' uniquely specifies the determiner form '*de*', the article can be selected at the moment number information becomes available. The selection process does not have to wait until gender information becomes available, since, regardless of the gender, the determiner will always be '*de*' (see Janssen and Caramazza, 2003 for a discussion). In the selection of a singular definite NP, however, gender information is necessary to select the determiner. Figure 5 illustrates the hierarchical control structure of the Dutch determiner system with respect to number and gender information:

Figure 5 Schematic representation of the relationship between number and gender in the selection of determiners in Dutch



Frame activation models

Frame activation models propose a different way of determiner selection. In a frame activation model determiners are represented by frames consisting of slots into which the features relevant for the selection of the determiner are inserted.¹⁴ We can distinguish between two different types of frame activation models, those which advocate unitized activation and those that advocate cascaded processing.

Unitized activation models

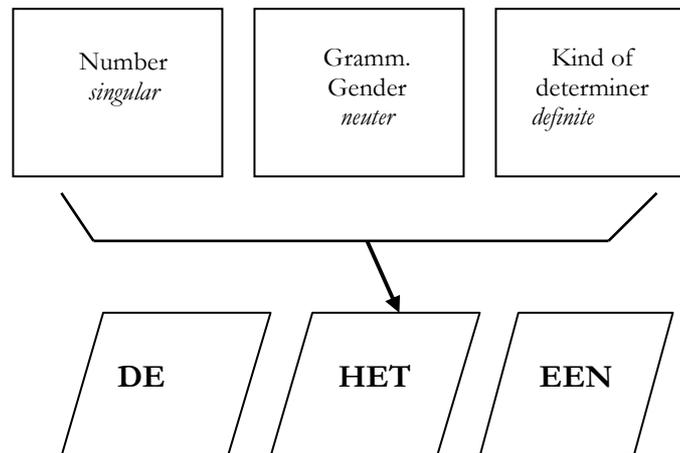
In a frame activation model based on *unitized activation* no activation will be sent to the determiner forms before the frame is fully filled (Miozzo et al. 1999). The central idea is that all activated features act together as a single 'information unit' to activate and retrieve the determiner form.

¹⁴ To avoid confusions: the frames in a frame activation model should not be confused with the frames previously mentioned in the discussion of the File Card Model. Frames in a frame activation model are filled with features activated during the selection process of the element (phonological, grammatical, semantic, depending on the theory) and form the basis for selection of the element. A frame in the File Card Model is filled with a functional category provided by syntax or by information from the context.

There is no individual contribution of the activated features to the activation process of the determiners. The internal structure of the ‘bundle of features’ that is addressing the retrieval of the determiner does not play any role. Only when the determiner frame is fully specified will the configuration of the filled frame activate the specific determiner with which it is associated. All the features that are associated with the determiner selection process have to be available before selection can take place, even the features that are not strictly speaking necessary for specifying the correct determiner form.

Let me illustrate this with an example from Dutch. The Dutch singular word ‘*boek*’ (book) has neuter gender, and hence the form of the definite article is ‘*het*’. In a unitized activation model, if and only if the article frame is fully specified with respect to the kind of determiner (*definite*), and the gender (*neuter*), and the number (*singular*) activation will be sent to the determiner form ‘*het*’. As long as the frame is not completed, no activation will be sent to any of the determiners in the set. Figure 6 illustrates determiner selection in a Unitized Activation Model:

Figure 6 Schematic representation of the Unitized Activation Model, all required information is used together as a bundle to retrieve the determiner.



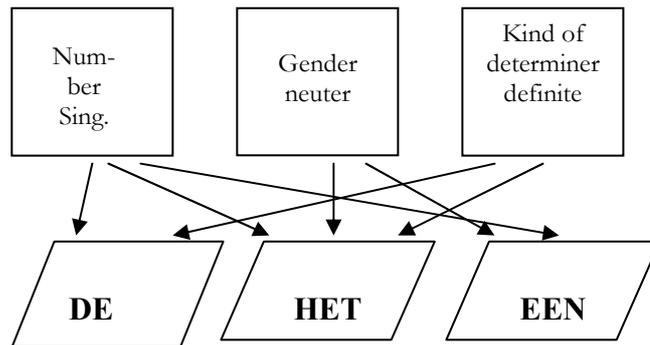
Cascaded processing (individual activation)

A frame activation model based on *cascaded processing* proposes that individual features that are activated send activation to the determiner with which they are associated. Using the example given above of the Dutch *'het boek'*, in a cascaded processing frame activation model the feature *definite* sends activation to *de* and *het*, the feature *singular* sends activation to *'de'*, *'het'* and *'een'*, and the feature *'neuter'* sends activation to *'het'* and *'een'*. In this way determiners receive activation from all the individually activated features. This means that more than one determiner will be activated in the selection process. In the example of *'het boek'* all determiners receive activation. The activation level of determiners will vary as a function of the amount of input they receive.¹⁵

What this type of model reflects is the general principle of cascaded processing (e.g. Roelofs, 1997, 1992, Dell, 1986), which assumes that activated nodes in the system continuously send activation to their linked nodes. In a model based on cascaded processing it is not the case that all informational features act together as a *single unitized* 'information chunk' and that the internal characteristics of this 'chunk' do not contribute individually to the process of determiner retrieval. Rather the opposite is true. Each piece of information contributes independently to the process of activation and the final process of determiner selection arbitrates between the pre-activated candidates. The candidate with the highest level of activation will win the competition process.

¹⁵ In addition, the activation level of determiners will, according to Caramazza (Alario and Caramazza, 2002), also vary with the frequency of the association. To illustrate this effect of the frequency of association with an example: if a specific noun (for example an inherently unique noun like *'zon'* (sun) occurs more often with the definite determiner *'de'* than with the indefinite determiner *'een'*, the strength of activation of *'de'* with that specific noun will be stronger than the strength of activation of *'een'*. Hence it will be easier to select *'de'* than *'een'* with that specific noun.

Figure 7 Schematic representation of the Cascaded Processing (Individual Activation) Model: each feature activates the determiner forms independently



2.6 Caramazza's determiner selection model

2.6.1 Introduction

Many studies on language processing have focused on the production of open class words (Levelt, 1989, Roelofs, 1997). The processing of closed class words has received attention in a number of studies that generally compare the processing of open class words with the processing of closed class words. An example is the study by Bradley (1978). He examined the performance of normal and aphasic speakers on open- and closed-class words and argued that normal speakers have a specialized closed-class word retrieval system, while aphasic speakers do not have this specialization.¹⁶ Less attention has been paid to the processing of the different types of closed class words (pronouns, determiners, inflection,

¹⁶ Using frequency sensitivity as a diagnostic for lexical recognition performance, Bradley found, for normal speakers that such sensitivity was present for open-class words, but absent for closed-class words. In contrast, agrammatic aphasics showed frequency sensitivity for both the open and the closed class. On these grounds Bradley advanced the hypothesis that the syntactic problems of agrammatics might be a symptom of the failure of a specialized closed-class retrieval system.

prepositions, etc.) on an individual level. We do find a large number of studies on pronoun resolution (Vasić 2006, Vincenzi and Di Domenico 1999, Zurif et al. 1993, among others). But other closed class elements, like for example prepositions, inflection but also determiners, have received less attention.

To my knowledge, the only comprehensive group of studies on the processing of articles are the studies by Caramazza and his colleagues (on determiner selection in Spanish and Catalan: Costa, Gallés, Miozzo and Caramazza, 1999; in Italian: Miozzo & Caramazza, 1999; in French: Alario and Caramazza, 2002; in Dutch: Janssen and Caramazza, 2003; in German and Dutch: Schiller and Caramazza, 2003). In his investigation of how determiners are selected Caramazza focuses in particular on what information is responsible for the lemma activation of determiners. Based on a large number of crosslinguistic experiments in Germanic and Romance languages Caramazza proposes the Primed Unitized Activation Model, which I will discuss in detail in this section.

As we will see, in his studies Caramazza suggests that determiners are ‘special’ from a processing point of view, in the sense that differently from what is found with lexical categories, which are universally processed in the same way, we find crosslinguistic differences in the processes that lead to selection of determiners. According to Caramazza, the determiner production process is tuned to language-specific properties.¹⁷ In this section I will give a description of Caramazza’s model, but I will also point to an intriguing aspect of the use of articles that seems to be at odds with his model.

2.6.2 Primed Unitized Activation Model, a description and basic assumptions

Alario and Caramazza (2002) suggest a ‘hybrid’ determiner selection model that combines aspects of both types of the frame activation model: the one with cascaded processing and the one with unitized activation. They suggest that determiners are represented by means of a language specific determiner frame with slots that must be filled with

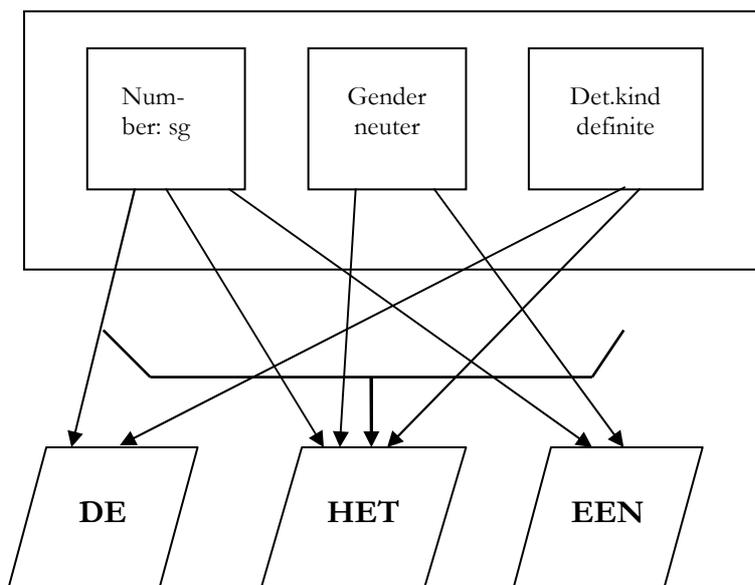
¹⁷ The same is true for speech perception; Mehler et al. (1993, 1996) have shown that processing routines (pre-lexical segmentation, word segmentation, etc.) are not identical in different languages but are fine-tuned to the specific properties of the native language

feature values. Each type of information about the determiner is represented by a different slot, so there is a slot for discourse/pragmatic information (+/-def), there is a slot for grammatical gender and there is a slot for information about the phonological form of the following word.¹⁸ What is important in their model is the fact that *all* slots have to be filled in before the required determiner form will be retrieved, as in a unitized activation frame model. So only when information about the discourse pragmatic meaning, the grammatical gender and the phonological form of the following word are all simultaneously active in the system can the appropriate determiner form be selected. Selection of the appropriate determiner form has to wait until all slots are filled in.

But, differently from what we find in a pure unitized activation model the authors propose that it is not the case that *activation* of the different determiner forms has to wait until all slots are filled in. In fact during the selection process all determiners corresponding with the feature content of a specific slot will be activated, like in a model based on cascaded processing. Let me illustrate this with the Dutch '*het boek*' example again. All determiner forms compatible with the information 'singular' ('*de*', '*het*' and '*een*') will be pre-activated to some degree when the gender information '*singular*' is specified. So in a kind of intermediary stage during the selection process all determiner forms specified for '*singular*' are activated, even the ones that may not be compatible with other feature specifications like for example, the discourse/pragmatical information. During the final process of determiner selection a choice is made between the activated candidates on the basis of their level of activation at the moment of determiner selection. See Figure 8 for an illustration of this process.

¹⁸ There will be a slot with information about the phonological form of the following word if the language, has determiners that depend on this form. For example, in Italian the masculine singular definite article that is used before words beginning with a consonant is 'il', '*il ragazzo*' (the boy), but if the following word begins with a 's + consonant', or a 'z' the form of the masculine singular definite article is 'lo': '*lo studente*' (the student), '*lo zio*' (the uncle).

Figure 8 *Primed Unitized Activation Model: note that there is activation from the individual features, as well as from the completely filled frame.*



To summarize, the three basic assumptions of Caramazza's Primed Unitized Activation Model are:

1. *Determiner selection follows a frame-based activation model, not a hierarchical activation model*
2. *Principle of unitized activation: All slots of the determiner frame have to be filled in, no determiner selection is possible if this condition is not fulfilled.*
3. *Principle of cascaded processing: individual determiner forms corresponding with the specification of a specific slot will be primed prior to final selection process.*

In addition, as I will show in the following section, Caramazza claims that there are important crosslinguistic differences in the time course of the determiner selection process. He proposes a distinction between so-called 'early-selection-languages' like Dutch and German and 'late-selection-languages' like Italian, Spanish, Catalan and French. This

difference depends on the information that is necessary to retrieve the determiner and the moment in the production process of the NP when this information becomes available.

2.6.3 Crosslinguistic differences: early selection languages versus late selection languages

Caramazza has conducted several experiments in a variety of languages on determiner selection in which he measured the reaction time in the naming of target DPs in the presence of a distractor word, with either congruent or incongruent gender specification (Janssen and Caramazza, 2003; Schiller and Caramazza, 2003, Miozzo and Caramazza, 1999; Costa, Miozzo and Caramazza, 1999; Alario and Caramazza, 2002).¹⁹ Interestingly he found that in Dutch and German the RTs of singular nouns were influenced by the grammatical gender of the distractor: the RTs were longer when the distractor had a different gender than the target word.²⁰ In Romance languages, however, there was no influence of the gender of the distractor word on the RT of the target word. Caramazza argues that the reason for this difference lies in the fact that in Romance determiner selection happens late during NP construction. In a number of cases in Romance languages the phonological properties of the following word are required to perform the selection. Thus, for example, in Italian whether the masculine determiner form *'il'* or *'lo'* (*'i'* or *'gli'* for plural) is selected depends on the phonological characteristics

¹⁹ For example the target word is *'de hond'* (the dog, common gender), and the distractor word in the condition with congruent gender is *'de koe'* (the cow), and in the condition with incongruent gender *'het schaap'* (the sheep, neuter gender).

²⁰ The fact that there was a gender congruency effect in the Germanic languages only in the condition with singular DPs and not in the condition with plural DPs shows that the gender congruency effect is related to the activation of competing determiners, and not to the activation of (in)congruent gender features of the noun. In Dutch and German plural NPs the plural definite determiners always have the same form, regardless of the gender of the noun; only in singular NPs the form of the definite determiner depends on the gender of the noun. Hence, if the delay in RT in the condition with incongruent gender were caused by only the gender features of the noun, and not by an effect of determiner competition, there should have been influence on RT in the plural, as well as the singular condition. However, there only was influence of incongruent gender in the singular condition, where competing determiners were activated.

of the following word. If the following word starts with a combination of ‘-s’ and another consonant the masculine article form is ‘*lo*’ in singular’ and ‘*gli*’ in plural. In other cases it is ‘*il*’ in singular and ‘*i*’ in plural. Thus we find for example: *il gatto* (the cat), and *lo strano gatto* (the weird cat). An implication of these phonological constraints on the determiner selection process is that in Italian, contrary to Dutch, both syntactic (the noun’s gender) and phonological information (the onset of the following noun or adjective) must be available before the form of the determiner can be selected.

Caramazza argues that these differences affect the way in which determiners are selected in both languages. Crucial in his line of reasoning is the fact that a word’s phonological content becomes available later than its gender feature (see Van Turenhout, Hagoort & Brown, 1998, for evidence). In Dutch determiners do not depend on the phonology of the words that follow them and can be selected relatively early in the course of NP production. In Italian, by contrast, the selection of a specific determiner form will occur relatively late in the course of NP production. This explains why gender-congruity effects are observed in Dutch but not in Italian. In Dutch the selection of the gender (and the number) of the target word is sufficient to initiate the selection process of the form of the determiner. Thus a gender-incongruent distractor word will activate a different determiner form, leading to a competition effect between the two activated forms at a crucial moment in the determiner selection process. Because of the competition process between the two different activated determiner forms, the selection threshold for the correct determiner will be reached later, and production will be delayed. In Italian, knowledge of the noun’s phonological form is necessary for the selection of a determiner form. Therefore, in Italian, the selection of a determiner form cannot begin immediately at the selection of the word’s gender feature, but must wait for the selection of the relevant phonological context. As a consequence, in languages like Italian determiner selection occurs so late that activation of the gender information of the competing noun, and the associated determiners, will have dissipated at the moment of determiner selection, and thus not be able to offer any competition.²¹

²¹ This does not mean, however, that in Italian a *determiner* competition effect can never be observed, Miozzo and Caramazza (1999) showed that in Italian a determiner competition effect does exist, but it becomes visible only at a very late stage during the

2.6.4 An apparent contradiction between data from acquisition studies and data from processing studies

Caramazza's studies offer a very useful insight in the determiner selection process and in crosslinguistic differences in this process. However, one question concerning article use in language remains unanswered or at first sight seems to be at odds with the explanations offered by Caramazza.

Why does lower processing time for adults correlate positively with more omissions in child speech (and special registers)?

According to Caramazza, Romance languages are 'late selection languages' and Dutch and German are 'early selection languages'. It is a well-known fact from language acquisition studies that children acquiring a Romance language omit less articles and for a shorter period of time than children acquiring a Germanic language (Guasti et al., 2008, Lleó and Demuth 1999, Chierchia, Guasti and Gualmini, 1999). Less well known is the fact that, as I will show in the next chapter, we find the same difference in omission pattern in adults' special registers. We find more omissions of articles in Dutch newspaper headlines than in Italian newspaper headlines. We are now faced with an apparent contradiction: In Dutch and German, articles apparently are processed more easily by adults (at least, they are processed faster) than in Romance languages. However, children acquiring a Germanic language have more difficulties in article production than children acquiring a Romance language. And in special contexts where 'speeded language production' seems to be essential, we find more omissions of articles by Dutch than Italian adult speakers. Hence, the lower processing time needed for article production in Germanic languages seems to imply more omissions, both in child and adult speech. We find more omissions in the language with the faster adult processing time. Although Caramazza's model does not aim at explaining acquisition data, the observed apparent contradiction is intriguing, and asks for a further explanation that can at least capture:

- the faster processing time for Dutch adults attested experimentally

selection process, and it is based not on competing genders but on competing phonological context.

- the slower acquisition rate for Dutch children
- the higher article omission rate in Dutch adult's special registers

In the following section and in chapter 4, I will propose how this contradiction can be solved. It is important to emphasize here that Caramazza gives a very detailed description of the processes of determiner *activation*. Determiners are activated on the basis of a determiner frame with language-specific slots that have to be filled in by semantic, grammatical and phonological features. However, in the insight he offers in the actual process of determiner *selection* one important aspect is missing, namely, the role of language-specific properties of the article set itself. Caramazza focuses on the processes taking place during the preparation (or activation) stage of selection and argues that languages differ in this respect. But, as far as the *actual* process of selection of the article from the article set is concerned, hence the process that takes place *after* all the slots in the article frame have been filled, he abstracts away from crosslinguistic differences. His model does not take into account that languages may differ with respect to the question of how difficult it is to select a specific article from the article set. More specifically, what is missing in his studies is the role of the differences in 'accessibility' or 'complexity' of the determiner set in different languages.

Languages may differ in the level of activation necessary for the selection of a determiner. Selecting an article in one language may be a more 'costly' operation, from a processing point of view, than selecting an article in another language. If article selection is a relatively 'easy' process in a language, the activation level necessary for determiner selection will be low. If the activation level is low, an article can be selected, in spite of the fact that the available processing resources of the speaker or the available processing time are limited. If, however, because of a higher complexity level of the article set, the necessary activation level is high the outcome of the process of article selection may be influenced by the amount of processing resources or time the speaker has available.²²

²² Compare the preparation for an exam: the result the student obtains depends on two factors:

1. the student's preparation (his 'activation')
2. the level of difficulty of the exam.

The question of course is whether it is possible to evaluate and compare this alleged level of complexity, and if so how? In chapter 4 I will propose a model that will enable us to solve this question.

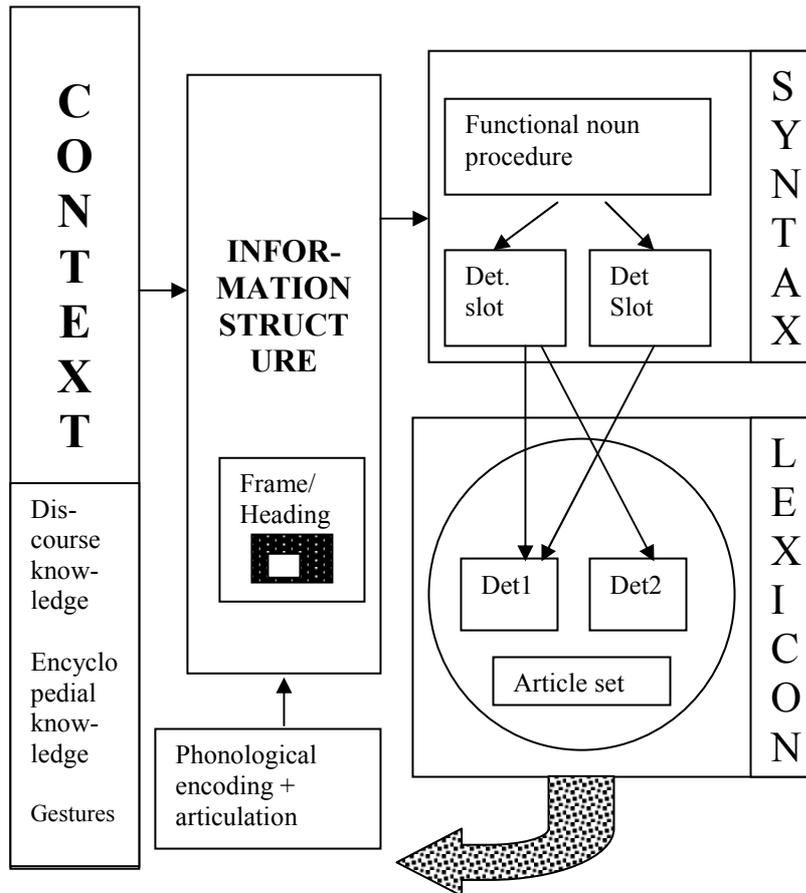
2.7 Conclusion: Model for article selection

2.7.1 Introduction

The model for article selection that I will propose combines the proposals of Levelt's model, Avrutin's model and Caramazza's model. This model describes the process of article selection under normal and specific circumstances, with normal speakers and speakers with limited processing resources. Figure 9 illustrates the architecture of this model for the processing of articles.

If the exam is 'easy' the student can pass, even if his preparation wasn't optimal. But if the level of complexity of the exam is high, the student can only pass the exam with optimal preparation, and under optimal conditions.

Figure 9 Model for article selection



I will first give a description of the components of the model and of the processes that take place in these components. I will then discuss the predictions the model makes for the production of a DP for:

- normal speakers in normal contexts
- speakers with limited processing resources
- normal speakers in special contexts
- differences between Dutch and Italian

2.7.2 The components

I propose that the language processing model (abstracting away from phonological realization) consists of three basic components:

- Context
- Syntax (which has access to the Lexicon)
- Information Structure

Let me first describe the different components:

Context:

The surroundings of the utterance in a broad sense: linguistic and non-linguistic. In Context we find the information-stores Levelt calls ‘discourse knowledge’ and ‘encyclopedial knowledge’. Discourse knowledge is the knowledge of what has been said before, and what is being said during the conversation as well as situational knowledge about the context of the conversation. Encyclopedial knowledge is the structured knowledge of the world the speaker has built up in the course of his lifetime. I suggest that our ‘system of thought’ also forms part of the context, as do the gestures we have at our disposal to support our utterances.

Syntax:

This component has the same function as the component called grammatical encoder in Levelt’s model. It operates on the input (‘preverbal message’ in Levelt’s terminology) it receives and translates the (non-linguistic) information into a linguistic form. In Levelt’s model the output of the grammatical encoder is called the ‘surface structure’ and forms the input for the phonological encoder. Since my study is not concerned with phonological realization, I abstract away from the retrieval of phonological forms.

Information Structure:

I follow Avrutin's proposal about a separate intermediary level between context and syntax, the Information Structure level. This intermediary level transmits the communicative intention to syntax and it subsequently receives the eventually retrieved overt elements from the phonological encoder. As the elements provided by syntax are not always directly interpretable for our cognitive system, at Information Structure level the operations that are necessary to make them interpretable are performed. Let me illustrate this with an example. Suppose Information Structure receives as input from syntax the DP *'the weird scientist'*. Use of the definite article presupposes that the particular scientist has been introduced into the discourse before, or belongs to a general knowledge store or to shared knowledge between interlocutors. A definite DP is referentially deficient. This means that use of a definite DP is only legitimate if it can be connected to another DP available in the context. In technical terms this connection operation is called 'incorporation', see Heims (1982) for details of this process. The processes at the level of Information Structure search the context for the information that is necessary to find the referent of *'the weird scientist'*, and in this way make the DP interpretable for our cognitive system. Therefore the level of Information Structure has to have access to context. Information Structure thus receives information from both syntax and context and is an intermediary level between these two components.

2.7.3 The process of article selection: a general overview

In this section I will give a general description of the process taking place in the different modules. In the next section I will discuss the predictions the model makes for different categories of speakers. At the same time, I will give a more detailed view of the processes, pointing to what is problematic in 'abnormal' discourse circumstances and explaining the reasons for these difficulties.

With Level I assume that the trigger for an utterance is formed by non-linguistic conceptions that contain all the necessary information to convert our meaning into language. Levelt calls these conceptions 'preverbal messages'. However, as we saw in section 2.3 verbal language is not the only method we have at our disposal to communicate our

intention. We can for example use gestures instead of words. In fact it is even possible to communicate our intention by using only gestures. Since the term 'pre-verbal message' anticipates *verbal* realization at a later stage I propose another term for these primary conceptions: '*communicative intention*'. This leaves open the possibility of alternative, non-linguistic ways of realization of the communicative goals.

In my view, the most important reason why we use language (verbal or non-verbal) in communication is the obvious fact that the interlocutors in the discourse cannot read our minds. This means that we need a technique to translate our communicative intentions in such a way that they can be interpreted by our interlocutors. Our communicative intention is conceived in our system of thoughts. This conception requires the speakers' conscious attention. Speakers have to take into consideration not only the meaning they intend to transfer, but also the knowledge interlocutors have about the object or event, about what has been said earlier in the discourse, about the presence or absence of the described object or event in the direct context, the type of discourse situation (formal, informal) and many other factors that play a role in the way in which the communicative intention has to be translated.

The communicative intention is passed from our system of thought, hence from context, through Information Structure, to syntax.. Under normal conditions (normal speaker, normal context) at the syntax level lexical elements are retrieved from the lexicon on the basis of the communicative intention. This means that at the level of syntax the communicative intention is translated into a linguistic form, which in my model, following Avrutin's proposal, will be passed to the Information Structure level, where the interpretability of the retrieved elements will be checked.

Let me start with a general observation on the translation process that takes place at the syntax level. It is a well-known fact that language is a far from perfect method for the expression of our thoughts. The problem with language is that it necessarily has to be far more restricted than our thoughts are. Often it seems to be impossible to fully express our thoughts by language. This is caused by the fact that language compresses our thoughts. An obvious question then is: why is this the case? Why does it appear to be inevitable that language is a restricted, compressed, reflection of our thoughts? It is obvious that this restriction is not caused by properties of language per se. It is generally assumed that all languages of the world offer in principle an unlimited range of

possibilities of expression. The bottleneck is not the expressive force of languages, the bottleneck is formed by the processing resources that are required to translate our thoughts into language and by the time we have at our disposal for the translation process. After all, we do not want to speak too slowly. Because of this restriction in both processing resources and time available for the processes that take place at the level of syntax, the amount of information we can translate into language is limited. This means that the output of syntax is very often a deficient translation of our communicative intention.

My model is concerned with article selection, hence selection of a syntactic form. Following Levelt, I assume that nothing in the speaker's communicative intention will by itself trigger a particular syntactic form. The selection of the article is triggered by properties of the noun, which means that the noun lemma (or the heading 'noun') has to be activated first. Even information about the accessibility status of the referent, which, at a later stage can (eventually) be expressed by an article, forms part of the communicative intention that leads to selection of the noun lemma.

To be fully interpretable for our cognitive system, the (noun) heading needs a frame. I argue that Information Structure will *first* search for a frame that is constructed at the level of syntax. In principle, Information Structure could at this stage also search for a frame at context level, but I will argue here that this is not the case. The most important reason for this claim is the assumption that the procedures at syntax and at Information Structure are automatic processes, taking place without the speaker's conscious attention. In normal conditions the frame is provided by syntax, in fact in most cases the frame is provided by syntax. In my view this can only be explained if we assume that the building procedure moves automatically from the Information Structure level to the level of syntax. Only if syntax cannot provide a frame, for reasons that will be explained later in this chapter, will the frame be searched for at context level.

The processes at Information Structure could, of course, as Avrutin proposes in his model, automatically search both modules, syntax and context, and compare the processing cost of the frame in both modules, since both modules are available. I will suggest, however, that the assumption that the building procedure automatically moves from Information Structure to the syntax level reflects the 'cheaper' option, in terms of processing resources that are required for the construction of the frame. In this case the frame will be produced in only one module.

The assumption that two possible ways for the construction of a frame compete and their processing costs will be compared, implies dual processing: both modules (syntax and context) have to be triggered to actually produce the frame. After all, the processing costs of a frame in a specific module can only be known if the frame is actually being produced by the module. The module that can provide the frame at the lower processing cost will be selected. However, *both* modules would have to, at least, initiate but maybe even complete the frame ‘production’ process before the processing cost can be known. For the speaker, this will lead to processing cost caused by *two* modules. From the perspective of economy a competition process between two modules would therefore not be the most efficient way to construct a frame.

Therefore, Information Structure will *first* search for a frame that is constructed at syntax level. At this level a so-called functional procedure is initiated on the basis of the grammatical properties of the lemma (e.g. noun, verb). Every grammatical type of lemma has its own functional procedures. These functional procedures are stored in our mental lexicon, and trigger the creation of a syntactic structure that is appropriate for the lemma. The resulting syntactic structure contains the information on the basis of which the syntactic elements (like articles, inflections, etc.) can be selected. Hence a noun lemma will start a functional procedure to create an appropriate syntactic environment for a noun. This environment can (depending on the semantic information in the noun lemma) contain a position that has to be filled with an article.

Combining the insight offered by the model of Caramazza with the model of Levelt, I argue that the functional noun procedure at syntax level contains the rules that specify the article frame, defining which (language specific) slots have to be filled in. I follow Caramazza and assume that *all* slots of the frame have to be filled in before the determiner can be selected.

The syntactic building procedure at syntax takes place automatically, non-intentionally. But this does not mean that this building procedure does not put demands on the speaker’s processing resources. I will turn to this observation later, in my discussion of speakers with limited processing resources and normal adults in special contexts.

As a next step in the process this output from syntax is passed to Information Structure (passing through the phonological encoder, after all, the output of syntax consists of *overt* lexical elements). Here the message formed up to this stage will be compared with the

communicative intention conceived by our cognitive system. Since the linguistic translation all by itself is usually not sufficient to fully express our intention, the processes that take place at Information Structure will search in the context for information that can make up for what has been omitted in the first stage. If context can supply this additional information, it will do so. We sometimes find gestures as an overt way of expressing that information has been provided by context. It is definitely not the case that all information in the communicative intention that has not been syntactically translated into language can be provided by context. Context cannot always make up for what has been left out by narrow syntax. To illustrate this with ‘the weird scientist’ example, if context cannot provide a referent for the referentially deficient DP, the utterance will be uninterpretable.

If we compare my model with Avrutin’s, I propose a different role for context: not the role of a competitor, fighting with narrow syntax for the right to encode information, but the role of a complement to narrow syntax. At the Information Structure level, context will contribute to the translation of the communicative intention in an interpretable form, by making up for what not has been encoded by syntax. This is only possible if specific contextual conditions are satisfied. Hence, in the model I propose there is no competition between context and syntax on the level of Information Structure. As I will discuss in more detail in the following sections, in my model a competition process does take place at the level of syntax itself, during the process of selection of elements from the lexicon. The different elements that belong to a particular set, like articles, fight for the right to be selected. This competition process puts a strain on the available processing resources: the stronger the competition between the elements within a set, the more processing resources are necessary to retrieve a particular element from the set. Therefore, in cases of strong competition between elements, speakers with limited processing resources and speakers in conditions in which the available processing time is limited cannot always carry out the selection process. If the selection process cannot be completed, we will find omissions.

2.7.4 The process in detail: The model's prediction for different speakers/conditions

Let us take a closer look at the processing of the article in the Dutch DP '*de vreemde wetenschapper*' and the Italian '*lo strano scienziato*' (*the weird scientist*) in several discourse conditions:

- normal speakers, normal discourse situation
- speakers with limited processing resources
- normal speakers, specific discourse circumstances

2.7.4.1 Normal speakers, normal discourse situation

Context: Both speaker and hearer visited a linguistic conference, where they met a very peculiar scientist. Recently, speaker saw this man at another conference, where hearer was not present. Speaker is now telling about the people he met at this new conference, and then he mentions also 'the weird scientist we met last year'. For hearer the naming of 'the weird scientist' immediately refers to the intended individual, on the basis of shared knowledge.

In the communicative intention the speaker will encode in an abstract way that there is shared knowledge between him and the hearer that has to be taken in consideration in the formulation. To illustrate the importance of the correct encoding of this discourse status, if instead of producing the utterance '*the weird scientist*' the utterance '*a weird scientist*' were produced, the hearer would not be able to find the intended referent, and the utterance would be uninterpretable for him. Further, the communicative intention will contain semantic information on the basis of which the noun '*scientist*' and the adjective '*weird*' can be selected from the mental lexicon. It will then be forwarded to the level of syntax, where, under normal circumstances, it will be translated into a linguistic message. I assume, like Levelt and Avrutin, that this syntactic translation process is a highly automatic and nonintentional process. A speaker will not, for every message, consider which of the various grammatical alternatives would be most effective in reaching the communicative goal. An obvious question then arises: Does this mean that the output is fully predetermined by the input? I will argue that this is not the case. Though it is not possible to influence the internal structure of the automatized process, there are factors within the level of syntax that influence the process and the output. These factors are:

- the processing resources that are required for the realization of the automatized processes. The fact that the processes are automatized does not imply that the processes do not make demands on processing resources. And processing resources are restricted, even in normal adult speakers. As a consequence, the number of informational units that can be processed in a given time span by narrow syntax is restricted.
- the time speakers have at their disposal: this factor follows automatically from the previous one: if the number of informational units that can be processed *in a given time span* is restricted, the total number of informational units that can be processed depends on the available processing time. The more time available, the more units of information can be processed.

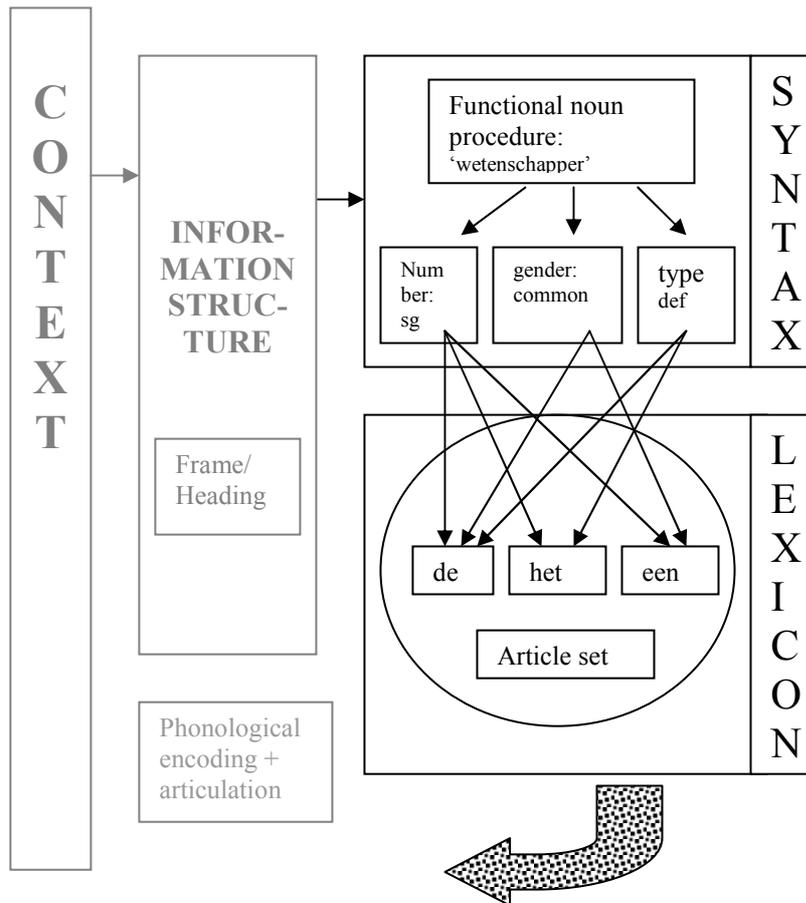
In a normal discourse situation with a normal adult speaker the available processing resources and the available time will suffice to translate the communicative intention of *'the weird scientist'* into a linguistic form. This means that the grammatical encoder will be able to retrieve from the lexicon the lexical and functional elements that constitute the linguistic message. Of course the question that immediately arises then is: Why is it the case that the available processing resources in normal adults in normal discourse circumstances will suffice for the production of a 'grammatically normally formed' utterance? Is this a coincidence, or can it be motivated? I will argue that it is definitely not coincidental. I assume that the capacity to form a grammatically well-formed utterance is the standard developmental measure towards which brain maturation processes of the language learning child are directed. If we talk about 'brain maturation' we mean that our brain capacity develops and is being structured in such a way that it is able to allocate the available processing resources and the available processing time in such a way that the brain regions involved in language processing can cope with the language-specific requirements for the selection of elements from the lexicon. Formulated differently: brain maturation not only means that more brain capacity becomes available, but also that the available capacity is structured in a more adult-like way (Gaillard et al., 2000, Holland et al., 2001). The available capacity will be distributed among the different processors in such a way that the infant will ultimately be able to perform in an adult-like way in a great variety of fields requiring cognitive resources, one of these fields being language.

Hence, in a normal discourse situation, with a normal adult speaker the information provided by syntax to the next level, the level of Information Structure (through the Phonological Encoder), will consist

of lexical elements corresponding in the best possible way with the communicative intention. This does not mean that the levels of Information Structure or context would not be necessary anymore. For instance, even in the simple example given earlier context is indispensable. The best possible translation syntax can provide for our communicative intention is *'the weird scientist'*, which is a referentially deficient definite NP that cannot be interpreted without additional contextual information.

In Figure 10 I illustrate the article selection process as it takes place at syntax level using the Dutch NP *'de vreemde wetenschapper'* (the weird scientist) as an example.

Figure 10 *Model for article selection in Dutch 'de vreemde wetenschapper'*



On the basis of the semantic information in the communicative intention a noun-lemma or heading is selected. The communicative intention contains the information on the accessibility status of the referent. In this case the accessibility index indicates that the referent is accessible in a shared knowledge store. Therefore a *definite* article has to be selected, which means that all definite articles will receive activation by the feature [definite]. In Dutch this means that ‘*de*’ and ‘*het*’ will be activated.

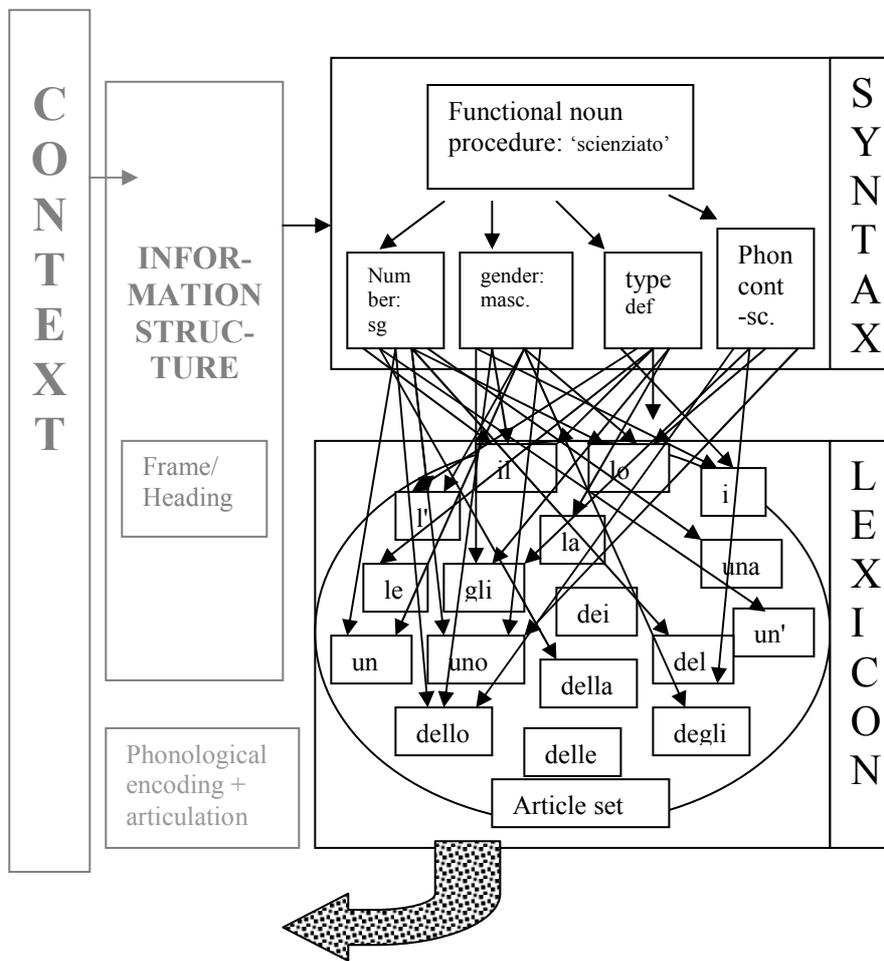
The noun lemma or heading ‘*wetenschapper*’ will activate the gender feature of the determiner: [common]. Hence, all articles corresponding with common gender will be activated. In Dutch this means that both ‘*de*’ and ‘*een*’ will be activated. I assume that the noun ‘*wetenschapper*’ will also activate the number feature of the determiner, ‘singular’, this will mean that all articles, ‘*de*’, ‘*het*’ and ‘*een*’ will be activated.

I will assume, with Caramazza, that the article that receives the strongest activation will be selected. The processor will compare the activation levels of the different articles and in this case ‘*de*’ will win the competition process.

It is important to note at this point that selection of an article not only depends on the level of activation of the individual article, but also on the activation level of the ‘competitors’. I will return to this observation later in this section.

In Figure 11 I illustrate the selection process for the Italian DP. The picture may lead to the intuitive conclusion that the process is more complicated than in Dutch, at least that it looks more complicated. However, this intuition is wrong, as I will show later on. Basically the process is the same as in Dutch. There is one additional slot that has to be filled in, with phonological context information. In addition, the set of articles contains more elements.

Figure 11 *Model for article selection in Italian 'lo strano scienziato'*



Summarizing, in a normal context with a normal speaker the maximally possible amount of information contained in the communicative intention will be encoded by syntax. But, because of its restricted capacity only a limited number of information units can be processed by the syntactic channel. This means that the message contained in the communicative intention will necessarily be ‘compressed’ by the processes in syntax. Therefore, at the level of Information Structure the message produced by syntax will be completed (‘decompressed’) with information coming from the context. If context can provide information in such a way that in combination with the compressed message from syntax the result is an interpretable utterance, expressing the communicative intention in the best possible way, the communication has reached its goal.

2.7.4.2 Speakers with limited processing resources

Let us now take a look at the prediction of the model for speakers with limited processing resources, as for example language acquiring children or aphasic speakers. As is well-known, the speech of these speakers is characterized by frequent omission of functional categories. But they do not always omit them, omission is optional. The model will have to account not only for this optionality, but also for crosslinguistic differences in omission patterns. For example, Dutch children omit the article more often than Italian children.

I do not assume that omissions are caused by the fact that the communicative intentions of agrammatic speakers or children differ from those of adults, certainly not in the case of a simple DP as in our example of *‘the weird scientist’*. Of course, it is reasonable to argue that, because of their limited brain processing resources they cannot conceive highly complicated communicative intentions in an adultlike way. However, we find omissions of functional categories even in the simplest utterances. It could of course be argued that even a simple utterance can require a complicated communicative intention. For instance, for our example with the definite article it could be argued that the correct encoding of the shared knowledge with the hearer might be problematic for people with limited processing resources (see for example Schaeffer and Matthewson, 2005 for child speech). In fact, there is evidence that it is even problematic for normal adult speakers (Keysar et al, 1998, Horton and Keysar, 1996). But then if that were the problem, we would not expect *omission* errors, rather we would expect *substitution* errors (in fact language acquiring children do for quite a long time make

substitution errors in the sense that they use the definite when the indefinite article would be appropriate (Schaeffer et al., 2005), and this can be a consequence of a non-adult like formulation of the communicative intention). Omissions, however, require a different explanation.

As is the case with normal adults and in normal discourse situations, the communicative intention will be passed to the level of syntax. The number of information units that can be translated into a linguistic form is restricted, even in normal adults. But under normal discourse circumstances this restriction is not problematic since the available adult brain capacity is structured in such a way that it can cope with the cost of retrieval of the proper lexical elements. In people with limited processing resources, however, the fine-tuned balance between available and required processing resources that we witness in normal adult speakers is either disrupted (in the case of agrammatic patients) or has not been established yet (in the case of language acquiring children). In these populations there are not enough processing resources available to translate the complete communicative intention into language. The syntactic formulator will therefore compress the communicative intention even more than in normal adult speech. The question arises then, why this should lead to omission of functional categories, and not to omission of lexical categories. Or, in our specific example, why ‘*lo*’ is omitted in Italian, and not ‘*scienziato*’, why is ‘*de*’ omitted in Dutch and not ‘*wetenschapper*’? Following a natural, conscious line of reasoning, this could be related to the fact that ‘*scienziato*’ is more informative than ‘*lo*’. Of course this is true, but this is a conscious line of reasoning. Yet, the processes at the level of syntax are highly automatized, non-intentional, computational processes. How could the processor know that ‘*lo*’ is less informative?

Another potential reason why ‘*lo*’ is omitted and not ‘*scienziato*’ could be that in order to activate the article, gender information has to be available, hence, ‘*scienziato*’ has to be selected first. And then, one might argue, after the processing of ‘*scienziato*’ there are not enough processing resources left for ‘*lo*’. There are at least two reasons why this argumentation can not be correct:

1. There are crosslinguistic differences in omission of articles in child speech and adult’s special registers, Dutch speakers omit more articles than Italian speakers. If articles are omitted because of the fact that after the processing of the noun not enough processing resources are left to process the article, we would not expect these crosslinguistic differences,

or we would need to explain why processing a noun demands less resources in Italian than in Dutch.

2. The experiments of Caramazza discussed earlier have shown that gender information on a noun becomes available and activates its determiner even if the noun itself is not selected (recall that in several of his experiments the gender information of a distractor noun became available and interfered, in the case of conflicting gender, with the selection of the article of the target noun). This shows that the argumentation that the noun has to be *selected* first is not correct. Rather, it has to be *activated*, and when it is activated the gender information contained in the noun lemma will activate the appropriate article form.

Instead I would like to propose a different account. The processor is not sensitive to differences in informative value, defined in an intuitive way as above, but rather to differences in the processing cost necessary to retrieve the elements from the lexicon. In chapter 4 we will see that the processing cost of retrieving a functional element from the lexicon is higher than the cost of retrieving a content word from the lexicon and I will show that these differences in processing cost are related to the level of complexity within the sets from which the elements have to be selected. I will also show that the original intuition was not completely wrong. The elements that intuitively are the most informative elements are usually the ones that can be selected with a low amount of processing cost. However, our intuition does not allow us to make more fine-grained distinctions. Intuitively, an article in Italian will be as uninformative as an article in Dutch. And we will see that in this respect our intuition is indeed wrong.

Summarizing, children and aphasic patients sometimes omit articles because of the fact that these elements require more processing resources to be selected than lexical elements. The syntactic processor only has a limited amount of processing capacity available, and therefore the selection process of the article cannot always be fulfilled. No article will be selected then, and the output of syntax will contain only the noun *'scientist'*.

2.7.4.3 Normal adults in specific contexts (telegram style, colloquial speech, headlines)

In the previous subsection I have argued that the reason for the omission of articles in child speech lies in the restricted processing resources of the non-mature child brains.

It does not seem reasonable, however, to argue that normal adults have less processing resources available when they are writing a telegram, a diary or a headline. Recall that the output of narrow syntax, given a certain input, depends on two factors:

- the available processing resources
- the available time

The second factor, the available time, plays an important role in all of the ‘special discourse’ conditions: after all, in these situations people write or talk in a more ‘speedy’ style, they have less time at their disposal than they have in normal discourse situations. As I will show in chapter 3 newspaper readers do not want to spend much time on reading newspapers, and prefer to read them in a very hurried way: the majority of readers only scan the headlines and read a few sentences of the articles that interest them. Headline writers will take this time restriction into consideration and will therefore aim at producing headlines that convey as *fast* as possible as *much information* as possible.

We can now repeat the same question we asked when we were looking at people with limited processing resources:

- why do we find omission of functional categories?
- why is this omission optional?
- why do we find differences between Dutch and Italian?

The reason why we find omissions in these special contexts is strongly related to the reason why we find omissions in speakers with limited processing resources: the number of informational units that can be processed at syntax level is limited. It is restricted by:

- the number of informational units that, given the available processing resources of the speaker, can be processed within a certain unit of time
- the amount of time speaker has at his disposal (and, in the case of headlines, the speaker/writer assumes the reader has at his disposal)

In people with restricted processing resources the available resources were the bottleneck, while in normal adults in special contexts the available time is the bottleneck. The outcome is the same. The syntactic processor has a limited amount of processing capacity available, and can

process only a restricted number of information units per unit of time. Therefore, if the available time is restricted, the competition process necessary to select the article from the set cannot always be completed. No article will be selected then, and the output of syntax will only contain the noun *'scientist'*. As a consequence, the output of syntax to information structure will even be more strongly 'compressed' than in normal situations.

2.7.5 Why more omissions in Dutch than in Italian?

In this section I will describe how the model I propose for selection of the article can predict crosslinguistic differences. Central to my argumentation is the claim that selecting an article in Italian, despite the apparent complexity of the Italian system, is a less demanding operation than in Dutch. If we look again at Figures 10 and 11 this may seem a very counterintuitive claim, but in chapter 4 I will show with an information-theoretical model that selecting an article in Italian is in fact easier than in Dutch.

In order to be selected, an element has to reach its 'threshold of activation', the level of activation necessary for actual selection. Each element has its own base level of activation, a level determined by frequency. The higher the frequency of an element, the higher the base level of activation is. The reason for this is the fact that higher frequency implies a stronger memory trace and easier retrieval of the element from the lexicon. The amount of information necessary to retrieve an element from the lexicon will be lower if the frequency of the element is higher. If the frequency is higher, the element will be closer to its activation threshold. Articles are among the most frequently used elements in both Dutch and Italian, and therefore this argumentation would predict less omissions of articles than of lexical categories. However, we find the opposite. How can we account for this? I suggest that the reason for this can be found in the fact that the reaction time for selecting an element out of the lexicon depends on an additional factor, besides frequency (and base level of activation). This factor is related to the question how well distinguishable the element is within the set, or stated differently, how strong the competition effect is among the different elements in the set. The more the elements look alike, the higher the threshold of activation that has to be reached for activation of the single element, because the element has to 'fight' to beat its competitors. This

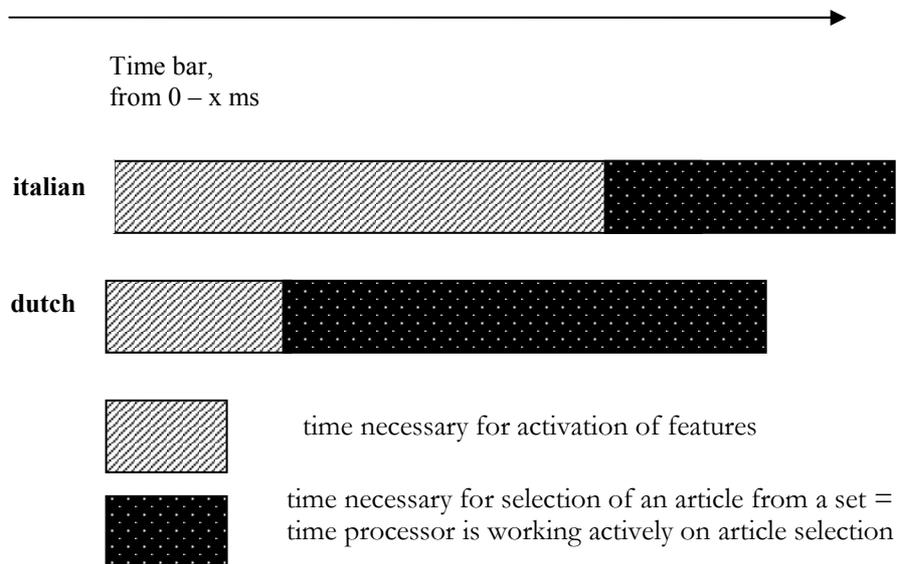
competition process demands processing resources, which is reflected in a longer reaction time (Moscoso del Prado et al., 2004; Kostic, 2004)

My proposal for the account of crosslinguistic differences is based on the claim that the competition between the different articles in the set is stronger in Dutch than in Italian. In chapter 4 I will show that it is possible to measure the competition effect within a set of lexical elements. I will introduce a measure that can be used to express the conspicuousness of the elements in the set. And using this measure, called Relative Entropy, I will show that Italian articles are more ‘conspicuous’ than Dutch articles. This means that Italian articles can be selected more easily among their competitors, and therefore require less processing resources than the Dutch articles.

Recall (from section 2.6) that Caramazza found that Italian is a ‘late selection language’ and that articles in Italian take longer to be processed. How is that possible, if we make the claim that it costs less processing resources (and even less processing time) to select an article in Italian than in Dutch? Isn’t this a contradiction? I will show that it is not. I will show that what is important is the distinction between ‘processing’ time and ‘selection’ time. If we look at the models for Italian and Dutch, we see that in Italian an extra slot of the article frame has to be filled in. As Caramazza argued, this slot depends on the phonological realization, hence contains information that becomes available later. Therefore the final selection process of an article in Italian necessarily has to start later than in Dutch. Crucially we have to distinguish here between two ‘time points’: the point marking the start of the *activation* process and the point marking the start of the actual *selection* process. Figure 12 illustrates the time course of the article production processes in Dutch and Italian, divided into these two stages. The figure shows that the actual selection process starts later in Italian. On the other hand, this actual selection process costs less time in Italian. I assume that activation of most of the features (gender, number and the phonological context), necessary for the selection of the article, is a by-product of other processes taking place, such as the selection of the noun. As is the case with by-products (spin-offs) in a production company, up to the stage where the products are processed individually, a by-product has no specific cost related to it. They come ‘for free’ because of the production of another product. In the same way, grammatical features for article selection originate from the selection of another element, a noun, for example. The selection of the noun itself does demand processing resources, and, hence there is specific processing cost related to this selection. But once the noun is

selected, the grammatical features necessary to fill the article slots are available. From the viewpoint of the article selection process, they do not require additional processing resources. They even become available if the article is not selected (Costa, Gallés, Miozzo and Caramazza, 1999; Miozzo and Caramazza, 1999; Xavier Alario and Caramazza, 2002; Schiller and Caramazza, 2003). This means that the process represented by the lefthand block of the time bar does not make demands on available processing resources, it only takes time until the necessary information becomes available. This process takes longer in Italian than in Dutch. The righthand block of the time bar represents the actual selection process. Here the processor is actively working on the selection of the article, hence this block does represent use of available processing resources by the processor. If selection is more difficult, for example because the elements look more ‘alike’, it will take longer and cost more processing resources than in the case where the elements are more conspicuous and can easily be distinguished. This second part of the process takes longer in Dutch than in Italian.

Figure 12 Schematic representation of the time course of the process of activation of the necessary features for article selection, and the actual selection process, in Dutch and Italian.



2.7.6 Concluding remark: separate storage of articles

The model rests on an important assumption regarding the storage of articles, namely that articles form a separate store in the mental lexicon. More generally, the model assumes that each type of functional category (articles, demonstratives, pronouns, inflections, prepositions, etc) forms a separate set within the mental lexicon, so there is an article set, an inflection set, etc. The reason for this assumption lies in the specific procedures that are necessary to generate the functional categories at syntax level. Let me explain this claim. In my model the heading triggers the construction of an article frame. This frame will be built (if processing resources and time are sufficient) at syntax level. As I have argued in section 2.7, I follow Levelt's proposal and assume that the heading (lemma) calls upon a functional procedure. These functional procedures are stored in our mental lexicon. The functional article procedure forms an article frame, with a number of (language specific) slots that have to be filled in. I propose that for this procedure only articles will be activated in the lexicon and this implies that articles have to be stored separately. I argue that this is because of the fact that they are attached to the specific functional procedure that is needed to activate them. Thus, the functional procedure not only contains rules, but also the elements that are required to implement these rules. The specific functional procedure is called upon by the information contained in the communicative intention, and each functional procedure is linked to its own specific elements. Thus the functional procedure for article selection contains the articles and the rules for their implementation (these rules contain information about the construction of the frames and language specific slots), the functional procedure for demonstratives contains demonstratives and the rules for their implementation, the functional procedure for inflection contains inflections and the rules for their implementation, etc.

This of course implies that retrieval of functional elements differs extensively from retrieval of content words, and that is exactly what has been found in all psycholinguistic studies on the processing of open and closed class words (Bradley, 1978; Garrett, 1982; Bock, 1989, among others).

2.7.7 Summary

The article selection process described here shows that in normal speakers, under normal discourse conditions the necessary elements of the utterance (both frame and heading) will be provided by syntax. However, since the operations taking place at syntax level, in spite of being automatic processes, put a strain on the available processing resources the necessary elements cannot always be provided by syntax. In the case of speakers with limited processing resources or normal speakers in specific contexts, like colloquial speech or headlines, the available processing resources or processing time are not always sufficient to select the elements from the lexicon. Consequently, the output of syntax does not always contain the elements that are necessary for the completion of the File Card at the next level, Information Structure. Whether or not the relevant elements will be provided by syntax is determined by the following factors:

- *processing resources available to speaker*: This explains the differences in the output of syntax in the case of speakers with limited processing resources when compared to normal speakers
- *processing time available to speaker*: This explains the differences in the output of syntax in the case of normal speakers in special contexts when compared to normal speakers in normal contexts
- *processing resources necessary to retrieve the element from the lexicon*: This explains the crosslinguistic differences between for example Dutch and Italian. Retrieving an article from the Italian article set demands less of the processing resources of the speaker than retrieving an element from the Dutch article set. Hence we find more omissions in Dutch.

Chapter 3 Article omission in Headlines and Child Speech

3.1 Introduction

Omission of articles is usually assumed to be a privilege of children or more generally, of speakers with limited processing resources.²³ However, this assumption is not correct. In this chapter I will present data on omission of articles by a category of speakers in which these omissions may be less expected, or may even come as a surprise, namely normal adults. I will show that normal adult speakers optionally omit articles in so-called special registers, like diary style, telegram style and in newspaper headlines.

The fact that omission of articles appears not to be restricted to categories of speakers with limited processing resources but can also be observed in people with normal speech processing capacity is problematic for many accounts of article omission that have been proposed so far, at least if we want to be able to account for these omissions within one and the same model. It is problematic for knowledge-based accounts that claim that omission of articles is caused by lack of knowledge of the correct use of articles. After all, it is not plausible to argue that adults have suddenly ‘lost’ the knowledge of the use of articles in their language when they use special registers. At first sight omission of articles by adults also seems to be problematic for accounts that claim that omission is caused by lack of processing resources. Why should normal adult speakers have less processing resources available when they use special registers?

Of course one could argue that omission of articles by adults in special registers is completely unrelated to omission of articles by speakers with limited processing resources, and consequently, that there is no need for a joint account. However, the data I will present will show that there are intriguing similarities between article omission by adults in special registers and article omission by speakers with limited processing resources, like children. We find the same crosslinguistic differences:

²³ Another category of speakers with limited processing resources in which omission of articles can be observed are agrammatic speakers (Ruigendijk, 2002; de Roo, 1999)

more omissions in Dutch than in Italian. We find the same positional effects: we find more omissions in sentence-initial position than in sentence-internal position, in adult speech as well as child speech. And in both categories of speakers we find the same effect of finiteness: more omissions in sentences with a non-finite verb than in sentences with a finite verb.

If we consider omission by adults and omission by children to be two completely unrelated phenomena, why should we find these similarities? I will argue that the findings on article omission in these two different categories of speakers challenge us to develop one unique model that can capture the reasons for omissions of articles we find in people with limited processing resources as well as the reasons for omissions we find in people with normal processing resources in special contexts.

3.2 Article omission in headlines

3.2.1 Introduction

In this section and the following one I will present data on article omission by normal adult speakers in special registers, in particular in newspaper headlines. I will present data coming from a database I have constructed on article use in Dutch and Italian newspaper headlines. Further I will present the results of experiments I have conducted with Dutch and Italian newspaper readers on the acceptability of headlines with and without articles. These data will show that article omission can be a ‘normal’ pattern in special registers, a finding that challenges the traditional assumption that omission of articles is always ungrammatical.

In chapter 2, section 2.4 we already saw examples of special registers used in colloquial speech. Example 1 repeats one of the examples that were given as an illustration of this style:

- (1) Q: Hebben jullie dat parket zelf gelegd?
Did you place this floor all by yourselves?
A: Ja, gigantisch karwei!
Yes, enormous job!

Colloquial speech is only one of the possible special registers in which adult article omission can be observed. Another example is the so-called 'diary style':

- (2) Had to stop, wet to skin (Haegeman 1990, from V. Woolf, Diary)
played grammophone.....so to Tower (Haegeman 1990, from V. Woolf Diary)

My study focuses on adult article omission in another type of special register: newspaper headlines. In headlines of newspapers articles are frequently omitted. Let me illustrate this with some examples of Dutch and Italian newspaper headlines:²⁴

- (3) NEDERLANDS DRUGSBELEID ONTZET FRANSE
REGERING
Dutch drug-policy horrifies French government
KRAB VEROVERT NOORDZEE
Crab conquers North Sea
- (4) LEGGE GASPARI, PERA MEDIA
Gaspari law, Pera mediates
COMANDANTE ARRESTATO PER SPIONAGGIO
Commander arrested for espionage

These examples already suggest that presence of articles is not necessarily something required by the rules of grammar. After all, if this were so it would be somewhat of a mystery why these rules can be violated in special circumstances. The challenge then is to explain what allows the omission in special registers, e.g. headlines.

3.2.2 Previous studies on headlines

The first linguistic study on headlines was conducted by Straumann in 1935. This was a descriptive study on the grammar of headlines in English newspapers. A fairly large number of studies have been conducted in which newspaper texts (including headlines) are studied

²⁴ To avoid misunderstandings in this study the headlines will be typed in capital letters.

from a socio-linguistic perspective (see for example for English: Arnold, 1969; Bell, 1991, for Italian: Dardano, 1981; Magni, 1992) or from a text-linguistic perspective (Van Dijk, 1988; Dor, 2003). Studies in which headlines are studied from a perspective different from a socio-linguistic one are scarce. Most of these studies are descriptive linguistic studies and concentrate on headlines in English (Mardh, 1980; Simon-Vandenberg, 1981).

To my knowledge, the first *crosslinguistic* study on headlines was conducted by Nortier (1995). In her study Nortier describes a striking discrepancy between Moroccan Arabic / Dutch code switching and Moroccan Arabic / French code switching. In Moroccan Arabic / Dutch code switching no definite Dutch articles are realized, while in Moroccan Arabic / French code switching a construction with a definite French article is very frequent. Looking for explanations for these differences Nortier examined omission of articles by *monolingual* speakers of both languages in specific circumstances. For this reason she investigated newspaper headlines in one French and one Dutch newspaper, to find out whether a similar difference could be observed in these specific contexts. The results confirmed her expectations. She found far more omissions of articles in Dutch headlines than in French headlines. In Dutch there was a striking difference between the number of definite articles used in headlines on the one hand and in 'normal' contexts on the other. In French the difference was almost negligible.

According to Nortier (1995:91), 'the most important and essential reason for code switching is the speakers' wish to express themselves as appropriately and economically as possible....The wish to communicate economically will lead the speaker to code-switch as economically as possible'. Nortier thus claims that the observed crosslinguistic differences in code-switching patterns, leading to more omissions of articles in Dutch than in French, are based on economy considerations. As we will see, this finding is compatible with my account of the crosslinguistic differences between article omission in Dutch and Italian. Our views differ on what exactly causes the differences in processing resources necessary to produce an article. Nortier suggests that the difference is caused by the clitical nature of definite articles in French, contrary to Dutch, which facilitates the realization of French articles. My account focuses on the processing cost of selecting an article from the article set.

To my knowledge the first work in the field of generative linguistics on newspaper headlines was the study by Stowell (1999). He discussed

several characteristic features of English newspaper headlines, like omission of determiners, omission of the auxiliary verb 'be' and use of present tense to report past events. Stowell proposed that the usage of these features is regulated by formal syntactic rules that resemble those of normal grammars. In sections 3.3.3 and 3.4.3.3 of this chapter, I will discuss Stowell's proposal in more detail.

The current study is, to my knowledge, the first linguistic study that focuses on omission of articles in headlines, from a crosslinguistic perspective as well as drawing a parallel with omission of articles in language acquisition. Since use of headlines in a comparative study with child speech is a fairly new topic in language acquisition studies, I believe it is useful to start with a brief introduction on headlines and news texts in general.

3.2.3 Headlines: the language and the functions

We roughly seem to know what headlines and news texts look like and what their characteristic features are. However, they have interesting properties that, if we only look on a superficial level, may remain concealed. I therefore believe that it is useful to provide some background information on headlines, and news texts in general. I will show that headline writers actually attempt to influence the way in which readers process the headline. More particularly, headline writers attempt to influence the processing *effort* of the readers. I will argue that the specific characteristics usually attributed to headlines (such as omissions of functional categories) play a very important role in achieving this goal.

The first question I would like to address is whether the language of headlines can be called 'language'. The language of headlines has often been considered an object unworthy of linguistic investigation, as it deviates too much from normal language. Let me start this section with an interesting anecdote on this topic. Sapir (1921) made the claim that 'headlines are language only in a derived sense'. He claimed that a sentence such as 'The mayor of New York is going to deliver a speech of welcome in French' could only be reduced by eliminating the 'contributory ideas' of 'of New York', 'of welcome', 'in French' (Sapir 1921: 37). But further than this, Sapir claims, we cannot go. Such a shortened form as 'Mayor is going to deliver' cannot be uttered, 'except, possibly, in a newspaper headline. Such headlines however are language

only in a derived sense'. In a diachronic study on the development of the headlines in *The London Times*, SimonVanden Bergen (1981) argued that this view is not correct. She correctly observed that Sapir was undoubtedly right in claiming that a sentence like 'Mayor is going to deliver' cannot be said, but that he was completely wrong in claiming that it could form a headline. The truncated form of this utterance is not at all the type of structure that can be found in English headlines, and this shows, according to SimonVandenBergen, that certain 'rules of cutting' were broken there. This indicates that headlines 'apparently have a grammar of their own, in the sense that there are certain rules which are obligatory, and others that are optional. The former must be strictly adhered to if the headline sentence is to be grammatical' (SimonVandenBergen 1981: 9). This made her conclude that Sapir's comment showed that, in fact, exactly the contrary of this claim was true. It showed that a study of the language of headlines was badly needed.

I will follow SimonVandenBergens proposal that headlines are a *functional* variety of language, a variety of language with specific linguistic features that can be attributed to the special function headlines have to fulfill. Examples of the specific linguistic features characterizing headlines are:

- frequent omission of functional categories, like articles, auxiliaries, copular verbs:

- (5) SMOKING BAN FORCED ON ITALY'S CAFES
 ENGLISHMAN'S CASTLE AT RISK
 MAN HACKED TO DEATH IN BELSIZE PARK

- use of present tense to denote past or future events:

- (6) MINISTER OF DEFENCE WHITE DIES AT 43
 HOLIDAY DREAM TURNS TO SCENE OF HORROR
 MINISTERS MEET AGAIN NEXT WEEK ON AGRICULTURE

- frequent use of ‘nominal constructions’, in which no verb is present:

- (7) MOBILE PHONE TUMOUR RISK
FESTIVE CHEER FOR MOURINHO
SILVER MEDAL PAIR ON VIEW

The next two questions that arise are:

- what are the specific functions headlines have to fulfill?
- why do these functions call for these specific characteristics?

Let us first examine the functions headlines have to fulfill and take a look at some of the proposals that have been made in the literature. From a descriptive viewpoint, Simon Vandenberg (1981:52) proposes that ‘the core function of the headline is best formulated as: provide information about the contents of the article, announce the topic which is discussed, further detailed or commented upon in the article. Of primary importance for the form is the limited space available. Further requirements are that the headlines should be attractive enough to draw the reader’s attention to the crucial point of the news and that in those cases where the reader does not go beyond reading the headlines of his paper the latter should not be misleading but still give accurate information’.

Dor (2003), an Israeli news-editor and linguist, has conducted a very interesting empirical study on the news-desk of one daily newspaper. He followed the decision making process leading to the choice of headlines for a large number of news items. The study focuses on the communicative function of headlines and is based on Sperber and Wilson’s (1986) Relevance Theory, a theory of cost-effectiveness. Sperber and Wilson claim that human cognitive processes are attuned to achieving the greatest possible cognitive effect for the smallest processing effort. Dor proposes that the function of a headline is ‘optimization of the relevance of the story’. Let us look at his proposal in more detail. Dor claims, following Sperber and Wilson, that the relevance (R) of a story is a function of the amount of contextual effects the reader can deduce from the story, hence the amount of information

the story contains for the reader (C) and the effort the reader has to invest in reading the story (E). The relevance of the story can thus be expressed in a formula: $R = C / E$. If we compare the headline with the complete article, then, of course, the information contained in the headline (measured in the number of 'contextual effects' the reader can deduce from the headline) is smaller. A short headline cannot possibly contain the same number of contextual effects as the complete article. However, the effort the reader has to invest in interpreting the headline is also smaller. After all, the headline is much shorter than the complete article, so reading a headline costs less processing effort than reading the complete article. Now, Dor suggests that a headline saves much more on the processing effort (E) than it loses on the contextual effects (C), and therefore, multiplies the relevance of the story (R).²⁵ He claims that this optimization of the relevance of the story contained in the headline is exactly the function of a headline. Stated differently, headlines aim at providing a maximum amount of information per unit of processing effort. I will return to this observation later in my study.

Other researchers have come up with similar, somewhat descriptive claims. Kronrod et al. (2001:696), for example, suggest that 'the purpose of the newspaper is to convey as much information as possible and *as fast as possible*. For the headline there is more information than space. This space restriction, as well as the wish to arouse curiosity, push for brief and vague expressions'. Bell (1991:189) says that headlines are a 'part of news rhetoric whose function is to attract the reader'. Van Dijk (1988) focuses especially on the fact that, because of the restricted time newspaper readers want to spend on the reading, the first part of the article, including the headline, has to be constructed in such a way that it provides the most important information of the article.

In summary, for the purposes of my study the most relevant functions of headlines are:

- optimize the ratio between informational effect and processing time; convey as much information as possible *as fast as possible*.
- optimize the ratio between informational effect and processing effort; convey as much information as possible *with the least possible processing effort*.

²⁵ Dor does not answer the question how the exact values of C and E can be found. He works with arbitrary assumptions and estimates about the relation between the values of C and E in different contexts (e.g. in headlines and in a complete newspaper article).

In conclusion, headlines have to provide the reader with *the best (informational) value for (cognitive) money possible*.

The value of a headline should not be underestimated. Headlines are important constituents of a newspaper, we may even claim that they are the most important elements of a newspaper. Several studies have shown that the reading time of newspaper readers is highly constrained. Readers do not want to spend much time on reading the news. They read in a 'speeded style' and many articles are read only partially (Van Dijk 1988 and references therein). Dor (2003:718) says: 'Most readers spend most of their reading time scanning the headlines, without reading the stories'. (Nir 1993: 24) suggests that 'for the modern newspaper reader, reading the headline of a news item replaces the reading of the whole story'. That is why a news item is constructed in such a way that even partial reading of only the first part of the text, or even only the headline, provides the most important information of the discourse. Studies on recall of the contents of newspaper articles have shown that in general recall of what readers have read is very poor. A large number of studies have been conducted on recall of news by newspaper readers, and all studies agree that readers who only skim the pages of the newspaper by scanning the headlines recall just as much of what they have read as readers who have read the complete stories (Van Dijk, 1988; Dor, 2003; among others).

Since the news contained in headlines is recalled very well by newspaper readers, we must draw the conclusion that headlines are very successful in performing their function, but also that apparently the characteristic features that are used to construct them have the intended effect. But that is no more than a mere description of the facts. It still leaves us with an intriguing question: why do these peculiar headline-features have the intended effect? Why does, for example, omission of functional categories lead to the effect of conveying as much information as fast and as cheap as possible? And, if these headline-features make a language both communicatively effective as well as efficient, why then don't we always talk in 'headlines'?

3.2.4 False beliefs about omissions in headlines

Let us first take a look at some interesting false beliefs about omissions in headlines. It is often suggested that:

- omissions are caused by space restrictions: less informative, redundant elements are omitted because of space limitations.
- headline writers work on the basis of a stylistic 'handbook', which is made up by the editors of the newspaper, and which prescribes what the headlines in their newspaper have to look like with respect to aspects like omission of functional categories, etc.

I will argue that these generally believed assumptions about headlines are at best simplified accounts without explanatory force, but in some cases they are even wrong.

'Omissions are caused by space restrictions; less informative, redundant elements are omitted because of space restrictions'

In many sociolinguistic and text linguistic studies it is assumed that omission of articles is a consequence of the fact that headline writers have to deal with space restrictions, and therefore omit the less informative, redundant elements. In fact, journalistic handbooks recommend omission of articles. Arnold (1969:93), for example, suggests: "Headlines eliminates articles. It says: DELEGATION GOES TO WHITE HOUSE. Adding an article is not a hanging offence, but it jars the reader just as it would if you told him 'I am going to the home' instead of 'I am going home'.

Stowell (1999), however, showed that this recommendation given in journalistic handbooks to always omit articles is too rigorous. Stowell observed that omissions in headlines are linguistically constrained, i.e. not every type of omission is permitted. For example, he observed that omission of an article before the direct object is impossible if the article has not been omitted before the subject.

- (8) CABBAGETOWN HOUSEWIFE FINDS RARE GOLD
COIN
* A CABBAGETOWN HOUSEWIFE FINDS RARE GOLD
COIN

Let me give another clear example of the fact that omissions are linguistically constrained. While constructing the database of headlines (see section 3.3), I observed that, as expected, we do find frequent omissions of auxiliary verbs:

- (9) VREDESPLAN VOOR HAITI VERWORPEN
Peace plan for Haiti rejected

NOS INTERNETSITE TOTAAL VERNIEUWD
Website NOS completely renewed

However, omission of auxiliaries does not occur freely. In the examples above an inflected form of the auxiliary ‘*zijn*’ (be) was omitted. In fact, all omissions of perfect tense auxiliaries in Dutch headlines are omissions of ‘*zijn*’ (be) or ‘*worden*’ (auxiliary used for passive constructions in Dutch). However, Dutch does have another auxiliary for perfect tense, ‘*hebben*’ (have). Some verbs require the use of ‘*hebben*’ to form the perfect tense, other verbs require the use of ‘*zijn*’. Only ‘*zijn*’ can be omitted, omission of ‘*hebben*’ is impossible in a headline:²⁶

²⁶ Even though omission of auxiliaries is not the subject of my study, for the interested reader I want to propose an account for the differences in omission pattern of ‘*hebben*’ and ‘*zijn*’ (have/be) in past participle constructions in headlines. This phenomenon is reminiscent of the unaccusative/unergative distinction as it has been developed by Burzio (1981) and Perlmutter (1978). Elaborating these proposals Haider (1984) and Hoekstra (1986) propose that past participle morphology blocks assignment of argumenthood to the external argument. Therefore a lexical verb with past participle morphology will assign argumenthood only to the verb’s internal argument (like an unaccusative verb). In order to repair the verb’s capacity to assign argumenthood to an external argument in constructions with a past participle the auxiliary is needed. This means that constructions of past participles without an auxiliary verb are only possible in the case of lexical verbs that only have an internal argument in their thematic grid (unaccusatives, unergatives used in a telic meaning, and verbs in passive constructions). Since these are exactly the verbs that take ‘*zijn*’ as auxiliary verb, only ‘*zijn*’ can be omitted.

This can be illustrated with the following headline-constructions. The Dutch verb ‘*springen*’ (jump), is an intransitive verb, and can be used with telic or non-telic meaning. If the verb is used with a telic meaning the auxiliary ‘*zijn*’ (be) is required to form a perfect tense, and can be omitted in a headline, see example (i)

- (i) MINISTER PRESIDENT IN SLOOT GESPRONGEN
Prime Minister jumped in ditch

However, if the verb is used with a non-telic, durative meaning the auxiliary ‘*hebben*’ (have) is required, and cannot be omitted from a headline, see example (ii)

- (ii) *MINISTER PRESIDENT GESPRONGEN IN SLOOT

In Dutch constructions of this type, if the PP precedes the verb, as in example (i) the verb can be used with both telic and non-telic meaning. ‘MP *heeft* in sloot gesprongen’

- (10) * REGERING VREDESPLAN VOOR HAITI VERWORPEN
 (REGERING HEEFT VREDESPLAN VOOR HAITI
 VERWORPEN)
 Government (has) rejected peace plan for Haiti²⁷

Now if omissions were only based on the low informativeness and redundancy of the element, we would not expect a difference between the auxiliaries *'hebben'* and *'zijn'*. The fact that there is one must be due to a linguistic constraint.²⁸

Another reason why space restrictions are no adequate explanation for omissions of functional categories in headlines concerns the fact that there are crosslinguistic differences in these omissions. I will show later that in Italian omission of articles seems to be far more restricted than in Dutch. In an experiment we conducted with Dutch and Italian newspaper readers (see section 3.4) we compared their judgements on headlines in which articles were either used or omitted.

- (11) (DE) REGERING VALT BALKENENDE AAN.

gives a non-telic, durative meaning, 'MP *is* in sloot gesprongen' gives a telic meaning. But, if the auxiliary is omitted only one interpretation of the verb is possible: the telic one. Hence, only 'is' can be omitted, 'heeft' cannot be omitted. If, as is the case in example (ii) the verb precedes the PP the verb can only be used with a non-telic meaning, requiring the auxiliary 'hebben' which cannot be omitted.

Summarizing, when *'zijn'* is omitted the lexical verb still has its capacity to assign argumenthood to its only argument, the internal argument. But when *'hebben'* is omitted the lexical verb cannot assign argumenthood to its external argument, and the sentence becomes ungrammatical.

²⁷ In English a headline like 'GOVERNMENT REJECTED PEACE PLAN FOR HAITI' is not problematic, but that is because of the fact that English simple past tense of regular verbs uses the same morphological form as the past participle in perfect tense. Therefore the English sentence may be interpreted as simple past tense. This is not possible in Dutch and Italian.

With irregular verbs, however, the effect becomes visible in English too:
 * GOVERNMENT SEEN PEACEPLAN

²⁸ Avrutin (1999) provides examples of the constraints on the omission of Tense in English Headlines. Omission of Tense is possible in matrix clauses

(i) UNIONS TO GO ON STRIKE

but ruled out in embedded contexts

(ii) *WORKERS HOPE THAT UNIONS TO GO ON STRIKE

(IL) GOVERNO ATTACCA PRODI
(The) government attacks Balkenende/Prodi

Interestingly we found that Dutch headline readers have a strong preference for headlines in which the article has been omitted, while Italian headline readers prefer the version with the article present. Space is just as restricted in Italian newspapers as it is in Dutch newspapers, therefore space restrictions cannot explain these crosslinguistic differences, pointing to the need for an additional explanation.²⁹

Another ‘false’ belief about headlines is that
‘Headline writers work on the basis of a stylistic ‘handbook’, which gives guidelines on what the headlines in the newspaper have to look like. In other words: they omit articles because they are told to do so.’

The central question is: how editors decide on what to omit in headlines. Do they work following guidelines? Illustrative in this respect is a study of Bell (1991), who, like Dor, works both as a newspaper editor and as a linguist. About the editing operations (including deletions) that are performed, he says: “The editing operations are rarely conscious, and editors are surprisingly unaware of what they are doing with language. Even for myself, as a journalist editing news copy and as a linguist analysing editing processes within the same day, I am largely unaware of the precise operations performed as I edit (and I imagine that becoming too aware could lead to paralysis!). In some cases, how we would describe an operation for linguistic purposes is demonstrably different from how the editor thinks of it.” (Bell, 1991:121)

This same finding about editors working largely ‘unconsciously’ or ‘intuitively’ was described by Dor (2003:707). ‘In general, news editors do not work with a very explicit definition of what headlines are. When asked to provide an explicit definition of what a headline is, senior newspaper editors usually give an answer of the type: ‘I don’t know what headlines are, but I can tell a good one when I see it’. This shows, according to Dor, that professional knowledge is practical, not theoretical. Editors do not work on a theoretical basis that provides them with guidelines for a good headline. Moreover newspaper editors

²⁹ In fact, as suggested by Guasti (p.c.), given that Dutch has shorter words than Italian, more space is available in Dutch headlines, and this should promote the use of articles, contrary to facts.

appear to have a very high rate of agreement on the preferred headline. This means, Dor argues, that experienced news editors know a great deal more about the functional properties of headlines than they ever explicate. In this sense, ‘headline production is more similar to an artistic activity than, say, to the practice of an exact science’. (Dor 2003:707)

Hence, while they are editing a headline, editors are largely unconscious of the details of what they are doing. This means that even if there is a ‘stylistic guideline’ for the newspaper editors, they do not use it. Just like normal adult speakers when they speak their native language, editors use their unconscious knowledge about their language.

3.2.5 Why do headline writers omit functional categories?

Why do the specific functions of headlines call for the specific characteristics of headlines, and how do these characteristics contribute to the fulfillment of these functions?

Dor’s argument that the knowledge about what makes a good headline is not theoretical, but practical, and based on the editor’s professional experience may not be completely correct. I will show (see section 3.4). that not only newspaper *editors*, but also newspaper *readers* have strong judgements on what makes a headline a good headline, without having any professional experience as headline-writers (or -readers, for that matter). And we will see that newspaper readers too show an extremely high rate of agreement on the preferred headline. This shows that both readers and editors have unconscious, intuitive knowledge about headlines. Dor is right in claiming that this knowledge cannot be theoretical, but he is wrong in suggesting that it is based on professional experience only.³⁰ There has to be an additional factor that explains what

³⁰ Professional experience is necessary, since writing a good headline requires far more than knowing whether or not a functional category, like an article for example, can be omitted (see for the details Dor, 2003). It requires a thorough evaluation process of the information the headline has to contain and the way this information has to be encoded to maximize the relevance of the headline and minimize the processing effort. This evaluation may very well be based on professional experience. Dor describes, using several ‘real’ examples, how the negotiations about the best version of a suggested headline take place in the news desk of his newspaper. For example he argues that if we compare ‘THE FIRST CASINO IN JERICHO WILL BE OPERATIONAL IN FEBRUARY’ with ‘THE FIRST CASINO IN JERICHO WILL BE

makes a headline a good headline. What can this reason be? Since for most readers it is important that headlines provide news and can be scanned quickly (Van Dijk, 1998; Dor, 2003), it seems reasonable to argue that readers will prefer the headline which provides them the highest amount of information for the least processing effort and time. I propose that it is legitimate to make the claim that the reader's judgement on headlines will be related to the amount of processing effort necessary to process a headline. Consequently, the editor's judgement will also be related to the amount of processing effort necessary for the reader to process a headline. After all, headlines are designed for the 'audience', the newspaper readers. Therefore, the editor's judgement will be based on his intuitive judgements about the reader's judgements.

Both Bell and Dor claim that the editing process of a headline is largely unconscious. This unconsciousness is not surprising, since the production of a headline is, just like the production of an utterance in normal speech, a highly unconscious process. The speaker conceives a communicative intention, he knows what is the output of the speech process, but he does not know what exactly takes place within the speech production process. The reason for this is the fact that he cannot look inside the processes taking place at the level of the Syntactic Formulator and Information Structure Level.

Both Dutch headline writers as well as headline readers 'know' intuitively that (a) is a better headline than (b), but they do not know why.

OPERATIONAL IN A YEAR' the headline with 'IN A YEAR' is better. The reason for this, according to Dor, is the fact that a headline with a statement like 'IN FEBRUARY' forces the reader to calculate the amount of time it will take till the casino will be operational. In the second version this amount of time is given, hence, it leads to less processing effort.

Interestingly we see that the headlines used in the study of Dor (translations of headlines in Hebrew) contain articles where, if they had been English or Dutch headlines, we would have expected omission. This is because they are literal translations of the Hebrew versions Dor used in his study. Obviously Hebrew is far less permissive with respect to article omission in headlines than English or Dutch. Dor does not discuss crosslinguistic differences in his study.

- (12) a. REGERING NEEMT DRASTISCH BESLUIT
government makes radical decision
- b. DE REGERING NEEMT EEN DRASTISCH
BESLUIT
the government makes a radical decision

In chapter 4 I will propose an account that offers insight in the processes taking place at the levels of the Syntactic Formulator and Information Structure and that explains, based on an information-theoretical approach, why (a) is a better headline than (b).

3.3 Database of Headlines in Italian and Dutch

In the previous section we saw that so far almost all linguistic studies on headlines have concentrated on one language, mostly on English. If we want to develop a theory that can account for article omission in adults' special registers, like headlines, and child speech, concentrating on only one language may give us a misleading view because it would be too restricted. I have therefore decided to conduct a crosslinguistic study on headlines. Previous studies on article omission in child speech (Guasti et al., 2008) have shown that there are differences between article omissions of Dutch and Italian children. Dutch children omit more articles and during a longer period of time than Italian children. It is therefore interesting to also compare article omission in headlines in these two languages.

In this section I will present the results of my corpus investigation of article omission in headlines in Dutch and Italian newspapers. In the next section I will present the results of an experiment on headlines in which readers were asked to give their judgements on headlines in which articles were used or omitted.

Before introducing the data of the omission of articles in modern newspapers let me start with an interesting observation on the development of article omission in newspaper titles over the years. In her diachronic study on the headlines in the London Times from 1870 – 1970 SimonVandenBergen (1981) found significant changes in the use of articles over the years: In the Times of 1870 31,5% of all nouns were preceded by an article, in the 1970 Times only 2,2%. These diachronic changes already suggest that in order to account for article omission, we

will necessarily have to look beyond a purely structural account. It is not the case that over time structural changes are impossible (see Lightfoot, 1999), but if article omission in headlines is purely the result of a structural change, why then have these structural changes taken place only in the headline registers, and not in normal speech?

3.3.1 Introduction: database set-up

The database is a collection of 1000 headlines of Dutch and Italian newspapers that appeared in the period October – December 2003. The newspapers I used were: *De Volkskrant*, *De Telegraaf* and *NRC (Nieuwe Rotterdamsche Courant)* for Dutch, and the *Corriere della Sera* and the *Repubblica* for Italian. I used the paper versions of the newspapers, bought at a newspaper stand, and not the digital versions. I did this for two reasons. First, because an initial comparison between the two versions of the same newspaper on the same day showed that there are differences between the digital and the ‘paper’ versions. I wanted to prevent interference from these differences with my results. Secondly, I had the intuitive idea that more care is taken in the construction of headlines for the paper version. After all, that is the one that is for sale in the shops and has to compete with the other newspapers, and it seems reasonable to assume that this competition partly takes place on the basis of the information contained in the headlines.

In previous studies on article omission in child speech a relation was found between omission of articles and the presence of a finite verb in the sentence and between omission of articles and the position of the article requiring noun in the sentence (Hoekstra & Hyams 1996, 1998; Clahsen et al. 1996; Baauw et al. 2005). Therefore these features were taken in consideration in the collection of the headlines. The collected headlines were examined for the following features. I will give a more detailed explanation for the reasons for these distinctions later in this section, in the discussion of the results of the search:

- Instances of non-legitimate omission of articles: headlines in which in article-requiring-noun contexts, which in normal standard adult grammar would have required the use of an article, articles were omitted.

- Instances of legitimate omission of articles: cases in which the omission of the article (use of a Bare Noun) is required by the standard rules of grammar of the language.
- Instances of correct use of the article: sentences in which all obligatory articles were present and used according to the standard rules of grammar.
- Absence of a verb in so-called 'noun phrases in isolation' contexts
- Presence of a finite verb
- Presence of a non-finite verb (participle) with either the auxiliary verb present or omitted
- Presence of a (root-) infinitive
- Position of the noun: sentence-initial or sentence-internal. In the case of nouns in sentence-internal position: whether the noun was used in combination with a preposition or not.
- Use and omission of articles within Hanging Topic Constructions

The spreadsheet used for my analysis was developed in such a way that it was possible to look at relations between all the abovementioned features. So it was, for instance, possible to look at the omission of articles before nouns in sentences with a finite verb, or to look at the omission of articles before nouns following a preposition in a sentence with a finite verb, or to look at a relation between the presence of obligatory articles and the presence or absence of a finite verb. Table 1 presents the overall percentages of article production and omission in both languages.

Table 1 Percentages of article omission in the headlines examined ($p < 0.0001$)

Language	Non-standard omission in verbal headlines	Omissions in noun-phrases-in-isolation	Total omission rate verbal and isolation contexts
<i>Dutch</i>	89,9	83,1	87,9
<i>Italian</i>	5,3	50,8	19,0
	Fisher's exact <.0001	$\chi^2 = 78.345$ $p < .0001$	Fisher's exact <.0001

This table shows the omission rates of articles in so-called 'obligatory' contexts, the contexts that in the adult grammar would have required the use of an article, like in the following examples:³¹

- (13) BANKOVERVALLERS VERLIEZEN GESTOLEN
PINAUTOMAAT
Bankrobbers lose stolen cash register
PITBULL FERISCE DUE BAMBINI
Pitbull attacks two children

and the omission of articles in nouns in isolation, like for example:

- (14) EERSTE DICHTBUNDEL VAN LIBRIS-WINNAAR
BENALI
First poetry collection of Libris-winner Benali
(15) COMPLEANNO DI CIAMPI
Aniversary of Ciampi

On the one hand the data in Table 1 show that there is a striking difference between omissions of articles in Italian and Dutch. On the other hand, they show that, in spite of the differences, article omission in

³¹ A headline of the type [Bare N – V – Bare N] with two article requiring contexts was counted as 2 omissions.

headlines is a truly real phenomenon, in Dutch as well as Italian. We also see that omissions are optional: sometimes articles are used, sometimes they are not. The large difference between article omissions in Dutch and Italian headlines makes it clear that, as I already suggested in the previous section, omission of articles in newspaper headlines cannot only be ruled by functional considerations, like space restriction, alone, since space is obviously restricted in Italian as well as in Dutch newspapers. The observation that article omission in headlines is optional gives rise to further questions. Where and when are articles used? What is it that triggers the production of an article? Or, to put it the other way round: Where and when are articles omitted? What is it that suppresses the production of an article?

Let us therefore examine the omission of articles in both languages in different sentence positions, in different types of linguistic context. For both languages I will compare:

- omissions in sentences containing a finite verb versus omissions in sentences containing no finite verb.
- omissions before nouns directly following a preposition and omission before nouns in non-preposition contexts.
- omissions in sentence-initial position versus omissions in sentence-internal position.
- omissions in so-called ‘noun-phrases-in-isolation’ contexts, contexts in which we only find a noun, that can eventually be modified by an adjective or a complement.
- omissions in so-called ‘Hanging-Topic’ constructions.

3.3.2 Effect of finiteness and preposition

3.3.2.1 Introduction

Here we will take a closer look at the relation between article omission and finiteness in the headlines in Dutch and Italian. Further we will examine the relation between article omission and the presence of a preposition. Let me first describe how I made the distinction between headlines *with a finite verb* and headlines *without a finite verb*. Headlines with a finite verb are headlines in which a verb shows number agreement with the subject. A real tense specification on verbs is usually absent in headlines. The tense of the verb is mostly present tense, even if the headline refers to past or future events (see example 6). In Dutch 95,9%

of all the headlines with a finite verb were in the present tense, in Italian 74,8 % .

- (16) BEJAARDE MAN VERDRINKT IN GRACHT NA
STARTEN AUTO

Ancient man drowns in canal after starting car

AJAX SLAAT DUBBELE AANVAL AF

Ajax repels double attack

PILOTA USA SI SCHIANTA E MUORE A INDIANAPOLIS

USA pilot crashes and dies in Indianapolis

ROGO UCCIDE QUATTRO RUMENI

Fire kills four Romanians

Headlines without a finite verb are:

- headlines with a past participle, used without an auxiliary (see example 17)
- headlines with an infinitive (see example 18)
- headlines in which the copular verb is omitted (see example (19), or in which a lexical verb is omitted. This last option is possible in constructions with frequently occurring lexical verbs, see example (20)
- headlines in which ‘er is’/ ‘er komt’ ‘c’è ci sara” (there is / there will be) is omitted (see example 21)

The following examples illustrate each of these types of headlines:

- headlines with past participle but without an auxiliary:

- (17) MAN GEDOOD IN COFFEESHOP

Man killed in coffeeshop

CONTROLE OP DRUGSGEBRUIK IN VERKEER
UITGESTELD

Control on drug use in traffic postponed

MILANO, DONNA UCCISA: ARRESTATO IL FIGLIO

Milan, woman killed, the son arrested

L'AUTOPARCO DI MILANO CONTROLLATO DALLA MAFFIA'

Milan carpark controlled by the maffia

- headlines with an infinitive:

(18) VANACHTER PC VEILIGHEID HAVEN TESTEN
Test safety harbour from behind computer

OM VIJF UUR OPSTAAN VANWEGE FNV-ACTIES
Rise at five o'clock because of FNV-strike

UN ERRORE CHIUDERE LE PORTE AI PAESI PIÙ DEBOLI

Mistake (to) close the doors for weakest nations

PIÙ FACILE COSTRUIRE NUOVE CENTRALI
Easier (to) construct new power plants

- headlines in which the copular verb is omitted:

(19) BLOTE KARIN BLOEMEN NIET NIEUWSWAARDIG
Nude Karin Bloemen not newsworthy

BAVARIAGEVOEL WEG UIT BRABANTS DORP
Bavaria feeling gone from village in Brabant

SANREMO: TONY RENIS DIRETTORE ARTISTICO
Sanremo: Tony Renis artistic director

LA PRINCIPESSINA RUFFO NUOVA FIAMMA DI WILLIAM
The Princess Ruffo new flame of William

- headlines in which a lexical verb is omitted: It is important to note that only a limited set of lexical verbs can be omitted. The omitted verbs are frequently used intransitive verbs such as *'gaan'* / *'andare'* ('go'), *'staan'*, ('stand'), *'zitten'* ('sit').

- (20) NEDERLANDSE UITGEVERS TEGENOVER ELKAAR IN VS
Dutch publishers face to face with each other in USA

VIJF VERDACHTEN VAST NA DOOD DIEVEGGE
Five suspects in prison after death thief

FISICA, NOBEL A DUE RUSSI E UN BRITANNICO
Physics, Nobel to two Russians and a Briton

LA SCALA VERSO L'INTESA, CON UNA NUOVA FORMULA
The Scala towards the agreement, with a new formula

- headlines in which 'er is' / 'er komt' 'c'è ci sarà" (there is / there will be) is omitted:

- (21) SCHERPERE CONTROLE OP VUURWERKSMOKKEL BELGIË
Sharper inspection of firework smuggling Belgium

KLACHT TEGEN DELL OM FOUTE REGELTJES
Complaint against Dell for wrong rules

I CATTOLICI: VERITÀ IMPORTANTI ANCHE NELLE ALTRE FEDI
The Catholics: important truths also in other religions

PATENTE, PER I CAMIONISTI CORSI DI RECUPERO FAI-DA-TE
Driving license, for the truckdrivers do-it-yourself resettlement courses

As the examples show, the category ‘headlines without a finite verb’ also contains headlines with no verb at all. The difference between this category and the category ‘noun phrases in isolation’ is that in the former category the verb is omitted, but there is still a position where the verb could be inserted. In ‘noun phrases in isolation’ no verb is possible in the structure. Therefore, in the ‘headlines without a finite verb’, a sentence-initial and a sentence-internal position could be distinguished. This made them more similar to verbal headlines than to ‘nouns in isolation’, since it was, in spite of the absence of a verb, still possible to give a judgement on the grammaticality or ungrammaticality of the omission of the articles on the nouns in the headlines. In ‘noun phrases in isolation’ this judgement was not possible.³²

Stowell (1999) and Avrutin (1999) have observed that a particular construction, the ‘TO+infinitive’ construction, frequently occurs in English newspaper headlines. Example (22) illustrates this construction:

- (22) GORE TO VISIT CHINA, WILL CONFER WITH ZHANG
(Stowell, 1999)
CLINTON TO SIGN THE CRIME BILL (Avrutin, 1999)
TYSON TO WIN THE FIGHT (Avrutin, 1999)

This construction has no equivalent in Dutch and Italian newspapers. Root infinitives, without TO, do occur, in Dutch and Italian (see example (18) but infrequently (Dutch: 2,5 % of the headlines, Italian 3,5%).

To examine whether the use of articles is influenced by the presence of a preposition, we counted the preposition contexts separately. Example (23) illustrates nouns that directly follow a preposition, example (24) illustrates nouns that do not follow a preposition.

- Article requiring noun, following a preposition.

- (23) HOCKEYTOP OP ZOEK NAAR MIDDENWEG
Hockey-top looking for middle course

DEL PIERO FIRMA CON LA JUVE FINO AL 2008

³² In the remainder of this study I will call headlines without a finite verb ‘headlines with no finite verb’

Del Piero signs for (the) Juventus until 2008

- Article requiring noun, not following a preposition:

(24) ROTTERDAM BESCHERMT GESUBSIDIEERDE ARBEID
Rotterdam protects subsidized jobs

IL GOVERNO DEVE AIUTARE L'ALFA
The government has to support (the) Alfa

3.3.2.2 Results

Let us first look at the effect of finiteness. Table 2 gives the omission percentages before article-requiring nouns, not following a preposition, in headlines with a finite verb and headlines with no finite verb.

Table 2 *Omission percentages of articles before article-requiring nouns not following a preposition, in headlines with a finite verb and headlines with no finite verb (the absolute numbers are indicated between brackets)*

	Finite verb (1)	No finite verb (2)
Dutch	92,9 (681/733)	98,5 (197/200)
Italian	2,1 (13/616)	33,3 (51/153)

The difference between omissions in finite verb headlines and no-finite verb headlines is significant in both languages: Dutch $\chi^2 = 7.884$, df 1, $p = .0050$, Italian: $\chi^2 = 16.616$, df 1, $p < .0001$. Further there is a very sharp difference between the two languages, both in finite verb contexts and in no-finite verb contexts.³³

Let us now look at the effect of the preposition. Table 3 compares the omission percentages before nouns following a preposition and nouns not following a preposition, both for headlines with finite verb and headlines with no finite verb.

³³ In both contexts the difference between Dutch and Italian is significant: Finite verb contexts $\chi^2 = 1101.017$, df 1, $p < .0001$; No-finite verb contexts: $\chi^2 = 11.186$, df 1, $p < .0008$.

Table 3 *Omission percentages of articles before nouns following a preposition and nouns not following a preposition verb (the absolute numbers are indicated between brackets)*

	Finite verb		No finite verb	
	Noun following preposition (1)	No preposition context (2)	Noun following preposition (3)	No preposition context (4)
Dutch	81,0 (213/263)	92,9 (681/733)	79,8 (130/163)	98,5 (197/200)
Italian	0,0 (0/304)	2,1 (13/616)	0,0 (0/146)	33,3 (51/153)

Table 3 shows that the presence of the preposition does have an effect: in both languages we find a lower omission rate in article-requiring noun contexts following a preposition. In Dutch this difference is significant in finite verb contexts: $\chi^2 = 29.878$, df 1, $p < .001$ and in no-finite verb contexts: $\chi^2 = 33.256$, df 1, $p < .001$. In Italian this difference is significant in no-finite verb contexts: $\chi^2 = 23.448$, df 1, $p < .001$. In finite verb contexts in Italian, where the omission rate is very low, regardless of the presence of a preposition, the difference is not significant: Fisher's exact $p = .1449$.

Hence, in both languages we find less omissions if a case-assigning category (finiteness or a preposition) is present. Especially in Italian we hardly find any omissions of articles after prepositions (see Table 3, column (3) and (4)).³⁴

3.3.2.3 Discussion

In several studies on article omission in child speech (Hoekstra & Hyams 1996, 1998; Clahsen et al. 1996; Baauw et al. 2005) and on article omission of agrammatic speakers (Ruygndijk, 2002), a relation was found between the presence of a finite verb (or the preposition) and the production of articles. Language acquiring children and agrammatic

³⁴ The data in Table 3 may suggest that in Italian there are no omissions at all with nouns following a preposition, but that is because the table only contains the verb-containing headlines. In headlines without a verb, omissions of articles before nouns directly following a preposition were present in Italian, but scarce. I will give examples and the exact percentages of omission later in this section when non-verbal headlines are discussed.

patients produce more complete DPs when a finite verb is present. One of the reasons that have been suggested in the literature to account for the effect of finiteness on the production of articles by people with limited processing resources is the difference in case assigning properties between finite and non-finite verbs. Ruygndijk (2002), for example, suggested that problems with the production of determiners by agrammatic speakers are caused by an underlying problem with case or case assignment. She found in an investigation of determiner production and case assignment by German and Russian agrammatic patients that once the (finite) verb or the preposition is realized, the production of determiners is less impaired than in the situation where the case assigner is absent.

Why should there be a relation between the presence of a case assigner (like a verb or a preposition) and the production of an article? Chomsky (1995) proposed that correctly case-marked subjects depend on the presence of a finite verb in AgrS. The production of correctly case-marked objects depends, in the same way, on the presence of a transitive verb in AgrO. Without a (finite or transitive) verb or preposition no case assignment is possible. Why is this important for the production of articles (or determiners in general)? It has been proposed that presence or absence of a case assigner affects the structure of the Nominal Phrase that accompanies the case assigner. According to Ouhalla, for example, (1994, 1999) the absence of a case assigner renders DPs (nouns with a determiner) ungrammatical, only NPs are allowed then.^{35,36} Since headlines are part of natural language this relationship between finiteness and determiner use should be present in headlines as well, and leads to the prediction that a headline without a case assigning

³⁵ The abbreviation NPs in Ouhalla's terminology is used for bare noun phrases that require a determiner but are realized without one.

³⁶ It is also possible to come to the same conclusion about the relation between finiteness and use of a DP formulating this claim differently, for example, using the feature checking theory (Chomsky, 1993). This was suggested for example by Clahsen et. al (1996) in a study on German child speech. According to the feature checking theory, DP-internal N-features [Gender] and [Number] have to be checked by an appropriate AGR-category. This means that fully specified DP-subjects have to be checked by AGR(S), fully specified direct DP objects by AGR(O). In this view co-occurrence of a root-clause infinitive (i.e. an utterance without AGR (S) with a fully specified DP subject constitutes a licensing conflict and a violation of the principles of Checking Theory.

verb or preposition cannot contain a DP, but only a NP. Only if a case assigning verb or preposition is present can a DP be realized.

Of course, the fact that a headline without a case assigning verb or preposition cannot contain a DP (but only a NP) does by itself not imply that the opposite is also true, e.g. that a headline with a case assigning verb necessarily has to contain a DP and could not occur with a NP. However, for a noun phrase to function as an argument, the category D is necessary (Stowell, 1991, Longobardi, 1994). The crucial nature of the category D for argumenthood is that certain referential properties typical of arguments, such as the semantic meaning of grammatical number are represented on the D position. Therefore, only a DP can be an argument, a NP cannot. NPs are predicative, and the role of D is to enable the noun to function as an argument. Summarizing,

- the presence of a case assigner is a necessary requirement for the realization of a DP
- the realization of a DP is a necessary requirement for a noun phrase to function as an argument.

Still, one could argue that for some reason, there is an option available for headline writers to use a noun phrase without an overt determiner as an argument. After all, a similar option also exists also in normal speech. Longobardi (1994) has suggested that in sentences like *Mary eats potatoes every day. I always drink wine.* an empty discourse operator is present in the D-position, which enables the noun phrases to function as an argument. Following this line of thinking, it could be argued that a similar type of empty discourse operator, active only in the headlines registers, fills the D-position, and thus saves the acceptability of the headline. However this assumption is problematic for a number of reasons.

First, it cannot straightforwardly account for the observed pattern of optionality, and is therefore incomplete. Particularly, it would leave us with the question: why headline writers sometimes use this option (and realize a DP in which the D position is filled by an empty discourse operator), and sometimes do not use this option (and realize a fully specified DP with an overt determiner)? Let me underline that this optionality constitutes a crucial difference between an eventual empty discourse operator active in headlines, and the empty discourse operator proposed by Longobardi for a sentence like *Mary eats potatoes every day.* The empty discourse operator proposed by Longobardi is *not* optional, since filling the D position with an overt determiner would change the

meaning of the utterance.³⁷ In headlines, however, the use of the empty discourse operator is optional. Use of an overt article does not change the meaning of the headline. Further, if languages offer the possibility to *optionally* use an empty discourse operator instead of an overt determiner, then why is this possibility not available for normal adult speech? Another problem for theories that relate article omission to the absence of case assigning elements is that they cannot explain why the effect of the presence of case assigners is stronger in Italian than in Dutch. In headlines with a finite verb, for example, the omission rate in Italian drops down to only 2,1%, where in Dutch it is still 92,9%. Nor can such theories explain why, overall, in all contexts, we find less omissions in Italian. After all, there are no differences between the two languages with respect to case realization. Both languages lack overt realization of case on the article. A theory that relates article omission to case assigning properties would predict that we should find more or less the same pattern in both languages.

3.3.3 Effect of position

Several studies on omission of articles by children, agrammatic speakers and adults in special registers have suggested that there is a difference in omission pattern, depending on the position in which the article-requiring noun appears, sentence-initial or sentence-internal. Let us take a closer look at some of these studies. In his study on headlines in English newspapers, Stowell (1999) observed a difference between omissions in sentence-initial and sentence-internal position. He found that determiner omission from the matrix subject DP is strongly preferred, and ‘arguably obligatory’ in a news flash register, while determiner omission from the direct object DP, though sometimes mildly preferable, is optional. Determiner omission from the direct object is impossible if an article has not been omitted from the subject. Stowell found more omissions of articles in sentence-initial position, and related this to effects of c-command.³⁸

³⁷ Or, depending on the context, even make it uninterpretable. For example, *Mary eats the potatoes every day*, cannot be interpreted in an out-of-the-blue context.

³⁸ Stowell (1999) suggests that the highest NP in the clause is marked +R if it is reduced (if it is not the complement of a D-head) and –R if it is not reduced. He then proposes that the constraint that regulates article omission can be expressed as follows:

Previous studies on article omission in child and agrammatic speech have found a similar pattern: more omissions from sentence-initial position (for child speech: Baauw et al. 2005; Guasti et al. 2004; for agrammatic speech: Ruygendijk, 2002; Guasti et al. 2004).

In fact, omission of elements from clause-initial position is not restricted to articles. Several studies have shown that also clausal subjects and topicalized object pronouns are optionally omitted from initial position in child speech, agrammatic speech and adults' special registers (e.g., Rizzi, 2000; Haegeman, 1990 and De Roo, 1999). It has been suggested that omission of functional elements from the highest clausal position is a universal option of UG (Guasti et al. 2004). This option is exploited under conditions of reduced capacity of the production system, in agrammatics as well as in children (Kolk, 2000 ; De Roo et al., 2003). I will discuss this proposal further at the end of this section when I discuss the results of the analysis of the headlines in Dutch and Italian.

Example (25) illustrates headlines with an article-requiring noun in sentence-initial position, in example (25a) the article is omitted, in example (25b) the article is used.

- (25) a. TURKS LEGER UIT KRITIEK OP
 ONDERWIJSPLAN
 Turkish army criticizes education plan
- b. IL GOVERNO SCEGLIE ORTIS
 The government selects Ortis

Example (26) illustrate headlines with an article-requiring noun in sentence-internal position, in example (26a) the article is used, in example (26 b) the article is omitted.

- (26) a. REGERING EN OPPOSITIE CLAIMEN DE ZEGE
 IN AZERBEIDZJAN

a +R NP cannot be c-commanded by a –R NP

This constraint is satisfied in the first clausal position, because there a +R NP is not c-commanded clause-internally. It is also satisfied in a clause-internal position provided that no appropriate 'antecedent' (a –R NP) is available.

Government and opposition claim the victory in
Azerbeidzjan

- b. ISRAELE ATTACCA CAMPO JIHAD IN SIRIA
Israel attacks Jihad settlement in Siria

Table 4 shows the omissions in initial and internal position in verbal headlines. Since the analysis of omission in the preceding section showed that there is an effect of the presence of a preposition on the omission of articles I have eliminated here the article-requiring nouns that directly follow a preposition. The results we obtain thus show us the effect of sentence position only, not mixed with an effect of preposition.

Table 4 Omissions in article-requiring-noun contexts in initial and internal position, in headlines with finite verb and headlines without finite verbs

	Finite verbs		No-finite verbs	
	Initial	Internal (non- prep.)	Initial	Internal (non- prep.)
<i>Dutch:</i>				
Contexts	456	277	186	14
Number of omissions	434	247	184	13
% omission	95,2 %	89,2 %	98,9 %	92,9 %
<i>Italian:</i>				
Contexts	273	343	86	67
Number of omissions	11	2	42	9
% omission	4,0 %	0,6 %	48,8 %	13,4 %

These findings confirm the findings of previous studies on the effect of position. In both languages we find an effect of position. In Dutch as well as in Italian in verbal headlines we find a higher omission rate from sentence-initial than sentence-internal position. In Dutch the differences are significant for finite verbs: $\chi^2 = .8.541$ df 1, $p = .0035$. For the headlines with no-finite verbs in Dutch the very low absolute number of internal contexts makes it difficult to draw a strong conclusion. The difference between omissions from initial and internal position does not reach significance (Fisher's exact 0.1966), but this may be caused by the low number of contexts.

In Italian the differences between omissions from initial and internal position are significant for finite and no finite verbal headlines: (finite

verbs: $\chi^2 = 7150$ df 1, $p = .0075$, no finite verbs: $\chi^2 = 19.679$, df 1, $p < .0001$).

As I mentioned earlier in this section, a proposal that has been put forward to account for omissions from sentence-initial position claims that omission of functional elements from sentence-initial position, the highest clausal position, is a universal option of UG (Guasti et al. 2004). This option is exploited under conditions of reduced capacity of the production system, as in agrammatics and children. However, if we want to use this proposal to account for the omissions we find in headlines we have to answer at least three questions:

- In our database we find far more omissions from initial position in Dutch headlines than in Italian headlines. This would suggest that the option to omit functional elements from initial position is used more often in Dutch. It is, however, not a priori clear why this universal option should be exploited more often in Dutch than in Italian.³⁹
- Children and agrammatics may exploit this option because of their reduced processing resources. But why should headline writers exploit this option? Of course it could be argued that ‘space restrictions’ in newspapers lead to the same effect on the output of the speech production system as restricted processing resources do, but it is not clear what such a claim would explain. Why should space restrictions have such an effect? Omission of an article from sentence-internal position saves exactly the same amount of space as omission from an article in sentence-initial position. Still, we find more omissions from initial position. And, even if we could show that the claim is legitimate, then, again, why do we find crosslinguistic differences? Space is equally restricted in Dutch and Italian newspapers.
- If omission from initial position is a universal option of UG, then it is not a priori clear why in Italian the effect of position

³⁹ Guasti (p.c.) has suggested that ‘one could argue that omission from initial position is a universal that is exercised to different degrees, possibly depending on a processing account. Null subject is an option in Italian and Chinese, but in standard Italian dialogue omission is higher than in Chinese. A processing approach may account for these differences’. This is an interesting observation that certainly merits further investigation.

stronger is in no-finite verbal headlines than in finite verbal headlines.

3.3.4 Noun Phrases in isolation

In the preceding section we examined omission of articles in *verbal* headlines, headlines with the verb either in finite or no-finite form. However, we do not only find verbal headlines in newspapers. A large number of headlines are phrases that not only do not contain a verb, but differently from the previously discussed headlines, could not contain one. These are headlines that look like ‘labels’ that announce the topics of the articles. They take a variety of structures, as exemplified in (27)

- (27) - a noun with modifiers
 THE CANADIAN BY-ELECTIONS
 - noun + preposition+noun,
 THE SECRET LIFE OF MOODY COWS
 - noun + conjunction+noun
 BARTJE EN DE GEORGANISEERDE MISDAAD
 Bartje and the organized crime
 - headline with initial preposition
 UNDER THE OLD CITY WALL

I have categorized these ‘label-like’ headlines as ‘Noun Phrases in isolation’. In her (1981) study SimonVandenBergen observed a striking change in the use of verbal headlines versus nominal headlines (‘NPs in isolation’in my terminology:). In 1870 the frequency of verbal headlines was only 3,75%. In 1914 it was 15,15%, in 1940 it had already increased to 26,77% and in 1970 the great majority of headlines in the Times were verbal sentences: 73,75%.

In my database of Dutch headlines of 2003 18,0 % of the headlines were ‘NPs in isolation’, and thus the rate of verbal headlines was 82%. In the Italian database 26,0 % were ‘NPs in isolation’, and the rate of verbal headlines was 74%.

For the category ‘NPs in isolation’ it is not possible to determine objectively whether omission of the article is in conformity with the rules of standard grammar. Both, NPs with an article as well as NPs without

an article are possible. I therefore counted use and omission of articles in these constructions, without making the distinction ‘grammatical omissions’ versus ‘ungrammatical omissions’. Earlier we saw that in verbal headlines the omission rate was influenced by the presence of a preposition. I therefore applied the earlier distinction (with or without preposition) to ‘NPs in isolation’ contexts as well. For nouns directly following a preposition it was possible to determine whether omission of the article is in conformity with the rules of standard grammar. Example (28) illustrates an ungrammatical omission from a NP following a preposition, example (29) illustrates a grammatical omission.

- (28) AANSLAG OP RADICALE PAKISTAANSE POLITICUS
Attack on radical Pakistani politician
- (29) EEN GELUKKIGE MANAGER ZONDER GELD
A happy manager without money

Table 5 *Headlines ‘nouns in isolation’, % of headlines without article (total number of headlines ‘NPs in isolation’ without article divided by total number of headlines in the category ‘NPs in isolation’, the absolute numbers are indicated between brackets)*

	All NPs in ‘NPs in isolation’ contexts	NPs not following preposition	NPs following preposition	
	Without article (1)	Without article (2)	Without article (3)	Of which: Agrammatical omission (4)
Dutch	83,1 (202/243)	86,3 (107/124)	79,8 (95/119)	65,1 (45/69)
Italian	50,9 (398/783)	59,3 (214/361)	43,6 (184/422)	5,6 (14/252)
Pearson chi-square	$\chi^2 = 18.481$, df 1, p <.0001	$\chi^2 = 5.325$, df 1 p =.021	$\chi^2 = 13.249$, df 1 p < .0001	$\chi^2 = 68.624$, df 1 p <.0001

The data in the table show that in both languages we find article omissions in ‘nouns in isolation’, but, just as with verbal headlines, we find more omissions in Dutch. Only for the preposition contexts was an objective decision on the grammaticality of article omission possible. That is why a comparison of agrammatical omission in preposition and non-preposition contexts in isolation constructions is not possible. We can, however, compare the total number of omissions in preposition and non-preposition contexts (columns (2) and (3)), and here we find that in both Dutch and Italian we find a lower omission rate after prepositions. The difference reaches significance in Italian, but not in Dutch.⁴⁰ And, when we compare the data in the last columns (3) and (4) we see, again, that the effect of the presence of the preposition is stronger in Italian than in Dutch. In Italian we observe a very low percentage of agrammatical omissions (5,6%) after a preposition, while in Dutch we still find an agrammatical omission rate of 65,1% after prepositions.

Comparing the omission percentages for Italian with those in the tables for verbal headlines, we observe that we find more omissions in ‘NPs in isolation’ constructions than in verbal headlines. For Dutch, however, we find less omissions in ‘NPs in isolation’ constructions, see Table 6.

Table 6 Omission in verbal headlines and omission in ‘NPs in isolation’ in Dutch and Italian

	Omission rate verbal headlines	Omission rate ‘NPs in isolation’	Pearson chi-square
Dutch	89,9	83,13	$\chi^2 = 9.709$, df 1 p = .0032
Italian	5,3	50,83	$\chi^2 = 555.410$, df 1 p < .0001

⁴⁰ Differences between column (2) and (3): Dutch: $\chi^2 = 0.098$, p = .07544, Italian $\chi^2 = 5.901$, p = .0151

3.3.5 Hanging Topics

In my analysis of the headlines of Dutch and Italian headlines I noticed that a particular type of construction is used very frequently in Italian newspaper headlines, but is hardly ever used in Dutch newspapers: the so-called Hanging Topic construction. Example (30) illustrates this construction, which occurs in 21,0 % of the Italian headlines

- (30)
- a. OSCAR, PER L'ITALIA CORRE SALVATORE
Oscar, for Italy takes part Salavatores
 - b. LAZIO, PER CONTINUARE ORA SERVE UN
MIRACOLO
Lazio, to continue now a miracle is needed
 - c. BANGLADESH, ELEFANTI INFEROCITI: 8
MORTI
Bangladesh, stampeding elephants: 8 victims
 - d. MEDICINA, NOBEL A LAUTERBUR E
MANSFIELD
Medical science, Nobel for Lauterbur and Mansfield
 - e. PATENTE, PER I CAMIONISTI CORSI DI
RECUPERO FAI-DA-TE
Driving licence, for the truck drivers do-it-yourselfes
resettlement courses

In these examples the first element of the headline introduces a topic, followed by a comma, and after the comma the rest of the headline provides further information on this topic. The part of the headline after the comma can be a 'noun in isolation' (examples 31), a verbal phrase with a finite verb (examples 30a and 30b), or a verbal phrase with a non-finite verb (example 30 c). Interestingly the use of these constructions affects article omissions in Italian. We find a high percentage of omission within these topicalized phrases. Example (31) gives a headline with omission from a topicalized phrase, example (32) is an example of a headline with omission after a topicalized phrase, and, example (33) shows that omission in these topicalized constructions is optional. We also find examples of the construction in which articles are used:

- (31) DECRETO ANTI BLACK-OUT, OK DEL SENATO
Decree anti black-out, OK of the senate

- (32) JENIN, ARRESTATO LEADER DELLA JIHAD
Jenin, arreste leader of the Jihad
- (33) TRAVIS, IL RITORNO, RIUNITI DA UN INCIDENTE
Travis, the return, reunited by an accident

Here I will present an analysis of article omission in Hanging Topic Constructions, which can capture the fact that we find a very high rate of article omission in this specific subset of Italian headlines.

For Hanging Topics, it is not possible to determine objectively whether omission of the article is in conformity with the rules of standard grammar. I counted use and omission of articles in these constructions, without making the distinction ‘grammatical omissions’ versus ‘ungrammatical omissions’.

Table 7 *Hanging Topics subdivided into Hanging Topics with article, and Hanging Topics without article, expressed as % of the total number of Hanging Topics (absolute numbers between brackets)*

Articles present	8,3 % (24/290)
No articles present	91,7% (266/290)

In order to answer the question why we find such a high rate of omission of articles in Hanging Topics we first need to find out what exactly is a Hanging Topic construction.⁴¹ Most studies on topic-like constructions (Prince, 1985; Escobar, 1995; Cinque, 1983) agree, following Chomsky (1977), on the distinction between (at least) two types of constructions that contain elements that have been dislocated to the left: Topicalization (if the construction contains a gap) and Left Dislocation (if the gap created by dislocation of the element is filled). The following example from Prince (1985) illustrates the two construction types:

⁴¹ Note that the percentage of omissions in Hanging Topic constructions is even higher than the percentage of ‘Nouns in Isolation’ without an article in Italian (59,3%, see table 5).

- (34) You see every Spielberg movie as soon as it comes out.
 No, E.T., I saw only yesterday (Topicalization)
 No, E.T., I saw *it* only yesterday (Left Dislocation)

The studies differ on the analysis of what exactly can constitute a topicalized element from a pragmatic point of view. I will follow the proposals of Escobar (1995) and Cinque (1983) as they offer the best account for the special characteristics (e.g. high omission rate and informative content) of Hanging Topics in Italian headlines.⁴² Escobar's (1995) analysis focuses on topicalized elements in Spanish, Cinque (1983) analyses topicalized elements in Italian. They both distinguish three types of left dislocation structures, which differ in syntactic and pragmatic functions.⁴³

- (35) Topicalization (TOP)
 Tuo fratello, hanno invitato, non te
 Your brother, (they) have invited, not you
- (36) Clitic Left Dislocation (CLLD)
 A tuo fratello, non gli hanno ancora dato il visto
 To your brother, (they) not to-him-have given the visa yet
- (37) Hanging Topic Left Dislocation (HTLD)
 Tuo fratello, invece, lui sì che aveva sempre fame
 Your brother, however, him yes that (he) was always hungry

Hanging Topics in headlines have the syntactic properties of Hanging Topic Left Dislocation structures (HTLD). Differently from Topicalization-structures (TOP), the fronted element in a Headline-Hanging Topic is not a stressed phrase and differently from Clitic Left Dislocation (CLLD) structures no resumptive clitic pronoun is used to fill up the gap of the dislocated element in the sentence. Hanging Topics, like HTLDs are clause-independent; the link between the dislocated

⁴² Note that these analyses are based on topicalization in normal adult speech, not on topicalization in headlines.

⁴³ The analyses proposed by Escobar and Cinque differ in several respects, but agree on the aspects of Hanging Topics that are relevant for my analysis of these in Italian newspaper headlines.

element and the sentence is provided by discourse, not by a pronominal copy. From a pragmatic point of view as well Hanging Topics in headlines show important similarities with HTLDs. Cinque (1983:9) claims that in HTLD the lefthand phrase is used to bring up or shift attention to a new or unexpected topic. He illustrates this claim by describing what an appropriate context for the HTLD example (37) above would look like:

- (38) (*context: a friend of two brothers recalling childhood with one of them*)
 ‘In those days I remember you would eat only occasionally and unwillingly.....
 Tuo fratello, invece, lui sì che aveva sempre fame
 ‘Your brother, however, him yes that (he) was always hungry

Escobar (1995: 84) argues that “the ‘topic’ in a HTLD structure is used to produce a change of the current topic of conversation and is more or less independent of the notion of new or old information”. It changes the discourse topic (like the English construction: *Let’s now talk about NP... or ‘As for NP.....’*). Escobar suggests that for Topics in HTLD constructions the information status (Hearer Old/ Hearer New) is of no concern at all. This is important for the analysis of omissions of articles in Hanging Topics in headlines. What matters in these constructions is the fact that the speaker (=headlines writer) wants to introduce a new topic of conversation, which may be old for some readers and new for others, but this is not important. What is important is the Topic itself, the concept expressed by the NP. And this means that it functions more or less like a ‘label’, a ‘book title’ or the caption of a photograph or picture . And these are exactly the cases in which the Italian grammar allows omission of determiners. According to Renzi (1988: 401), in these cases determiners are omitted because the NPs will be considered ‘intrinsically determined’, so the content of the (omitted) determiner can be derived from the context. The omitted determiner can be either definite (as in examples (39) and (40) or indefinite (as in example (41), (all from Renzi, 1988).

- (39) *vagone letto*
 Sleeping carriage
 (written on a train carriage)

- (40) guardaroba, segretaria
 wardrobe, secretariat
 (written on door department office)
- (41) carta della Francia
 map of France
 (written under a geographic map, Renzi argues that the omitted
 article here is an indefinite article, comparing: * la cartadella
 Francia/una carta della Francia
 (*the map of France/ a map of France)

We may thus conclude that article omissions from Hanging Topics are in conformity with the rules of standard Italian grammar.

3.3.6 Summary

Let me summarize the findings of the analysis of article omission in Dutch and Italian newspaper headlines.

- In all linguistic context types we find more article omissions in Dutch than Italian headlines.
- There is an effect of finiteness. In both languages the rate of omission is lower in headlines without a finite verb. In Italian the effect of finiteness is stronger than in Dutch (see Table 2).
- There is an effect of preposition. In both languages the rate of omission is lower if the noun directly follows a preposition (see Table 3).
- There is an effect of position. In both languages we find a higher omission rate in initial than in internal position. In Dutch the difference is significant only in finite sentences (possibly due to the low number of internal contexts in no-finite sentences). In Italian the difference is significant in both finite and no-finite sentences (see Table 4).
- In Italian we find more omissions in non-verbal headlines (nouns in isolation and hanging topics) than in verbal headlines. In Dutch we find more omissions in verbal headlines.
- We find a very high omission percentage in Hanging Topics. These constructions are only used in Italian headlines.

3.4 The experiment

3.4.1 Introduction

I argued that headline writers strive to give the reader the best informational value for his cognitive effort. Hence, given the restricted space, they want to provide the reader the most information as possible for the least processing effort. According to this argumentation, omissions seem to be based, at least partly, on intuitive judgements of headline writers on the processing effort of readers. It may be useful to conduct an experiment to find out whether these judgements of headline writers are correct. After all, in normal everyday newspaper life readers have no choice. They are confronted with the newspaper headlines as they are, and as headline writers think the readers like them. But, to my knowledge, no one has ever asked the Dutch newspaper readers whether they really prefer the headlines in which articles are omitted. Perhaps, if readers had a choice, they would prefer headlines with articles present. And a similar question can be asked to Italian newspaper readers. In short, do the headline writers make the correct assumptions, based on their intuitive knowledge of the relation between the amount of information contained in an element and the processing effort necessary to process the element?

3.4.2 Set-up of the experiment

In the previous section with the database of headlines we saw that article omission in headlines is influenced by a number of factors:

- position of the noun: initial / internal
- presence / absence finite verb
- verbal / non verbal headline
- hanging topic

The goal of the headlines experiment was to test whether readers show sensitivity to these conditions. To give an example, in headlines we find more omissions before nouns in sentence-initial position than before nouns in sentence-internal position. So the question was whether we would find a similar effect in the readers' judgements, i.e. whether they

have a stronger preference for omission of articles before nouns in sentence-initial position than before nouns in sentence-internal position.

The experiment consisted of pairs of headlines, based on real headlines from the database. Each test item consisted of two versions of the same headline, forming a minimal pair. In one version the article was present, in the other the article was omitted.

Figure 1 *Example questionnaire*

nr.	KRANTENKOP (HEADLINE)	BESTE KRANTENKOP? (best headline)		
		A	B	Geen voorkeur No Preference
1	A. Het enthousiasme voor Europa neemt af. B. Enthousiasme voor Europa neemt af. (The) enthusiasm for Europe is diminishing			
2	A. Een vrachtwagen ramt ventilator. B. Vrachtwagen ramt een ventilator. (A) truck rams (a) fan			
3	A. Complete Bandidos-bende is gepakt. B. De complete Bandidos-bende is gepakt. (The) complete Bandidos-gang has been caught			

The participants were informed that the experiment was on ‘headlines’. They were asked to give their judgements on which headline they preferred. If they had no preference for one of the two versions, they could indicate this in a separate column of the questionnaire. The participants were newspaper readers, aged between 18 – 65 years old, with no background in linguistics. We tested 34 participants in the Italian version of the experiment and 50 participants in the Dutch version.⁴⁴

⁴⁴ I acknowledge Maria Teresa Guasti for her help in preparing and conducting the Italian headlines experiment.

The experiment had 16 conditions, 10 sentences per condition and was divided into two parts. Each participant had to respond to 80 items. This was done in order to prevent decrease of the attention caused by a (too) large number of test items. Each part contained 5 sentences of each condition, so all participants passed judgements on 5 sentences of each condition.

The test sentences in the conditions with verbal headlines were selected in such a way that when an article had to be used according to standard rules of grammar in either the subject or the object position, the other noun in the sentence (if present) was either a name or a noun that did not require an article (mass noun or plural). This was to prevent judgements from being influenced by the presence of another article in the sentence. So, for example, in the condition that tested article use or omission before a subject noun a sentence like: (DE) CDA-TOP BEKRITISEERT BALKENENDE ((the) CDA-top criticizes Balkenende) was considered to be a good test sentence, while a sentence like (DE) CDA TOP BEKRITISEERT DE REGERING ((the)CDA-top criticizes the government) was not a good test sentence, and therefore not used. The judgement on article use or omission before the subject might have been influenced by the presence (or absence) of the article before the other noun in the sentence (and of course vice versa).

In one separate condition we particularly aimed at testing the influence on judgements by the presence of another noun with or without an article in the sentence (see section 3.4.3.3). In that separate condition we tested sentences in which we alternated the presence of the article: either before the subject noun or object noun.

- (42) (DE) CDA TOP BEKRITISEERT DE REGERING
 (the)CDA-top criticizes the government
 DE CDA TOP BEKRITISEERT (DE) REGERING
 The CDA-top criticizes (the) government

The experiment was set up in such a way that, per condition, half of the items presented had the article present in the first (A-) sentence and the other half of the items had the article present in the second (B-) sentence. The sentences were balanced for the use of the definite articles 'de' and 'het' in Dutch and 'il' and 'la' in Italian. The Italian sentences with indefinite articles were balanced for the use of 'un' and 'una'. In the Dutch sentences with indefinite articles no gender distinction was

necessary, as the indefinite article in Dutch is not specified for gender. All Dutch sentences with indefinite articles used ‘*een*’.

All conditions consisted of a version for definite articles and a version for indefinite articles. This made it possible to determine whether judgements were influenced by the type of article. In the analysis of the headlines database the effect of article type (definite-indefinite) could not be calculated in objective way, because of the fact that not in all cases of omission was it possible to decide which article had been omitted.⁴⁵

3.4.3 The conditions and the results

I will now describe the different test conditions, and give examples of test sentences. I will also give the prediction that could be made on the basis of the results of the analysis of the database, and present the experimental results. In the following discussion I will make a division based on the factors that in the database were found to influence the pattern of article omission:

- presence or absence finite verb
- position of the noun: initial / internal
- nouns in isolation and Hanging Topic constructions

3.4.3.1 Effect of finiteness

In order to find out whether the participants’ judgements on determiner use were influenced by the presence or absence of a finite verb, two versions of one and the same headline were tested, in two conditions:

- The headlines in *condition I* were headlines in which the finite auxiliary verb was omitted. The (definite or indefinite) article was omitted in one version and used in the other version of the headline.
- In *condition II* the same headlines were used, but this time with the finite auxiliary verb present.

Since only auxiliary BE can be omitted in headlines (omission of HAVE does not occur, see section 3.2.4 for a discussion) the headlines were all passive (or unaccusative) constructions, in which only the participle is

⁴⁵ Appendix A gives an overview of all headlines used in the experiment.

used and the article-requiring noun is in sentence-initial position. Example (43) illustrates the headlines used in these conditions.

- | | | |
|------|----------------------------|--|
| (43) | <i>I NO-AUX, NO-Art</i> | VREDESPLAN VOOR HAÏTI
VERWORPEN
Peace plan for Haiti rejected. |
| | <i>I NO-AUX, YES-Art</i> | HET VREDESPLAN VOOR
HAÏTI VERWORPEN
The peace plan for Haiti rejected. |
| | <i>II YES-AUX, NO-Art:</i> | TURISTA È DISPERSA IN
SVIZZERA
Tourist is lost in Switzerland |
| | <i>II YES-AUX, YES-Art</i> | UNA TURISTA È DISPERSA
IN SVIZZERA
A tourist is lost in Switzerland |

The results of these conditions I and II will be analyzed jointly. This will enable us to find out whether judgements on determiner use were influenced by the presence or absence of a finite verb.

Prediction

The analysis of the headlines in my database showed that article omission is influenced by the presence of a finite verb (see Table 2). Basing ourselves on this analysis, the experimental prediction for the experiment is that the Dutch participants will prefer the version of the headline without the article, but their preference for the articleless version of the headline will be less strong in condition II, where the finite auxiliary verb is present, than in condition I, where the finite auxiliary verb is omitted. As Table 2 showed, we only find a very low rate of article omissions in Italian headlines with a finite verb (omissions in finite verbal headlines in Italian: 2,1%, in Dutch: 92,9%). Therefore the prediction is that in the experiment the judgements of the Italian readers on conditions I and II will differ more strongly than the judgements of the Dutch readers. For Italian a stronger effect of the presence of a finite auxiliary verb on the judgement on article omission is expected. It may very well be the case that in the Italian experiment the preference for the version with either the article present or omitted changes ‘abruptly’ with the presence or absence of a finite verb. In other words, if the finite auxiliary verb is omitted, the readers will prefer the version of the

headline with the article omitted, if the finite auxiliary verb is present the readers will prefer the version of the headline with the article present.

Results

I will report the participants' preferences for *omission* of the articles. As I mentioned earlier, participants did in fact have three choices in the questionnaire: version with article – version without article – no preference. However, in most conditions the number of 'no-preference' answers was fairly low (0-2% of the answers), hence the preference for use can be more or less considered to be the reverse of the preference for omission.

Table 8 *Preferences for omission in headlines with either the finite auxiliary present or omitted, comparison Dutch - Italian*

Condition	Dutch	Italian	One way Anova (Dutch vs. Italian)
NO-AUX	87,6	82,9	F(1,82) = 3.442, p= 0.676
YES- AUX	62,6	12,9	F(1,82) = 66.124, p <.001

Effect of finiteness, within the two languages (NO-AUX- versus YES-AUX)

In both Dutch and Italian the difference between the preferences for omission in the condition with and the condition without auxiliaries was significant.⁴⁶

Comparison Dutch and Italian

In both languages participants have a stronger preference for omission of the article when the finite verb is omitted, but, as was the case for the omission rates found in the database (see Table 3), this difference was significantly stronger in Italian than in Dutch (Anova, interaction

⁴⁶ For Dutch, the difference between the NO/YES AUX conditions with definite articles (NO-AUX-Def – YES-AUX-Def) was significant : Anova repeated measures Wilks' Lambda F(49,1) = 27.323, p <.001, and also the difference between the NO/YES AUX conditions with indefinite articles (NO-AUX-Ind – YES-AUX-Ind): Anova repeated measures Wilks' Lambda F (49,1) = 26.098, p <.001). For Italian, for the difference between the NO/YES AUX conditions with definite articles: Anova repeated measures Wilks' Lambda F(33,1) = 187.326, p <.001, for indefinite articles: Anova repeated measures Wilks' Lambda F (33,1) = 300.515, p<.001

language/omission aux $F(1,82) = 38,150$, $p < .001$). The differences between Dutch and Italian in the headlines where no finite auxiliary or copular verb was present were not significant; the differences in the headlines where the finite auxiliary verb was present were significant (see Table 16 for the results of the statistical analysis).

If we compare Table 2 with the percentages of omission found in Italian newspaper headlines in finite and no finite headlines to Table 8 with the preferences for omission of the Italian participants, we see that the preference percentage for omission in the experiment is 82,9% when the finite auxiliary verb is omitted and 12,9% when the finite auxiliary verb is used. The actual omission in Italian headlines found in the database is 33,3% when no finite verb is used, and 2,1% when a finite verb is present. We do find a similarity in pattern, but we observe that the omission rates are somewhat different. We will see later that this difference should not be surprising. One of the reasons why it arises is the way the test items were formulated: participants had to choose from a minimal pair, the version with and without the article. This construction has undoubtedly accentuated the absence / presence of the article in the headline. Since I wanted to measure the effect the presence or absence of the article had on the preferences of the readers, this was a good effect for the experiment. In real life newspaper titles headline writers do not have a choice between two versions differing only in one respect. As we saw in the study of Dor (2003), headline writers have a very wide choice of alternative headlines, and the role of the article may very well be less pronounced than in the experiment. What is important is the similarity in pattern that we seem to find.

Summary Finiteness effect**Table 9**

Nr.	Condition	Example	Prediction (on basis of database analysis)	Results
<i>I</i>	NO-AUX	(HET)VREDESPLAN VOOR HAÏTI VERWORPEN (The) peace plan for Haiti rejected	Pref.omission article in I > Pref.omission article in II	As predicted
<i>II</i>	YES- AUX	(HET) VREDESPLAN VOOR HAÏTI IS VERWORPEN (The) peace plan for Haiti has been rejected.	Dutch pref for omission > Italian pref. for omission.	As predicted
			Difference in preferences for omission articles I-II stronger in Italian than in Dutch	In condition II, <i>with</i> auxiliary verb as predicted, but in condition I, <i>without</i> auxiliary verb no difference between Dutch and Italian

3.4.3.2 Effect of position (1): sentence-initial and sentence-internal

The conditions in this category consisted of headlines with a finite verb, with an article-requiring noun in either (sentence-initial) subject position (conditions III and IV) or (sentence-internal) object position (conditions IV and V). Examples (44) and (45) illustrate both condition types:

- (44) *III .+ IV. Headlines with finite verbs, with or without a definite/indefinite article, before sentence- initial subject:*

III: *IniSub- Def-YES* : HET ENTHOUSIASME VOOR
EUROPA NEEMT AF
The enthusiasm for Europe is
diminishing

III: *IniSub- Def-NO* : ENTHOUSIASME VOOR EUROPA
NEEMT AF
Enthusiasm for Europe is diminishing

IV: *IniSub- Indef- Y:* EEN AMSTERDAMMER BEDREIGT
BALKENENDE
A man from Amsterdam threatens
Balkenende

IV: *IniSub- Indef- N* AMSTERDAMMER BEDREIGT
BALKENENDE
Man from Amsterdam threatens
Balkenende

- (45) *V.+ VI. Headlines with finite verbs, with or without a definite/indefinite article, before sentence-internal object*

V: *IntObj- Def-YES* DEL PIERO CERCA LA PARTITA
PERFETTA

Del Piero is looking for the perfect match

V: *IntObj- Def-NO* DEL PIERO CERCA PARTITA
PERFETTA

Del Piero is looking for perfect match

VI: <i>IntObj-Indef-Y</i>	CLOONEY COMPRA UNA VILLA A COMO Clooney buys a residence in Como
VI: <i>IntObj-Indef-N</i>	CLOONEY COMPRA VILLA A COMO Clooney buys residence in Como

Prediction conditions III, IV, V and VI

The database analysis for omission from sentence-initial subject nouns in finite sentences showed a preference for omission in Dutch of 95,2%, and in Italian of 4,0 % (see Table 4).

This leads to the following experimental prediction for the conditions III and IV: in Dutch the headline without a determiner will be preferred; in Italian the headline with a determiner will be preferred.

The database analysis for article omission from sentence-internal (object) nouns in finite sentences showed a preference for omission in Dutch of 89,2 % , and in Italian of 0,6% (see Table 4).

The resulting experimental prediction for the conditions V and VI is that in Dutch the headline without the article will be preferred while in Italian the headline in which the article is present will be preferred. In both languages the preference for article omission in the sentence-internal, object conditions (IV and VI) should be less than in the sentence-initial, subject conditions (III and V).

Results

Table 10 shows the preferences for *omission* of the articles in the four conditions in both languages.

Table 10 Preferences for omission of the articles in the conditions *IniSub Def* – *IniSub Indef* – *IntObj Def* – *IntObj Indef* in Dutch and Italian

	Dutch	Italian
III. <i>IniSub Def</i> : definite article in sentence-initial subject	88,8	21,2
IV. <i>IntObj Def</i> : definite article in sentence-internal object	79,6	18,8
V. <i>IniSub Indef</i> : indefinite article in sentence-initial subject	91,7	80,6
VI. <i>IntObj Indef</i> : Indefinite article in sentence-internal object	84,4	45,9

IniSub Def and IntObj Def, Dutch and Italian compared

In both conditions the differences between Dutch and Italian are significant (subjects: $F(1, 82) = 4,81$, $p = 0,0309$; objects: $F(1, 82) = 12,008$, $p < .001$). In both conditions, with a definite article in subject and in object, we find a very strong preference for omission in Dutch. The low percentages show that in Italian the general preference is *not* to omit, but to use the definite article, both in subject as well as object. It is important, however, to note that although there is a strong preference in Italian not to omit the article, we still find optionality. The preference for use of the article in Italian is not absolute (nor is the preference for omission of the article in Dutch).

IniSub Def and IntObj Def, compared within the two languages

Analysis of the differences within the two languages between the *IniSub Def* and *IntObj Def* conditions in the experiment, shows that for Dutch the difference between subject and object was significant ($F(1, 82) = 12,221$, $p < 0,01$). That is to say that in Dutch there is a stronger preference to omit the article before the sentence-initial subject. For Italian the difference between the preferences for omission of the

definite article before the subject and the object was not significant: ($F(33,1) = .327, p = .571$).

Ini Sub Indef and IntObj Indef, Dutch and Italian compared

In both conditions the differences between Dutch and Italian are significant (subjects: $F(1, 82) = 5.57, p = 0.021$; objects: $F(1,82) = 36.49, p < .001$). In both conditions, with indefinite articles in subject and in object, we find a stronger preference for omission in Dutch.

IniSub Indef and IntObj Indef, compared within the two languages

In both languages the differences in preference for omission between the subject and object condition (*IniSub Indef* and *IntObj Indef*) were significant. In both languages we find a stronger preference for omission of the article before the subject (Dutch: ($F(49,1) = 6.76, p = .012$), Italian: ($F(33,1) = 50.912, p < .001$))

Effect of definiteness: IniSub def versus IniSub Indef and IntObj Def versus IntObj Indef in Dutch and Italian:

We used test items with definite and indefinite articles. This made it possible to examine whether the type of article influenced the participants' judgements. In Dutch we found no significant effect of definiteness. Preferences for article omission from subject or object did not differ significantly between the conditions with definite and indefinite articles:

- *IniSub Def* versus *IniSub Indef*: $F(49,1) = 1.945, p = .169$
- *IntObj Def* versus *IntObj Indef*: $F(49,1) = 2.309, p = .135$

This is not so for Italian. In Italian the preference for omission is stronger with indefinite articles, both from subject ($F(33,1) = 186.914, p < .001$) as well as from object noun phrases ($F(33,1) = 24.135, p < .001$).

Summary effect of position**Table 11**

Cond.	Example	Prediction (on basis of database analysis)	Results
<p><i>III.</i> <i>IniSub</i> <i>Def</i></p> <p>def. in subject</p>	<p>(HET)ENTHOUSIASME VOOR EUROPA NEEMT AF</p> <p>(The) enthusiasm for Europe is diminishing</p>	<p>Dutch: pref.omission > pref. use</p> <p>Italian: pref. omission < pref. use</p> <p>Dutch: pref.omission > Italian pref.omission</p>	<p>As predicted</p> <p>As predicted</p> <p>As predicted</p>
<p><i>IV.</i> <i>IntObj</i> <i>Def</i></p> <p>def. in object</p>	<p>DEL PIERO CERCA (LA) PARTITA PERFETTA</p> <p>Del Piero is looking for the perfect match</p>	<p>Dutch: pref.omission > pref. use</p> <p>Italian: pref. omission < pref. use</p> <p>Dutch pref.om > Italian pref.om</p> <p>Dutch + Italian: pref. om <i>IniSub Def</i> > pref.om.<i>Int Obj Def</i></p>	<p>As predicted</p> <p>As predicted</p> <p>As predicted</p> <p>Dutch as predicted, Italian pref.om <i>IniSub Def</i> = pref. om. <i>IntObj Def</i></p>

Cond.	Example	Prediction (on basis of database analysis)	Results
<p>V. <i>IniSub</i> <i>Indef</i></p> <p>indef. in subject</p>	<p>(EEN) AMSTERDAMMER BEDREIGT BALKENENDE (A) man from Amsterdam threatens Balkenende</p>	<p>Dutch: pref.omission > pref. use</p> <p>Italian: pref. omission < pref. use</p> <p>Dutch pref.om > Italian pref.om</p>	<p>Dutch: as predicted</p> <p>Italian: pref.om > pref.use</p> <p>As predicted</p> <p>+ Italian definiteness eff.: pref.om def. < pref. om. indef.</p>
<p>VI. <i>IntObj</i> <i>Indef</i></p> <p>indef. in object</p>	<p>CLOONEY COMPRA (UNA) VILLA A COMO Clooney buys (a) residence in Como</p>	<p>Dutch: pref.omission > pref. use</p> <p>Italian: pref. omission < pref. use</p> <p>Dutch pref.om > Italian pref.om</p> <p>Dutch + Italian: pref. om <i>IniSub Indef</i> > pref.om. <i>IntObj Indef</i></p>	<p>As predicted</p> <p>Less strong than predicted</p> <p>As predicted</p> <p>As predicted</p> <p>+ Italian definiteness effect: pref.om def. < pref. om. indef.</p>

3.4.3.3 Effect of position (2): order of omission

In this condition the headline contained two article-requiring nouns. In one version of the headline the article was used before the sentence-initial subject and omitted before the object, in the other version of the headline the article was omitted before the sentence-initial subject and used before the object. Hence, in each version an article was present. Participants were ‘forced’ to make a choice between use of article before the subject noun or before the object noun. Example (46) illustrates both versions:

- (46) VII. + VIII. *Either omission before subject or omission before object, with definite/indefinite article*

VII- A: *Def-Sub:YES, Obj:NO* DE KRAB VEROVERT
NOORDZEE
The crab conquers Northsea.

VII- B: *Def-Sub:NO, Obj:YES* KRAB VEROVERT DE
NOORDZEE
The crab conquers Northsea.

VIII- A: *Ind-Sub:YES, Obj:NO* UNA BOMBA DISTRUGGE
VETRINA
A bomb destroys shop-window

VIII- B: *Ind-Sub:NO, Obj:YES* BOMBA DISTRUGGE UNA
VETRINA
Bomb destroys a shop-window

Prediction

Stowell (1999) claimed that article omission from the direct object is impossible if the article has not been omitted from the subject. Hence, headlines of the type VII-A and VIII-A should not be possible. And, in fact, I did not find any headline in my database of Italian and Dutch headlines that was not compatible with this claim. Therefore, the prediction for these two conditions in the experiment was that the participants would have a very strong preference for the versions (VII-B and VIII-B) in which the determiner was present in the object and

omitted from the subject, and that they would reject the A-versions of the headlines.

Results:

Table 12 shows the preferences for the article used in either the subject or the object noun, in the conditions with definite and indefinite articles.

Table 12 *Preferences for article used on either subject or object noun (forced choice) in Dutch and Italian.*

Condition	Preference for article used in:	Dutch	Italian
DEF	Initial subject (VII-A)	18,0	38,2
DEF	Internal object (VII-B)	72,4	42,4
IND	Initial subject (VIII-A)	16,8	14,7
IND	Internal object (VIII-B)	65,2	71,8

I mentioned earlier that participants did have three preference choices in the questionnaire: *version with article* – *version without article* – *no preference*. In most conditions the number of *no-preference* responses was fairly low (0 – 2% of the responses). In the conditions discussed in this section however the rate of *no-preference* responses was higher than in the other conditions, which shows that participants were less certain about their preference.⁴⁷

In Dutch the difference between the preferences for versions A and versions B (see Table 12) is significant. Participants have a stronger preference for the B versions, both in the condition with definite articles ($F(49,1) = 61.220$, $p < .001$) and in the condition with indefinite articles ($F(49,1) = 66.371$, $p < .001$). At a first glance this means that Dutch confirms the prediction made by Stowell. However it is important to note that Stowell's account makes a very strong claim. It predicts that

⁴⁷ For Dutch for 9,6% of the items in the condition with definite article and 18 % of the items in the condition with the indefinite article participants indicated that they had no preference for one of the two versions. For Italian for 19,4% of the items in the condition with definite article and 16,5% of the items in the condition with the indefinite article participants indicated that they had no preference for either of the two versions.

omission from the object will be *impossible* (because of structural constraints) if the article has not been omitted from the subject. Hence we should expect that the A-versions, with the article omitted from the object, would be rejected in (almost) all cases. This is not what we found. Although there is a stronger preference for the versions in which the article is omitted from the subject, the results also show optionality. The B-versions in which the article is omitted from the subject is preferred in 'only' 72,4 % of the test items with definite articles and 65,2 % of the test items with indefinite articles. This is too low for an absolute rule-based account. What we find is a 'preference' but not an absolute grammatical judgement, and Stowell's principle is claimed to be a grammatical principle. What is more, the high number of '*no – preference*' responses by the participants in the conditions in this section is not what would be expected with an absolute rule-based account.

Moreover, the results for Italian raise even more doubt about such rule-based account. In Italian only the version with the *indefinite* articles seems to support Stowell's prediction. Here we observe a stronger (but again, not absolute) preference for the B-versions.⁴⁸ But, in the condition with the *definite* articles in Italian we see no significant difference between the preferences for the two versions ($F(33,1) = .256$, $p = .616$). This means that Italian participants have no clear preference for one of the two versions, when definite articles are used. This finding is not compatible with Stowell's claim, which was expressed in the form of the following constraint:

- (47) "a [+R] NP cannot be c-commanded by a [-R] NP"
 where a NP without an article is marked [+R], and a NP with an article is marked [-R].

This condition is satisfied in sentence-initial position, because there a [+R] NP is not c-commanded clause internally. Stowell suggests that it is also satisfied in clause-internal position if no c-commanding [-R] NP is present in the sentence. However, on this account we would expect the Italian participants to have the same preferences as the Dutch participants. It is not clear how on Stowell's account we can explain the preferences of the Italian participants with the definite determiners in

⁴⁸ The difference between the preferences for the A and B version in Italian, in the condition with indefinite articles, is significant: $F(33,1) = 66.049$, $p < .001$.

this condition. Stowell would have to stipulate that the marking for [+/-R] is different for definite and indefinite articles, and, moreover that they are different for Dutch and Italian. But as far as I see there is no independent evidence for this claim.

Summary order of omission:

Table 13

Nr	Condition	Example	Prediction (based on Stowell's claim)	Results
VII- A def	Def-Sub: YES- Obj:NO	DE KRAB VEROVERT NOORDZEE The crab conquers Northsea	pref B versions > pref A versions, with very strong difference. Expected: strong rejection of A- versions	Dutch: pref B versions > pref A versions, but difference less strong as expected, no strong rejection of A versions. Italian: pref A versions = pref B versions
VII- B def	Def- Sub:NOObj : YES	KRAB VEROVERT DE NOORDZEE Crab conquers the Northsea		
VIII A ind	Ind- Sub:YES- Obj:NO	UNA BOMBA DISTRUGGE VETRINA A bomb destroys shop-window	pref A versions > pref B versions, with very strong difference. Expected: rejection of A- versions	Dutch and Italian: Pref B versions > pref A versions, but difference less strong as expected, no strong rejection of A versions.
VIII -B ind	Ind- Sub:NO- Obj:YES	BOMBA DISTRUGGE UNA VETRINA Bomb destroys a shop-window		

3.4.3.4 Omissions in Nouns in Isolation and Hanging Topic Constructions

The experiment also included ‘Noun Phrases in isolation’ (NI) and Hanging Topic (HT) Constructions. Examples (48) and (49) illustrate these conditions.

- | | | | |
|------|----|-----------|--|
| (48) | IX | NI- D NO | EERSTE NEDERLAAG VOOR
ORANJE
First defeat for Oranje (Dutch national
team |
| | | NI- D YES | DE EERSTE NEDERLAAG VOOR
ORANJE
The first defeat for Oranje |
| (49) | X | HT- D NO | SPARATORIA A NASSIRYA, FERITO
UN CARABINIERE
Shooting in Nassirya, a policeman injured |
| | | HT- D YES | UNA SPARATORIA A NASSIRYA,
FERITO UN CARABINIERE
A shooting in Nassirya, a policeman
injured. |

Prediction

Analysis of the database (see Table 5) showed a high omission percentage of articles in Nouns in Isolation constructions, in both Dutch and Italian (though somewhat less in Italian than in Dutch). The prediction therefore is that in both, Dutch and Italian, nouns-in-isolation headlines the version of the headline without the article will be preferred.

In the previous section we saw that Dutch newspapers never use Hanging Topic constructions. We also saw that in Italian we find a very high rate of article omission in these constructions (see Table 7). I argued that with a Cinque/Escobar type of account we could account for the high omission rate in Italian Hanging Topic constructions. Since I wanted to test the same conditions in Dutch and Italian, I decided to include the Hanging Topic condition in the Dutch version as well.

For both, Dutch and Italian I expected a very strong preference for omission.

Results

As Table 14 shows, we find a high preference for omission of the article in both conditions, in both languages.

Table 14 *Preferences for article omission in conditions Nouns in Isolation and Hanging Topics*

Condition	Dutch	Italian	One way Anova Dutch-Italian
Nouns in Isolation	82,6	83,8	$F(1,82) = 0.008,$ $p=.9292$
Hanging Topics	73,2	87,3	$F(1,82) = 2.640,$ $p=.1042$

In the analysis of the headlines in Dutch and Italian newspapers I found the following pattern of omission (see tables (6) and (7), repeated in Table 15.

Table 15 *Omission of articles in non-verbal headlines in Dutch and Italian newspapers*

Headline type	Dutch	Italian
Nouns in isolation	83,13	50,83
Hanging Topics	-	91,7

If we compare this with the results of the experiment, we observe a difference. In the experiment in these conditions the Italian participants had a stronger preference for omission than the omission rate observed in the Italian headlines database. Again, this may be an effect of the construction of the test items. When minimal pairs are used, the difference between article used or omitted becomes very prominent, far more than it will be in the actual process of headline-composition. Therefore, in such an experiment the role of the article may very well be more important for the preferences for omission or use than it is in the actual production process of a headline. And it is of course not a priori clear that we should find exactly the same rates in the experiment and the database. The experiment is about preferences for omission, the database shows what is actually omitted. These are two different aspects. What is important however, is that there is a similarity in pattern. This similarity in pattern suggests the existence of a factor that influences

both, the preferences for omission of the readers and the actual omissions found in the database.

Summary Noun Phrases in Isolation and Hanging Topic Constructions

Table 16 *Noun Phrases in Isolation and Hanging Topic Constructions.*

Condition	Example	Prediction (on basis of database analysis)	Results
NI-Def	EERSTE NEDERLAAG VOOR ORANJE First defeat for Oranje (Dutch national team)	Strong preference for omission, in both, Dutch and Italian	as predicted
NI-Indef	UNA NUOVA STRAGE A ISTANBUL A shooting in Nassirya, a policeman injured		
HT-Def	DE MEDIAWET, REGERING IN GROTE PROBLEMEN Television law, government in big problems	Strong preference for omission in Italian. No Hanging Topics in Dutch headlines database.	as predicted
HT-Indef	SPARATORIA A NASSIRYA, FERITO UN CARABINIERE A shooting in Nassirya, a policeman injured		

3.4.3. Summary results

Let me start with a very important general observation on the results of the experiment, and on what they tell us about the interpretation of the readers' preferences in my study. The fact is that, playing the devil's advocate, one could make the claim that the results I found in the experiment do not reflect readers' intuitions on what makes a good headline, but just reflect what type of headline they find in their national newspapers. Hence, Italian readers, for example, will just prefer the versions with the article because that is what headlines in the Italian newspapers look like. The results of the experiment show that this argumentation cannot be correct. One of the counterarguments against this claim is that it then becomes very difficult to account for differences in judgements depending on for example the position of the noun, the absence or presence of finiteness or the type of determiner. Why would, in the Italian experiment, the preference for omission of the article before a subject noun be 80,6% if the omitted article is an indefinite article and only 21,2 % if the omitted article is a definite article (see Table 10)? After all, it does not seem plausible to argue that readers keep track of omissions in headlines in a detailed way, recording omissions of definite and indefinite articles or omissions in finite sentences and non-finite sentences separately.

Another counterargument is offered exactly by the results in the conditions with Noun Phrases in Isolation and Hanging Topics (see section 3.4. 3.4), where Italian participants showed a high preference for omission of the article. If Italian readers would simply prefer headlines in which articles are used, because that is the general pattern they observe in their newspapers, then we would find this same preference also in headlines with nouns in isolation and hanging topics, contrary to fact. Of course, one could argue that Italian readers prefer the headlines with omissions in the conditions with nouns in isolation and hanging topics because that is the specific pattern they observe in their newspapers. There are at least two reasons why this cannot be correct. In the first place, as I argued earlier, it is very unlikely that readers keep statistical track of in which type of constructions articles are omitted in their

newspapers and in which type of constructions articles are not omitted. Furthermore, even if they did keep a statistical track, this is not reflected by their preferences, as these diverge from the omission pattern observed in the actual headlines. In the database of actual headlines in the 'NPs in isolation' constructions articles were omitted in 50,83 % of the cases (see section 3.3.4). Hence, if the readers' judgements were a reflection of the pattern they observe in their newspapers, we would expect no strong preference for either of the two versions. However, there was a strong preference for omission, differently from the pattern found in the actual Italian headlines.

This means that readers do not base their judgements on a kind of template of stereotype headlines in Italian (or Dutch), which would indicate that, for example, articles cannot be omitted in Italian headlines. They base their judgements on intuitions about their language. The crucial question then is what is the nature of these 'intuitions'? Is it intuitive knowledge about the structure of their language or is it intuitive knowledge about the processing effort of certain constructions in their language? If it is intuitive knowledge about the structure of the language, it will be hard to account for the crosslinguistic differences in omission and preferences in omission. It will also be hard to account for the fact that the observed preferences are never 'absolute' preferences for one of the versions. If the preferences are based on structural motivations we will not expect to find this optionality, we will expect to find clear rejections of one of the two versions.

Headlines database and experiment compared

In our comparison of the headlines-database and the results of the headlines-experiment we find similarities on a number of aspects, but also differences.

Let me first summarize the similar patterns of omissions or preferences for omission:

- In (almost) all experimental conditions and in all the contexts distinguished in the database in Dutch the preference for omission is stronger than in Italian.
- Finiteness effect: in both database and experiment we find a stronger preference for omission if the verb is a non-finite verb. This effect of finiteness is stronger in Italian than in Dutch.
- Effect of position: in both database and experiment we find a higher omission rate from sentence-initial than on sentence-internal

position. In the Italian experiment there was no position effect in the experimental conditions with definite determiners

Let us then look at the *differences* between the experimental results and the findings in the database:

- Finiteness effect (section 3.4.3.1): the Italian participants' preference for omission in the experiment is higher than the omission attested in the headlines. I argued that this difference is caused by the construction of the test items: minimal pairs, different only in use or omission of the article. This necessarily pronounces the difference between presence or absence of the article. Headline writers, when composing the headline, have to cope with far more varying factors, in which use or absence of the article is only one of the elements that can be used to make the headline as informative as possible with a limited amount of processing effort.
- 'NPs in Isolation': In the experiment the Italian participants had a stronger preference for omission than the omission attested in the headlines.

Furthermore the experiment revealed some interesting aspects on the preferences of omission that, because of practical reasons, could not be attested in the database:

- One interesting aspect of the preferences for omission for which no evidence was found in the database, because of the absence of the necessary contexts, concerns the order effect of articles omitted (see section 3.4.3.3). Stowell (1999) claimed that article omission from the object is impossible if the article has not been omitted from the subject as well. The experimental results for Dutch, and for Italian with indefinite articles, seem to provide some support for Stowell's prediction, at least the results point towards the direction predicted by Stowell. However, this preference still was optional, there was no absolute grammaticality judgement, as would be expected in a rule based account. The experimental results for Italian with definite articles are not compatible with Stowell's prediction.
- Another aspect of the preferences for omission that could not be attested in the database concerns the effect of determiner type (definite versus indefinite determiners). In Italian we found a stronger preference for omission of indefinite articles than definite articles. In the Dutch experiment no difference is attested. This

effect could not be calculated for the database, since in the database it was not always possible to determine what type of article had been omitted. In a fairly high number of cases both article types were possible.

3.5 Article omission in Child Speech

3.5.1 Introduction

Let us now look at article omission by another category of speakers: language acquiring children. It is a well-known fact that in their early productions children omit articles, like they omit other functional categories. And this happens in spite of the fact that in the input articles occur very frequently, just like other functional categories. Of course it could be argued that the reason for this omission of functional categories in child speech is that they are 'less meaningful' than lexical categories, and hence, can be omitted more easily. However, if this argumentation is correct we would not expect crosslinguistic differences in omission patterns, since articles are generally less meaningful than lexical categories in all languages.

In this section I will present data of a study on article omission in Dutch and Italian child speech. In the next section I will compare these data with the data on article omission in Dutch and Italian headlines. A part of this study was performed as a joint project with Maria Teresa Guasti and Anna Gavarrò (Guasti et al., 2004, 2008). In our study we used spontaneous, longitudinal data, coming from the CHILDES data base (MacWhinney and Snow, 1985). We decided to use longitudinal, spontaneous data and not data from a cross-sectional experimental setting for a number of reasons. First, longitudinal data show whether there are any developmental patterns emerging over time and provide insight into the developmental stages individual children pass through to reach the adult target. This pattern cannot be observed in cross-sectional experiments. Secondly, spontaneous data show the extent to which article omission occurs in natural child speech. A cross-sectional experimental context gives insight in what several children are doing at a given time in a specific setting, but it does not give insight in children's use of articles in natural speech situations. Thirdly, in order to come to

an understanding of the full development process of article acquisition, we have to start with examining language production of the child at a very young age. But language experiments for language production under the age of 2;6 are problematic (Thornton, 1996) and therefore it would not be possible to come to a good understanding of the early stages of the article acquisition process using experimental data. In this study I want to propose a model of language development that can account for article omission and crosslinguistic differences, hence it is important to have a good understanding of what the complete pattern of language development in natural child speech looks like, starting from a very early age.

To come to broader generalizations of what takes place during the language (article) acquisition process we used data of four Dutch and four Italian children. The structure of this section is as follows: I will start with a description of how the data were collected and according to which points of view they were analyzed and explain how the children were grouped into developmental stages. I will then discuss this analysis, and compare these results with the results found in the headlines analysis and experiment. Furthermore, I will discuss previous accounts of article omission in child speech and discuss my concerns with these previous accounts. I will argue that, if we want to come to a model that can account for the observed patterns of omission as well as the crosslinguistic differences, we will need an account based on processing considerations.

3.5.2 Method

Subjects and Data

I analyzed data from four Dutch (Abel, Peter, Sara, Tom) and four Italian speaking children (Diana, Martina, Raffaello, Rosa), between 1;7 and 3;0 years old, at several data points evenly spread across the age range. All data used were taken from CHILDES (MacWhinney and Snow, 1985; Dutch: from the Groningen-corpus, Bol 1996, and the corpus of Van Kampen, Van Kampen 1994; Italian: Cipriani et al. 1989).⁴⁹ Table 17 gives a summary of the data used in the study, including information about the files used, the age range during the

⁴⁹ The Italian children are from Tuscany and speak the central variety of Italian.

period investigated, the MLU range and the number of article-requiring utterances, in which the article was either used or omitted. Care was taken to have roughly the same number of utterances for both languages. Since the lengths of the files differed, this means that in the case of shorter files the number of files examined was higher. All the files used included at least three contexts that required the use of an article.

The MLU was calculated for all children on the first 100 utterances of each file used (unless the number of utterances was less than 100, in which case it was based on the total file). The total number of utterances was 1790 for Italian and 2005 for Dutch.

Table 17 *Child Speech Data used in the study*

	Files	Age range	MLU range	Utterances with article-requiring nouns
Abel	111,201,203,205,207, 210, 211, 301	1;11-3;0	1;2-3.7	487
Peter	200, 202, 204, 205, 207, 208	1;11-2;08	1.8-3.5	537
Sara	2, 7,11,17,20, 27	1;7-2;11	1.2-3.0	428
Tom	110, 202, 205, 206, 209, 210, 301	1;10-3;0	1.2-3.1	553
Diana	2-10	1;10-2;6	2.0-4.5	457
Martina	2, 4, 6, 8, 11, 13, 16	1;7-2;7	1.2-2.7	500
Raffaello	3, 5,7, 9, 10, 12, 13, 14, 16	1;9-2;11	1.3-2.7	405
Rosa	5,7,10,12, 13,14,16,18, 20, 21	2;0-2;12	1.4-2.8	398

The ages of the children during the period investigated were roughly the same in both languages (Mann Whitney, $z=-.901$, $p=.367$; age in months: Italian $M=27.7$ $SD=4.8$, age range: 19 to 36, Dutch: $M= 28.8$ $SD=4.8$, age range: 19 to 36). Thus differences could not be attributed to the fact that some children are older than others. Eliminated from the utterances in the files were unclear sentences, immediate repetitions of the same sentence, clear imitations of adult input, idiomatic expressions, rimes and songs, and routine sentences. The relevant utterances were analyzed by hand.

In my analysis I made a distinction between different types of utterances. The goal was to come to a comparison of the data on omission in child speech with the data on omission in headlines. Therefore the division of utterance types was, as far as possible, the same

as the one used for types of headlines. For child speech I distinguished the following categories:

Noun Phrases in isolation

These were Noun Phrases used without phrasal context, used (also by adults) for example in lists or in naming an object in answer to a question, as for example in (50) and (51).

- (50) N: Ja, wat is dat dan? [Tom, 11011, 571]
 What is that?
 T: paard
 horse
 N: paard hè?
 horse, yes?
 C: en wat zit hier nou in de boom?
 And what is here in the tree?
 T: ap (=appel)
 Apple
- (51) M: E questa? Lo sai. [Rosa, 10, 692]
 And this? You know.
 R: Lunca (=mucca).
 Cow
 M: Una mucca!
 Hai visto, te lo ricordi.
 A cow! You see, you remember it.
 M: Poi.
 And next.
 R: Poi, cane.
 Then, dog.

But in child speech 'nouns in isolations' are also used as substitutes for complete phrases in situations where adults would not do so, as in the following examples:

- (52) a. G: Wat een leuk varkentje is dat hè? [Abel, 20302, 77]
 what a nice pig isn't it?
 A: pa...paard
 ho...horse
 G: Dat is een paard, ja
 that is a horse, yes
- b. J: Wat is ze aan het doen? [Sara, 09, 324]
 What is she doing here?
 S: B(l)oe(me)tje
 flower
 J: Bloemetjes aan het plukken, ja!
 Picking flowers, right!

The reason for distinguishing between nouns in isolation and more-word utterances is that in the case of nouns in isolation it is sometimes difficult to decide whether articles are needed or not. As we saw in the examples (50) and (51), in naming an object or in lists and in answers to questions an article is not always necessary. Making a distinction between these contexts and non-adultlike contexts of use of nouns in isolation is not always feasible. Thus it is possible that with nouns in isolation omission of articles would be overestimated, and that is why omissions before nouns in isolation were counted separately.

Non-verbal utterances

These were constructions with nouns in a phrasal context in which the verb was omitted, like for example:

- (53) a. Hond niet lekker bad. [Sara 09, 432]
 Dog not nice bath.
- b. F: Wat heb je daar nou weer? [Peter 20822, 488]
 What have you got there?
 P: Peter lucifertje.
 Peter match.
- c. Quetto poi qui un pentolone. [Diana 10, 618]
 This then here a pan.

- d. Ora pone (=sapone). [Rosa 10, 51]
Now soap.

Verbal utterances with a non-finite verb

These were constructions with nouns in a phrasal context in which a non-finite verb was used.

- (54) a. ja, zo blok gooien [Abel 20506, 386]
yes, throw block like that
- b. kraanwagen rijden [Peter 20007, 69]
lorry drive
- c. kleurtjes kijken [Tom 20507, 577]
colors look
- d. boekje lezen (=lezen) [Sara 11,242]
book read

Verbal utterances with a finite verb

These were constructions with nouns in a phrasal context in which a finite verb was used.

- (55) a. P: schaapje gaat klein stukje lopen. [Peter 20529, 511]
Sheep is going for a little walk.
- b. S: (ko)nijntje is weg. [Sara 11, 117]
Rabbit is gone.
- c. A: ikke ga koekje eten.
[Abel 20506, 1275]
I am going to eat cookie.
- d. T: daar kan andere auto rijden. [Tom 20926, 407]
Other car can drive there.
- e. E poi si mette qui la tazza. [Diana 10, 608]
And then you put the cup here.

- f. Appetta ora metto la bambola sull'aroplano.
[Diana 10, 74]
And now I put the doll on the airplane
- g. Nonno corre. [Rosa 10, 389]
Grand-dad is running.

Utterances in which the noun was preceded by a preposition

These were constructions with nouns directly following a preposition:

- (56) a. T: daar bloed op de rails. [Tom 20926, 801]
There blood on the rails.
- b. A: meer (=meneer) Gerard gaan pelen (=spelen) met de
blokken. [Abel 20729, 446]
Mister Gerard going to play with the blocks.
- c. S: jij on(der) tafel. [Sarah 17, 409]
You under table.
- d. P: in gaatje. [Peter 20203, 1461]
In hole.
- e. Io lo metto nel fligo (=frigo). [Diana 10, 294]
I will put it in the refrigerator.
- f. Io so'venuta a giocale colle babbi (=Barbie).
[Diana 10, 809]
I have come to play with the Barbies.

All nouns that in the normal adult language would have required the use of an article but were produced without an article were counted as omissions. For the Dutch data this means that plural and mass nouns without an article were considered ungrammatical only in those contexts in which the use of an article was obligatory in the target language. For Italian, plural and mass nouns without an article were considered ungrammatical unless they occurred as direct object or after a preposition. Copular sentences in which the distinction between argument or predicate was impossible to draw (e.g. in sentences like

‘This tree is a maple’ or ‘John is a doctor’) were not included in the counts. The reason for this is that predicates in these constructions can be used without an article both in Dutch as well as in Italian.

3.5.3 Results and Analysis

3.5.3.1 Introduction

I first analysed whether there were differences in article omission in both languages, collapsing together data for the whole period investigated for all more-word utterance types.

Table 18 shows the overall omission rate: the total number of article-requiring nouns used without an article divided by the total number of article-requiring nouns. A significant difference was found between Italian and Dutch, analogous to the results of the headlines, also in child speech we find more omissions of articles in Dutch than in Italian (Mann Whitney $z=-3.012$, $p=.003$).

Table 18 *Overall omission rate: mean and standard deviation in whole period and all contexts under investigation.*

	Overall omission rate	
	M	SD
Dutch	62.71	29.63
Italian	38.17	31.85

Since we want to come to a model of child language development, we need insight in the developmental patterns of children in both languages. In order to investigate the development of children’s omissions of articles, their performance has to be matched based on an independent measure. Age is not a reliable measure of linguistic development since children at the same age may be more or less advanced (Bates et al., 1995). Differences in morphology and lexicon make it problematic to use MLU as a measure of linguistic development in crosslinguistic studies. Another drawback in the use of MLU as a measure of linguistic development concerns the fact that it does not evaluate qualitative aspects of syntactic development (see Klee & Fitzgerald 1985 for a discussion). Following Guasti et al. (2008), in this study I used as an independent measure of linguistic development the rate of verbal utterances (VU). This rate expresses the total number of the child’s utterances containing a verb divided by the total number of utterances

(Valian, 1991). The rate of VU was calculated for each individual file. On the basis of this calculation the data of article use and omission were divided into two classes: the first class includes observations obtained when children's rate of VU was between .03-.29 and the second one when it was between .30 and .60. (VU ranges between .05-54, M=30.54, SD=15.75 in Dutch; range between .03-.60, M=27.00 SD=14.75 in Italian). Table 19 reports the means and SD of article omission at the two different stages of linguistic development.

Table 19 Omission in different periods of linguistic development as measured by rate of VU

	Stage 1 0.03-0.29 VU		Stage 2 0.3-0.6 VU		Mann-Whitney stage1-stage 2 compared
	M	SD	M	SD	
Italian	58.29	29.5	16.9	17.18	$z=-4.097, p < .0001$
Dutch	87.05	14.84	41.6	21.84	$z=-4.372, p < .0001$
Mann-Whitney Italian-Dutch	$z = -2.528, p = .011$		$z = -3.395, p < .0001$		

Table 19 shows that there are differences between Italian and Dutch. In both stages omission of articles is higher in Dutch than in Italian. In the second stage Italian children have almost stopped omitting articles and perform almost adult-like, while Dutch children still omit articles quite often. The Mann-Whitney test shows that in both stages the differences between Dutch and Italian are significant

The data show that article omission in child speech is optional. Sometimes the article is used, sometimes it is omitted. In the omission data on headlines this same optionality was found. This means that, as was the case for omission in headlines, we can ask the following questions for omission in child speech: Where and when are articles used? What is it that triggers the production of an article? Or, alternatively: What is it that suppresses the production of an article? In order to find an answer to these questions in the analysis of omission in headlines, I made a distinction in different types of linguistic contexts in which I examined the omission of articles. I will make a similar distinction in linguistic contexts for the omission data in child speech to

find out whether the factors influencing omission in headlines also influence omission in child speech.

3.5.3.2 Relation with finiteness

In headlines omission of articles is influenced by the presence or absence of a finite verb. We find more omissions of articles in headlines (and a stronger preference for omission by the readers) if no finite verb is present in the headline. We saw that the effect of finiteness was stronger in Italian than in Dutch. Let us take a look at the data on omission in child speech in utterances with 'phrasal context with finite verb' versus utterances with 'phrasal contexts without a finite verb'. In the analysis of headlines the headlines with no finite verb were either 'headlines with a non-finite lexical verb' or 'headlines in which the lexical verb was omitted'. In section 3.3.2.1 I showed that the category of 'headlines in which the lexical verb was omitted' did not have a high rate of occurrence, as these types of headlines were only possible with a limited type of verbs in a limited range of contexts. I also argued that the reason for treating these headlines on a par with the headlines in which a non-finite lexical verb was used was the fact that, in spite of the omission of the lexical verb, it was still possible to distinguish a sentence-initial from a sentence-internal position, and that this characteristic made the headlines without lexical verb more similar to verbal headlines than to nouns in isolation, where no distinction in position could be made.

In child speech, especially in the earlier files, the situation is slightly different because of the fact that children are more permissive with respect to omission of lexical verbs than headline writers. For this reason I made a comparison of the child data of omission in phrasal utterances with a non-finite verb and phrasal utterances without a verb. For Dutch this comparison revealed no significant difference. For Italian I hardly found any utterances with a non-finite verb. In Italian 2,5% of the total number of verbal utterances were utterances with a non-finite verb. In Dutch 18% of the total number of verbal utterances were utterances with a non-finite verb. Phrasal utterances without verb are present in the speech production of Italian children: in Italian 25,8% of the phrasal utterances are without a verb, in Dutch: 23,0 %. So, for Dutch both types of utterances with no finite verb are present in the speech output, but there is no difference between omission rates in both types of utterances. In Italian only one type of utterance with no finite verb is

present in the speech output (the phrasal utterance with no verb). I therefore decided to treat the two types of utterances as one type, as I did in the analysis of headlines.

Table 20 for Dutch and Table 21 for Italian show for the individual children, and for the four children together, the number of article-requiring nouns in which an article was either used (DP) or omitted (NP), and the omission rate (the number of article requiring nouns in which an article was omitted divided by the number of article requiring nouns) in utterances with no finite verb and utterances with a finite verb. In this analysis the so-called preposition contexts, the nouns directly following a preposition, were excluded. Hence, the same distinction as was made for the headlines.

Table 20 *Article use and omission in finite and no-finite verb contexts in Dutch, both developmental stages mapped together*

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

	Peter		Tom		Abel		Sara		Total all children	
	Fin V	No Fin V	Fin V	No Fin V	Fin V	NoFin V	Fin V	No FinV	Fin V	No Fin V
DP	131	12	62	25	63	39	25	10	281	86
NP	112	79	31	45	38	64	45	25	226	213
Om. Rate (%)	46,1	86,8	33,3	64,3	37,6	62,1	64,3	71,4	44,6	71,2
	$\chi^2=43.195$		$\chi^2=14.157$		$\chi^2=11.295$		$\chi^2=.0263$		$\chi^2=14.847$	
	p<.0001		p=.0002		p=.0008		p=.6084		p<.0001	

Table 21 *Article use and omission in finite and no-finite verb contexts in Italian, both developmental stages mapped together*

	Rosa		Raffaello		Diana		Martina		Total all children	
	FinV	No Fin V	Fin V	No Fin V	Fin V	No Fin V	Fin V	NoF nV	Fin V	No Fin V
DP	145	56	141	26	209	28	116	16	611	126
NP	65	43	25	7	34	5	58	20	182	75
Om. Rate (%)	30,9	43,4	15,1	21,2	14,0	15,2	33,3	55,6	23,0	37,3
	$\chi^2=4.087$		$\chi^2=0.383$		$\chi^2=0.032$		$\chi^2=5.393$		$\chi^2=16.514$	
	p=.0434		p=.5358		p=.8576		p=.0202		p<.0001	

Looking at the level of children individually, we see that the correlation between finiteness and article omission is in Dutch statistically significant for three of the four children, and in Italian for two of the four children.⁵⁰

Let me compare these findings with the results found in the headlines. Recall that in both the headlines-experiment and the database in both Dutch and Italian, the difference between omissions in headlines containing a finite verb and headlines in which the finite verb was omitted was significant. Participants showed a higher preference for omission if the finite verb was omitted. Thus we can conclude that the omission pattern of some of the children is similar to the pattern found in headlines, but not for all children. There are individual differences. One of the reasons for the differences between the headlines and the child speech data may be that the child speech data in Tables 20 and 21 present the omission data of all files, in all stages. However, it may very well be the case that we do find a stronger finiteness effect, in more children, if we examine the omission data of the developmental stages separately. The Italian data in particular suggest that this may be a plausible assumption. As Table 21 shows the finiteness effect is less strong for the two Italian children that perform most ‘adultlike’ (Raffaello and Diana) than for the other two children. Hence there may

⁵⁰ In the remainder of this study the statistically significant results are shaded in grey in the tables.

be a (stronger) finiteness effect in the earlier developmental stages. Therefore, I made an analysis of article omission in utterances with no finite verb and utterances with a finite verb for developmental stage 1 and developmental stage 2. Tables 22 to 25 show the results.

Table 22 *Article use and omission in finite and no-finite verb contexts in Dutch, developmental stage 1*

	Peter		Tom		Abel		Sara		Total all children	
	Fin V	No FinV	Fin V	No FinV	FinV	No Fin V	Fin V	No Fin V	Fin V	No Fin V
DP	1	0	1	4	4	12	1	3	7	19
NP	23	73	7	24	20	50	10	11	60	158
Om. rate	95,8	100	87,5	85,7	83,3	80,6	90,1	78,6	89,5	89,3
	$\chi^2=0.346$		$\chi^2=0.017$		$\chi^2=.0.083$		$\chi^2=0.082$		$\chi^2=.0004$	
	p=.5563		p=.8975		p=.7739		p=.7751		p=.9483	

Table 23 *Article use and omission in finite and no-finite verb contexts in Italian, developmental stage 1*

	Rosa		Raffaello		Diana		Martina		Total all children	
	Fin V	No Fin V	Fin V	No Fin V	Fin V	No Fin V	Fin V	No Fin V	Fin V	No Fin V
DP	28	17	9	5	24	15	31	5	92	42
NP	26	26	16	5	15	3	19	13	76	47
Om. rate	48,1	60,5	64,0	50,0	38,5	16,7	38,0	72,2	45,2	52,9
	$\chi^2=1.007$		$\chi^2=0.0146$		$\chi^2=1.793$		$\chi^2= 4.924$		$\chi^2=1.050$	
	p=.3156		p=.7025		p=.1806		p=.0265		p=.3055	

Table 24 *Article use and omission in finite and no-finite verb contexts in Dutch, developmental stage 2*

	Peter		Tom		Abel		Sara		Total all children	
	FinV	No Fin V	FinV	No Fin V	FinV	No Fin V	FinV	No Fin V	FinV	No Fin V
DP	130	12	61	21	59	27	24	7	274	67
NP	89	6	24	21	18	14	35	14	166	55
Om. rate	40,6	33,3	28,2	50,0	23,3	34,1	59,3	66,7	37,7	45,1
	$\chi^2=0.128$		$\chi^2= 4.908$		$\chi^2=1.072$		$\chi^2=.0.111$		$\chi^2=1.860$	
	p=.7205		p=.0267		p=.3004		p=.7395		p=.1717	

Table 25 *Article use and omission in finite and no-finite verb contexts in Italian, developmental stage 2*

	Rosa		Raffaello		Diana		Martina		Total all children	
	Fin V	No Fin V	Fin V	No Fin V	Fin V	No Fin V	Fin V	No Fin V	Fin V	No Fin V
DP	117	39	132	21	185	13	85	11	519	84
NP	39	17	9	2	19	2	39	7	106	28
Om. rate	25,0	30,3	6,4	8,7	9,3	13,3	31,4	38,9	17,0	25,0
	$\chi^2=0.364$		$\chi^2=.0.169$		$\chi^2=.003$		$\chi^2=.130$		$\chi^2=3.604$	
	p=.5463		p=.6810		p=.9553		p=.7184		p=.0576	

Surprisingly, in both languages the finiteness effect disappears (with the exception of one child) if we divide data over developmental stages. In both languages children omit less in stage 2 than in stage 1, but, differently from what was found for all the developmental stages mapped together, there is no finiteness effect in the stages viewed

separately.⁵¹ Clearly the question arises why we find a finiteness effect if we examine all stages together, but not if we examine the stages separately. In order to account for this pattern, let us first take a look at the rate of finite verbal utterances in both stages, as illustrated in Table 26.

Table 26 *Finite verbal utterances expressed as a percentage of the total number of utterances, in Dutch and Italian, in stage 1 and stage 2*

	Rate of finite verbal utterances		
	Stage 1	Stage 2	
Dutch	27,5	78,3	$\chi^2= 49.160$, $p<.0001$, Fisher's exact: $p<.0001$
Italian	65,4	84,8	$\chi^2= 5.052$, $p=.0246$
	$\chi^2= 25.098$, $p<.0001$, Fisher's exact: $p<.0001$	$\chi^2= 0.833$, $p=.3615$	

The development in the rate of finite verbal utterances shows an increase in the use of finite utterances from stage 1 to stage 2. If we compare this development with the development of article use (see Tables 21 to 24), we find that from stage 1 to stage 2 the use of articles increases as well. Hence, in stage 2 children use more finite utterances and, at the same time, they also use more articles. Note however, that children's use of articles is not directly influenced by the presence of a finite verb in the utterance, they also use more articles in non-finite verbal utterances. Use of finiteness and use of articles are related in the time course of the children's linguistic development, but use of finiteness by itself does not influence article use. Children do not use more articles because of the presence of a finite verb in the sentence. In stage 1 and stage 2 we do not find significant differences between omission percentages in utterances with finite verbs and utterances with no finite verb.

⁵¹ Also on a division in developmental stages based on MLU or different content words used no finiteness effect was present in the developmental stages viewed separately.

Let us now turn to the question how we can account for the (misleading) finiteness effect found in the data of the developmental stages mapped together. From stage 1 to stage 2 the number of finite utterances increases and the number of articles produced increases as well. Therefore there is a diachronic mismatch between the categories ‘finite utterances’ and ‘no-finite utterances’ in Tables 20 and 21 (all developmental stages mapped together). The ‘finite utterances’ for the major part consist of utterances produced in stage 2, while the ‘no-finite utterances’ for the major part consist of utterances produced in stage 1. This results in an apparent finiteness effect, but this is a misleading one, since if we look at the results in the stages separately we see no difference between omission rates in finite and no-finite verbal utterances. Use of finiteness and use of articles are only related in the time course of children’s linguistic development, but not inter-related. Use of finiteness does not directly affect the use of articles. What the results really show is that the more likely a sentence is to contain a finite verb, the more likely it is that an article is produced.

These findings are in conformity with the results reported by Van der Velde (2004), in a study on article omission in French and Dutch acquiring children aged between 3 and 6 years old. Van der Velde too made a distinction in developmental stages, and then found that there was no direct relation between use of finiteness and use of articles.

Let us now consider the differences in finite and no-finite verbal utterances, focusing on the position of the noun.

3.5.3.3 Division based on sentence position of the noun

It has been claimed in several studies on child language that there is a difference in omission pattern depending on the position of the noun: sentence-initial or sentence-internal. However, the studies differ with respect to the direction of this difference. Gerken (1991) claims that children omit more determiners in initial position than in internal position, but Baauw, de Roo & Avrutin (2002) observed a pattern in the opposite direction: more omissions from sentence-internal position by one of the children under investigation. A similar observation has been made for child German by Schoenenberger, Penner & Weissenborn (1997). I investigated whether our data confirmed the existence of a difference in omission pattern depending on the position of the noun,

what the direction (initial/internal) of this difference was and whether there were cross-linguistic differences in the observed patterns.

I analyzed the position effect for finite verbal utterances and no-finite verbal utterances separately. Let me start with the finite verbal utterances. Tables 27 and 28 show the omissions from sentence-initial and sentence-internal position of the individual children in all developmental stages mapped together, Tables 29 to 32 show the results of the individual children in stage 1 and stage 2.

Finite verbal utterances, all developmental stages mapped together:

Table 27 *Article use and omission from sentence-initial and sentence-internal position in finite verbal utterances Dutch, both developmental stages mapped together*

	Peter		Tom		Abel		Sara		Total all children	
	Ini	Int	Ini	Int	Ini	Int	Ini	Int	Ini	Int
DP	44	87	7	55	9	54	4	25	64	221
NP	72	40	12	19	10	26	17	28	111	113
Om.rate	62,1	31,4	63,1	25,6	52,6	32,5	80,9	52,8	63,4	33,8
	$\chi^2 = 21.592$		$\chi^2 = 7.946$		$\chi^2 = 1.889$		$\chi^2 = 3.881$		$\chi^2 = 39.629$	
	p < .0001		p = .0048		p = .1693		p = .0488		p < .0001	

Table 28 *Article use and omission from sentence-initial and sentence-internal position in finite verbal utterances in Italian, both developmental stages mapped together*

	Rosa		Raffaello		Diana		Martina		Total all children	
	Ini	Int	Ini	Int	Ini	Int	Ini	Int	Ini	Int
DP	26	119	7	134	16	193	9	107	58	553
NP	19	46	1	24	2	32	5	53	27	155
Om rate	42,2	27,9	12,5	15,2	11,1	14,2	35,7	33,1	31,8	21,9
	$\chi^2 = 2.766$		$\chi^2 = 0.043$		$\chi^2 = .000$		$\chi^2 = 0.039$		$\chi^2 = 3.643$	
	p = .0963		p = .8356		p = .9896		p = .8438		p = .0563	

Finite verbal utterances stage 1:**Table 29** *Article use and omission from sentence-initial and sentence-internal position in finite verbal utterances Dutch, stage 1*

	Peter		Tom		Abel		Sara		Total all children	
	Ini	Int	Ini	Int	Ini	Int	Ini	Int	Ini	Int
DP	0	1	0	1	0	4	0	1	0	7
NP	9	14	0	7	5	14	10	0	24	35
Om. rate	100	93,3	0	87,5	100	77,8	100	0	100	83,3
	$\chi^2 = .0626$				$\chi^2 = .0243$		$\chi^2 = .2.227$		$\chi^2 = 2.889$	
	p= .4288				p= .6221		p= 0.1356		p= .0892	

Table 30 *Article use and omission from sentence-initial and sentence-internal position in finite verbal utterances Italian, stage 1*

	Rosa		Raffaello		Diana		Martina		Total all children	
	Ini	Int	Ini	Int	Ini	Int	Ini	Int	Ini	Int
DP	4	24	4	5	5	19	2	29	15	77
NP	6	20	1	15	2	13	3	16	12	64
Om. rate	60	45,5	20	75,0	28,6	40,6	60	35,6	44,4	45,4
	$\chi^2 = .231$		$\chi^2 = 3.136$		$\chi^2 = .0243$		$\chi^2 = .340$		$\chi^2 = .008$	
	p=.6309		p=.0766		p=.8690		p=.5601		p= .9279	

Finite verbal utterances stage 2:**Table 31** Article use and omission from sentence-initial and sentence-internal position in finite verbal utterances Dutch, stage 2

	Peter		Tom		Abel		Sara		Total all children	
	Ini	Int	Ini	Int	Ini	Int	Ini	Int	Ini	Int
DP	44	86	7	54	9	50	4	24	64	214
NP	63	26	12	12	5	12	7	28	87	78
Om. rate	58,9	23,2	63,2	18,2	35,7	19,4	63,6	53,8	57,6	26,7
	$\chi^2 = 27.393$		$\chi^2 = 12.592$		$\chi^2 = .944$		$\chi^2 = .067$		$\chi^2 = 39.356$	
	p < .001		p = .0004		p = .3312		p = .7951		p < .001	

Table 32 Article use and omission from sentence-initial and sentence-internal position in finite verbal utterances Italian, stage 2

	Rosa		Raffaello		Diana		Martina		Total all children	
	Ini	Int	Ini	Int	Ini	Int	Ini	Int	Ini	Int
DP	22	95	3	129	11	174	7	78	43	476
NP	13	26	0	9	0	19	2	37	15	91
Om rate	37,1	21,5	0	6,5	0	9,8	22,2	32,2	25,9	16,0
	$\chi^2 = 2.763$		$\chi^2 = .209$		$\chi^2 = .313$		$\chi^2 = .061$		$\chi^2 = 2.934$	
	p = .0965		p = .6476		p = .5759		p = .8053		p = .0867	

The results in the no-finite utterances are:

No-finite verbal utterances, all developmental stages mapped together:**Table 33** Article use and omission from sentence-initial and sentence-internal position in no-finite utterances Dutch, both developmental stages mapped together

	Peter		Tom		Abel		Sara		Total all children	
	Ini	Int	Ini	Int	Ini	Int	Ini	Int	Ini	Int
DP	4	8	13	12	18	21	1	8	36	49
NP	40	39	25	21	35	29	14	11	114	100
Om. rate	90,9	83,0	65,8	63,6	66,0	58,0	93,3	57,9	76,0	67,1
	$\chi^2 = 0.652$		$\chi^2 = 0.036$		$\chi^2 = 0.406$		$\chi^2 = 3.741$		$\chi^2 = 2.481$	
	p = .4195		p = .8497		p = .5239		p = .0531		p = .1153	

Table 34 *Article use and omission from sentence-initial and sentence-internal position in no-finite utterances Italian, both developmental stages mapped together*

	Rosa		Raffaello		Diana		Martina		Total all children	
	Ini	Int	Ini	Int	Ini	Int	Ini	Int	Ini	Int
DP	16	40	5	21	10	18	11	5	42	84
NP	11	32	3	4	2	3	10	10	26	49
Om.rate	40,7	44,4	37,5	16,0	16,7	14,3	47,6	66,7	38,2	36,8
	$\chi^2 = 0.011$		$\chi^2 = 0.637$		$\chi^2 = 0.034$		$\chi^2 = 0.630$		$\chi^2 = 0.002$	
	p=.9176		p=.4249		p=.8544		p=.4274		p=.9688	

*No-finite utterances stage 1:***Table 35** *Article use and omission from sentence-initial and sentence-internal position in no-finite utterances Dutch, stage 1*

	Peter		Tom		Abel		Sara		Total all children	
	Ini	Int	Ini	Int	Ini	Int	Ini	Int	Ini	Int
DP	0	0	2	2	6	6	1	2	9	10
NP	38	35	13	12	24	26	8	3	83	76
Om. rate	100	100	86,7	85,7	80,0	81,25	88,9	60,0	90,2	88,4
	Fisher's exact p=1.000		$\chi^2 = .0006$ p= .9408		$\chi^2 = .0016$ p= .9009		Fisher's exact p=.5055		$\chi^2 = .0024$ p=.8764	

Table 36 *Article use and omission from sentence-initial and sentence-internal position in no-finite utterances Italian, stage 1*

	Rosa		Raffaello		Diana		Martina		Total all children	
	Ini	Int	Ini	Int	Ini	Int	Ini	Int	Ini	Int
DP	5	12	3	2	8	7	1	4	17	25
NP	7	19	3	2	1	2	6	7	17	30
Om. rate	58,3	61,3	50,0	50,0	11,1	22,2	85,7	63,6	50,0	54,6
	$\chi^2 = 0.032$		$\chi^2 = .0000$		$\chi^2 = 0.400$		$\chi^2 = 0.230$		$\chi^2 = 0.040$	
	p= .8588		p= 1.0000		p= .5271		p=.6314		p=.8428	

No-finite utterances stage 2:**Table 37** Article use and omission from sentence-initial and sentence-internal position in no-finite utterances Dutch, stage 2

	Peter		Tom		Abel		Sara		Total all children	
	Ini	Int	Ini	Int	Ini	Int	Ini	Int	Ini	Int
DP	4	8	11	10	12	15	0	6	27	39
NP	2	4	12	9	11	3	6	8	31	24
Om. rate	33,3	33,3	52,2	47,4	47,8	16,7	100	57,1	53,5	38,1
	$\chi^2 = 0.000$		$\chi^2 = 0.096$		$\chi^2 = 3.084$		$\chi^2 = 1.916$		$\chi^2 = 2.285$	
	p= 1.000		p= 0.7565		p= .0791		p= .1663		p=.1306	

Table 38 Article use and omission from sentence-initial and sentence-internal position in no-finite utterances Italian, stage 2

	Rosa		Raffaello		Diana		Martina		Total all children	
	Ini	Int	Ini	Int	Ini	Int	Ini	Int	Ini	Int
DP	11	28	2	19	2	11	10	1	25	59
NP	4	13	0	2	1	1	4	3	9	19
Om. rate	26,7	31,7	0	9,5	33,3	8,3	28,6	75	26,5	24,4
	$\chi^2 = 0.001$		$\chi^2 = 0.0209$		$\chi^2 = 0.036$		$\chi^2 = 1.206$		$\chi^2 = 0.056$	
	p=.9720		p= .6479		p=.8494		p=.2720		p= .8124	

Summarizing, in *finite* verbal utterances we find an effect of position (more omissions in sentence-initial position) in Dutch in stage 2. In Italian there is a tendency to omit less in sentence-initial position, but this difference is not significant. In *no-finite* verbal utterances we find no significant differences between the omission rates from sentence-initial and sentence-internal position, neither in Italian, nor in Dutch. Let us compare these findings with the results found in the headlines.

- In the headlines-experiment in Dutch the difference between omissions from sentence-initial and sentence-internal position was significant. Participants showed a higher preference for omission from sentence-initial position. In Italian the difference between omission from sentence-initial and sentence-internal

position was significant for the condition with the indefinite determiners, but not for definite determiners.

- In the database of headlines the omission rate was higher in sentence initial position in both languages.

Formulated differently, in

- Dutch: we found more omissions from initial than internal position in finite verbal utterances in child speech (stage 2), the experiment and the database, but not in non-finite verbal utterances in child speech.
- Italian: we found more omissions from initial position in the experiment (but only with indefinite determiners) and in the database, but not in child speech.

3.5.3.4 Division based on type of article

Several studies on the acquisition of articles in Dutch have shown that there are differences in the acquisition patterns of definite articles. Van der Velde (2004), for example, found in a production task with Dutch and French speaking 3, 4 and 6 year old children that Dutch-speaking children frequently use the non-neuter definite article *de* instead of the neuter definite article *het*, whereas gender errors are totally absent in French. Van der Velde suggests that the definite non-neuter article *de*, due to its lexical specification, is taken as a default form in early child Dutch.⁵² Zonneveld (1992) argued, on the basis of a production task with 4 to 7 years old Dutch children, that children systematically use '*de*' before words they do not know, and that they use '*de*' as the unmarked option (or as a kind of 'resort' option) before words of which they know the meaning, but not the gender.

In the spontaneous speech data of the children in our study we found that children more often use '*de*' instead of '*het*' than vice versa (in total we found 54 substitution errors of '*de*' instead of '*het*', and 3 substitution errors of '*het*' instead of '*de*'). But, there are other interesting patterns we can observe that cannot straightforwardly be accounted for by arguing

⁵² Van der Velde proposes that '*het*' can be used only in one context type, specified by the feature combination: [+def,-pl,+neuter], '*de*' can be used in two different context types, specified by the feature combinations [+def,-pl, - neuter] and [+def, +pl]. Since '*de*' can be used in different context types it is the less specified, or 'elsewhere' form, which is used by children as a default form before they acquire the specific forms.

that ‘*de*’ is used as a default form in early child speech. First, we investigated whether ‘*het*’ was omitted more often than ‘*de*’ in the spontaneous speech data of the children in our study. Table 39 illustrates the data on use and omission of ‘*de*’ and ‘*het*’ of the individual children (in verbal utterances, preposition contexts excluded)

Table 39 Omission and use of ‘*de*’ and ‘*het*’ in obligatory contexts in verbal utterances, preposition contexts excluded

	Peter		Tom		Abel		Sara		Total all children	
	De	Het	De	Het	De	Het	De	Het	De	Het
DP	84	27	36	0	21	6	10	0	151	33
NP	109	48	15	10	12	17	19	17	155	92
om. rate	56,5	64,0	29,4	100	36,4	73,9	65,5	100	50,6	73,6
	$\chi^2=.0969$		$\chi^2= 14.429$		$\chi^2= 6.242$		$\chi^2= 5.601$		$\chi^2=18.173$	
	p=.3249		p <.0001		p=.0126		p=.0180		p <.0001	

As the results in Table 39 show, three of the four children omit ‘*het*’ significantly more than ‘*de*’. In fact, two children did not at all use ‘*het*’ in verbal, non-preposition contexts. Since these two children did use ‘*het*’ in preposition contexts it is not the case that their article paradigm consists of only ‘*de*’ and ‘*een*’ and that they always use ‘*de*’ when a definite article is needed. Apparently, use of a preposition facilitates the use of the article.⁵³ Table 40 shows the use and omission of ‘*de*’ and ‘*het*’ in preposition contexts.

⁵³ We find less omissions of both, ‘*de*’ and ‘*het*’ in preposition contexts (table 40) than in non-preposition contexts (table 39), the difference is significant: ‘*de*’: $\chi^2 = 54.512$, df 1, $p <.0001$; ‘*het*’ $\chi^2 = 11.597$ df 1, $p = .0007$. Note that also in preposition contexts the difference in omission of ‘*de*’ and ‘*het*’ is significant.

Table 40 Omission and use of 'de' and 'het' in preposition contexts

	Peter		Tom		Abel		Sara		Total all children	
	De	Het	De	Het	De	Het	De	Het	De	Het
DP	39	15	78	5	28	4	21	3	166	27
NP	24	5	3	7	7	6	2	4	36	22
om. rate	38,1	25	3,7	58,3	20	60	8,7	57,1	17,8	44,9
	$\chi^2=.642$		$\chi^2= 27.060$		$\chi^2= 4.267$		$\chi^2= 5.136$		$\chi^2=14.783$	
	p=.4231		p <.0001		p=.0389		p=.0234		p <.0001	

Further it is important to note that the two children who do use 'het' in verbal contexts use it almost exclusively in sentence-internal position (86% of the occurrences of 'het' in verbal utterances were on nouns in sentence-internal position). In sentence-initial position in verbal utterances the article 'het' is almost always omitted, by all children. It is not a priori clear how to account for these findings by assuming that 'de' is used as a default form in child speech.

We can make another observation with respect to the use of 'het' by looking at the use of the articles before diminutive nouns. Dutch diminutives have the interesting property that regardless of the gender of the base noun from which they are derived, they always take the determiner 'het'. Zonneveld (1992) observed a very high rate of correct use of the article 'het' before diminutive nouns. In the same study children did have problems with the correct use of 'het' before non-diminutive nouns. Zonneveld argued that an explanation for these findings could be the fact that diminutives are used very frequently by young children. Following Scharlaekens (1980), he argued that the formation of the diminutive is one of the earliest acquired morphological 'skills'. This account would predict that in our spontaneous child speech data too we could find differences between the omission of the article 'het' before diminutive nouns and the omission of the article 'het' before normal, non-diminutive nouns. In a separate study, together with Phaff (Phaff, 2006), also based on analysis of spontaneous speech data in Childes-files, we examined whether there were differences in the omission of articles by children before diminutive nouns and non-diminutive nouns. Contrary to the findings of Zonneveld (1992) we did not find less, but significantly more problems with the production of the article 'het' before diminutive

nouns. We found that children omit the article ‘*het*’ more often before diminutive nouns than before non-diminutive neuter nouns (omissions of ‘*het*’ before diminutive nouns: 47, 2%, omission of ‘*het*’ before non-diminutive neuter gender nouns 29,6%, $\chi^2 = 19.633$, $p < .0001$).

Next we examined whether there was a difference between the omission rates of definite articles before diminutives of common gender nouns (*de hond* (the dog) = common gender, diminutive: *het hondje*) and definite articles before diminutives of neuter gender nouns (*het boek* (the book) = neuter gender, diminutive: *het boekje*). Table 41 shows the results

Table 41 Omission rate of definite articles before non-diminutive nouns and before diminutive nouns of common gender and neuter gender

	Normal, non diminutive nouns	Diminutive nouns
Common gender nouns	24,9 (<i>de hond</i>)	50,6 (<i>het hondje</i>)
Neuter gender nouns	29,6 (<i>het boek</i>)	37,6 (<i>het boekje</i>)
	$\chi^2 = 2.460$, $p = .1168$	$\chi^2 = 4.121$, $p = .0424$

The results in Table 41 show that producing the correct article ‘*het*’ before a common gender diminutive noun is significantly more problematic for children than producing the correct article ‘*het*’ before a neuter gender diminutive noun, hence ‘*het hondje*’ is more difficult than ‘*het boekje*’. It is not clear how to account for these findings with an account based on frequency of the diminutives in child speech. There is no reason to assume that common gender nouns occur less often as diminutives in child speech or in child directed speech than neuter gender nouns. If we look at the frequency distribution of common and gender neuter nouns in Dutch we would expect the opposite pattern. Common gender nouns occur far more frequently than neuter gender nouns.⁵⁴ Why would articles before diminutives of common gender nouns be more difficult to process than articles before diminutives of neuter gender nouns?

⁵⁴ According to the corpora in the ‘Corpus Gesproken Nederlands’ (see chapter 4, section 4.3, for more details) the frequency distribution of the Dutch articles is: ‘*de*’: 48%, ‘*het*’: 18%, ‘*een*’: 34%)

Moreover, the results of Phaff's study on child speech are in conformity with the findings of Schiller and Caramazza (2003) for Dutch adults. Schiller and Caramazza found a longer naming latency with Dutch speaking adults for naming the diminutive (including the article) of a common gender noun.⁵⁵ They argued that the results showed that the gender feature of the base noun plays is activated in the article selection process. This leads to a stronger competition effect in the case of a diminutive of a common gender base noun (where both, 'de' and 'het' are activated) than in the case of a diminutive of a neuter gender base noun (where only 'het' is activated). In the next chapter I will present a similar account, based on the effect of competition between articles, to explain the differences in omission patterns of 'de' and 'het' in Dutch child speech.

3.5.4 Results child speech summarized and compared to headlines

Let me summarize the results found in the analysis of the child speech data:

- Overall, in all analysed contexts and stages we find more omissions in Dutch than in Italian.
- When we look at the effects of finiteness in child speech: if we look at the data of all children, and all stages collapsed together there is an effect of finiteness. We find more omissions in no-finite phrases than in finite phrases. But if we look at the data of the individual children and if we split the data in developmental stages, the finiteness effect disappears. I have argued that finiteness and article use are related in the time course of linguistic development: if children use more finite verbs, they also use more articles. However, use of finiteness by itself does not directly influence article use: within the different developmental stages we find no difference in omission rate in finite and no-finite utterances, we only find differences between the developmental stages. It is not clear how to account for these

⁵⁵ Caramazza and Schiller only found a difference with the naming of DPs, not with bare nouns. This strongly suggests that the difference has to be related to the naming of the article.

results in a structural account that relates omission to properties (f.e. case assigning properties) of the finite verb. The results are, however, compatible with a processing account: difficulties that children have with the realization of finiteness and the realization of articles are related to the children's reduced processing resources. In stage 1 children's processing resources are more reduced than in stage 2, therefore in stage 1 children produce less finite verbs and at the same time, less articles.

- Looking at the effects of position: in Dutch in finite verbal utterances there is an effect of position, for 3 of the 4 children, in the stages collapsed and in stage 2: we find more omissions from initial than from internal position in finite verbal utterances. In Italian in finite verbal utterances there is no significant difference in omission from initial and internal position. In no-finite verbal utterances we find no effect of position, neither in Dutch nor Italian.
- Effect of article type: Dutch children have more difficulties with 'het' than with 'de', resulting in more omissions and more substitution errors.

If we compare the data on child speech with the results found in the analysis of the database of headlines and the results obtained in the experiment with headlines:

- We find the same crosslinguistic differences in the overall preference for omission or overall omission rates: in Dutch the preference for omission is stronger than in Italian.
- We find the same relation between finiteness and article use: use of finiteness and use of articles are related, both in headlines and in child speech, in Dutch and Italian. However the data show that the relation is indirect; there is no 'constraint' that violates omission of articles in the case of the presence of a finite verb. If finiteness is produced, then articles are more likely to be produced as well.
- In Dutch in the database of headlines, in the experiment with headlines and in child speech we find more omissions from sentence-initial than sentence-internal position, (in headlines the difference reached significance only in phrases with finite verbs). In Italian child speech the difference between omission in

sentence-initial and sentence-internal position is not significant, in the Italian experiment with headlines we find more omissions from sentence-initial position only in the conditions with indefinite determiners, but not in the conditions with definite determiners.

- In Dutch child speech we found a difference in the omission rate of *'de'* and *'het'*: children have more problems with *'het'*. We find more omissions, and we find substitutions of *'het'* by *'de'*. In the headlines database and experiment we did not find significant differences in the preferences for omission of *'de'* and *'het'*.

We find similarities in omission patterns, but there are also differences. The somewhat complicated pattern of omission that results from our analysis, makes it hard to account for these findings with a structural, knowledge-based account. Structural accounts predict a direct, straightforward relation between for example finiteness of the verb and article use. My findings show that there is only an indirect relation. There are no strong constraints that prohibit article omission in the presence of a finite verb. Therefore these findings ask for another type of account.

3.6 Conclusion

Already the fact that traditional approaches cannot account for the differences in omission pattern in child speech asks for a different approach, the fact that we find a similar pattern of differences in omission in adult speech is very useful because this suggests that the reason for these differences may be related to the processing cost of articles. It does not seem reasonable to suggest that adults in special registers have lost knowledge of the semantic or syntactic conditions of article use, nor does it seem plausible that (phonological) motivations can cause omission of articles in a specifically written register like headlines. Of course, we can try to find two separate accounts, one for omissions and crosslinguistic difference in adult special registers, and one for omissions and crosslinguistic differences in child language acquisition, but, because of the similarities in both populations the best approach is to find a model that can account for both. In the next chapter I will propose a model for the processing of articles that accounts for both: development of article use in child speech as well as article use of adults in special registers.

Chapter 4 An information-theoretical approach to omission of articles

4.1 Introduction

In the previous chapter I presented data on article omission in Dutch and Italian headlines and in Dutch and Italian child speech. The data showed that we find more omissions in Dutch than in Italian headlines. In an experiment with Dutch and Italian newspaper readers we found a higher preference for omission of articles in Dutch. And in Dutch child speech I observed more omissions, for a longer period of time. The fact that we find omissions not only in child speech but also in adult speech already suggests that omission does not seem to violate strong grammatical constraints. In addition, I showed that a very important characteristic feature of the pattern of omission, both in Dutch and Italian child speech and headlines is the optionality of article omission: sometimes the article is omitted, sometimes it is used. This optionality also suggests that omission does not seem to obey strong grammatical constraints or specific grammatical rules. I showed that the pattern of omission is influenced by particular features like finiteness, sentence position of the article-requiring noun and the type of article (definite or indefinite, and for Dutch child speech: *'de'* or *'het'*).

- As for finiteness, I argued that in child speech, the increase in probability of finite verb use correlates with the probability of the use of articles. There is, however, no direct relationship between finiteness and omission of articles. There is no constraint that prohibits omission of articles if the verb is finite. We find omissions of articles also in sentences that contain a finite verb. And, moreover, if we divide the data over different developmental stages we find no difference in omission of articles in finite sentences and non-finite sentences.
- As for position, I found more omissions in sentence-initial position, but in child speech only if the verb is finite, and in Italian finite headlines only if the article is indefinite. Hence, the effect is not due to sentence-position only, but also depends on aspects like finiteness of the verb and type of article.

Structural accounts alone cannot explain these findings. Traditional knowledge-based accounts of article omission in child speech cannot account for article omission in adults' special registers, like headlines. Why should adults in special registers disobey their knowledge on the use of articles? What is more, structural accounts predict a strong categorical judgement, and do not predict an optionality in the omission pattern.

Intuitively, the data suggest that children and headline writers in Dutch and Italian are sensitive to a particular property of articles that influences omission. This property leads to a higher omission rate of articles in Dutch child speech and in Dutch adults' special registers. What is the nature of this property? It does not seem reasonable to argue that this property is related to the grammar of Dutch and Italian. Optionality of article use is not an option of the grammars of Dutch and Italian. Both languages have so-called 'obligatory' contexts of article use, and contexts in which the article has to be omitted, but no contexts where an article may or may not be used, at least not without changes in meaning, (e.g. *Ik heb boeken gekocht* (I bought books, generic meaning) versus *Ik heb de boeken gekocht* (I bought the books, specific meaning)). Optionality has to be caused by a less rigid factor or condition, a factor that sometimes leads to omission, but not always. A factor that appears to be sensitive to contextual features, like the type of speaker (child or adult) and the speech situation (adults in normal registers versus adults in special registers) but also to linguistic features, like the use of a finite verb in a sentence, the position of an element within the sentence or the type of article that is processed.

In this chapter I will show that we can explain the observed patterns of omission, in child speech and adult speech, as well as the crosslinguistic differences with an account based on processing considerations. Language use depends on two important conditions: a speaker's knowledge of the language and the speaker's performance during the processing of language. Since optionality of omission cannot be straightforwardly explained by lack of knowledge, it has to be related to the performance of the language user during the speech production process, hence, to the processing of language. It thus follows that the property influencing article use and omission has to be related to the processing of articles in Dutch and Italian. Since we find more omissions in Dutch than in Italian, processing of articles in Dutch somehow has to be more 'difficult' than processing articles in Italian. Intuitively, this may be correct, but how can we measure the difficulty of article processing in

a specific language? In this chapter I will propose a method to measure the difficulty of processing articles, based on an application of Shannon's Information Theory (Shannon and Weaver, 1949).

4.2 Aspects of the Information Theory and its application to language processing

4.2.1 Introduction

In chapter 3 I argued that when an article is to be produced, the targeted form needs to receive the highest activation among the members of the set it is a member of. The studies of Caramazza and his colleagues have shown that other members of the article set too receive a certain amount of activation. Thus the final output is the result of a competition between all members of the set. In a sense, the article selection process can therefore be characterized as resolving *uncertainty* with regard to which element of the set will win the competition. Uncertainty is a notion that can be measured, and in fact it is the basic notion of the Information Theoretical approach which finds its roots in Shannon and Weaver (1949).

Information Theory was introduced as a means to study and solve problems of communication or transmission of signals over (technical) communication channels. In the early 1940s it was thought that increasing the transmission rate of information over a communication channel increased the probability of error. Shannon surprised the communication theory community by proving that this was not true. Not the transmission rate of the information but the complexity level of the encoded information (in Shannon's terminology: the 'uncertainty' of the encoded information) determined the probability of errors. Shannon named this complexity level of the encoded information H , 'entropy', analogous to the use of this word in thermodynamics and he proposed a model to calculate this entropy level.

We can think of the human language processing brain in terms of channels transmitting information in order to communicate. And we can then hypothesize that in the processing of human language too errors (for example, omission of articles) occur when the entropy ('uncertainty') of the encoded information exceeds the channel capacity. This leads to

the following challenging questions: can we measure the entropy of, for example, the information encoded in articles in human languages? Are there differences in the entropy values of articles in different languages? Do these differences make the right predictions with respect to omission of articles?

First of all, however, a very crucial question has to be answered. Information Theory was introduced as a means to study and solve problems of *technical* data transmission. Is it legitimate to apply these information theoretical notions to the transmission of information or, communication, in human language, hence to language processing? The answer to this question will be the topic of the next section.

4.2.2 Earlier applications of Information Theory to Language Processing

Earlier approaches to the use of mathematical, statistical methods for the study of language focused on the *description* of language structure (see Chomsky, 1956 for references). Chomsky (1956) argued that a statistical approach to language can shed no light on the problems of grammar. One of his objections was that there is no general relation between the frequency of a string (or its component parts) and its grammaticality. To illustrate his objection he used a sentence pair like (1)

- (1) a. Colourless green ideas sleep furiously.
 b. Furiously sleep ideas green colourless.

(a) is a grammatical sentence, (b) is not. Chomsky argued that if we consider the problem of grammar in part as that of explaining and reconstructing the ability of an English speaker to recognize (a) as grammatical and (b) as ungrammatical, a statistical order of probability model cannot solve this problem. After all, he reasoned, a statistical model cannot distinguish between (a) and (b).

Chomsky was right in claiming that we cannot, at least not in an efficient way, use statistical methods to come to the description of a language model. But, it is important to underline that he was talking about the use of statistical methods for the description of language structure. Using statistical methods for the study of language *processing* is a completely different field of application. Here, as I will show in this

chapter, the use of statistical methods can help us to come to a model of language processing that can account in an elegant and transparent way for findings on language processing in different groups of speakers and in different languages, which cannot easily be explained by a structural account.⁵⁶

Clark (2001) underlines the importance of taking into consideration the possible role of statistical learning in language *acquisition* studies. After all, he argues, ‘languages are learned in the real world and in real time. It may be that real-world learners are oblivious to these statistical properties, but it would be surprising if such a learner ignores a potentially useful source of information given the demanding constraints that the natural environment places on the learning problem. The very fact that the learner will have to face some degree of uncertainty in distinguishing different candidate hypotheses suggests that statistical methods will be of great utility to learners’. (2001:128)

The most convincing evidence that human language processing brains are sensitive to information theoretical notions comes from a series of studies by Kostić (Kostić, 2004, Moscoso del Prado et al., 2004). In a series of reaction time experiments on inflected nouns in Serbian Kostić found that these were sensitive to two values defined in Information Theory: the uncertainty (or information value) of the *individual elements* and the uncertainty (or information value) of the *system*. In the next section I will explain these notions and show how we can use them for measuring the complexity level of the article selection process.

4.2.3 A brief introduction to Information Theory

In this section I will introduce the reader to the basic notions of Information Theory that are important for the argumentation that will follow in this chapter (for a more detailed discussion of Information Theory see Schneider, 1984; Cover & Thomas, 1991; Baddeley et al., 2000). First of all, it is important to keep in mind that the word ‘information’ in Information Theory is used in a special sense, which should not be confused with its ordinary usage. In particular, information should not be confused with ‘meaning’. The word

⁵⁶ Maybe it would not be literally impossible to account for the findings by a structural account, but such an account will necessarily have to be so complex that it will be of little use or interest.

‘information’ in Information Theory relates not so much to what one *says*, as to what one *could* say. That is, information in Information Theory is a measure of one’s freedom of choice when one selects a message.⁵⁷ Since freedom of choice necessarily implies uncertainty, we can also say that information in Information Theory is a measure of uncertainty. This concept of information applies not to individual messages (as the concept meaning would), but rather to the situation as a whole, whereby the notion ‘information’ essentially indicates that in this particular situation there is a certain amount of uncertainty that a message will be selected. The greater the freedom of choice, the greater the uncertainty that a particular message will be selected. And therefore, the amount of information contained in the message will be greater if the message is selected in a situation of greater uncertainty. Thus, greater freedom of choice, greater uncertainty and a greater amount of information go hand in hand: in fact, in Information Theory ‘information and uncertainty find themselves to be partners’ (Shannon & Weaver, 1949:27).

The concepts *information* and *uncertainty* are technical terms related to the process of selecting one or more entities out of a set of entities. Suppose that we have a mechanism that can produce 3 symbols, A, B or C. As we are waiting for a symbol, we are uncertain as to which symbol it will produce. Once a symbol appears and we see it, our *uncertainty* decreases, and we observe that we have received some information. Thus, information, in Information Theory, reflects uncertainty. How can we measure this uncertainty?

The simplest way would be to say that, in our example, we have an ‘uncertainty of three symbols’, or that the information contained by the set is ‘three units of information’. And if we consider a second mechanism, which, for example, produces symbols 1 and 2, we could say that this second mechanism contains ‘two units of information’. If we combine the two mechanisms there will be six possibilities, A1, A2, B1, B2, C1, C2. This mechanism then contains six units of information. We therefore have to *multiply* the number of symbols of the individual sets. However, as the following example illustrates, the fact that we have to

⁵⁷ In order to avoid confusion, it is important to note that the message in the information theory can be anything from a single letter to a long text. We will be applying this approach to the process of article selection. Therefore, a message, for our purposes, is the targeted article (not to be confused with the message as an output of the conceptualizer).

multiply the number of symbols is confusing. It is not the way we intuitively think about the amount of information we receive.

- | | | | |
|-----|---------|----------|------------------------|
| (2) | Set (A) | gives us | 3 units of information |
| | Set (B) | gives us | 2 units of information |

However, the combination of Set (A) + Set (B) does not give us $3 + 2 = 5$ units of information, but $3 * 2 = 6$ units of information

Intuitively we would like our measure to be *additive* rather than *multiplicative*. The easiest way to achieve an additive measure is using logarithms. If we first take the logarithm of the number of possible symbols we can *add* the logarithms instead of multiplying the number of symbols. In the above example the first device makes us uncertain by $\log(3)$, the second by $\log(2)$ and the combined device by $\log(3) + \log(2) = \log(6)$. Therefore, uncertainty in Information Theory is measured by using logarithms.

In Information Theory the base of the logarithms is 2. The base of the logarithms determines the quantity in which the units are expressed. When we use the base 2 the units are in *bits*, base 10 would give *digits* (Schneider, 1984; Cover & Thomas, 1991). Thus, if a mechanism produces one symbol, we are uncertain by $\log_2 1 = 0$ *bits* of information. There is no uncertainty about what the mechanism will do next, and thus, the symbol, if it appears, has no informative value (from an Information Theoretical perspective). If the mechanism produces two symbols it contains $\log_2 2 = 1$ *bit* of information, if it produces three symbols $\log_2 3 = 1,585$ *bits* of information.⁵⁸

So far, the formula for uncertainty is $\log_2(M)$, with M being the number of symbols in a set. In our argumentation we assumed that all elements were equally likely to occur. If the probability distribution of elements is unequal, the amount of information ('uncertainty') the system contains not only depends on the number of elements (M), but also on the probability with which they occur. This can easily be seen in an example of a set in which there are 3 possible symbols, but one of them never appears. Our uncertainty will then be: $\log_2 2 = 1$ *bit* of

⁵⁸ Let me illustrate this with a simple example: $\log_2 4 = 2$, since $2^2 = 4$, $\log_2 8 = 3$, since $2^3 = 8$. Thus: $\log_2 2 = 1$, since $2^1 = 2$, $\log_2 3 = 1,585$, since $2^{1,585} = 3$.

information, and not $\log_2 3 = 1,585$ bits of information.⁵⁹ And if the third symbol rarely appears relative to the other two symbols, then our uncertainty should be larger than 1 bit, but not as high as $\log_2(3)$ bits. In order to calculate the uncertainty of individual elements in the case of unequal probability of the symbols we will have to adapt the formula so that it can take into account the probability with which any symbol can appear. Equation (3) shows how this can be done.

$$\begin{aligned}
 (3) \quad \log_2(M) &= -\log_2(M^{-1}) \\
 &= -\log_2(1/M) \\
 &= -\log_2(p) \text{ (in which } p \text{ is the probability of} \\
 &\quad \text{occurrence)}
 \end{aligned}$$

If all elements have the same probability of occurrence, then $-\log_2(p)$ equals $\log_2(M)$. Let me illustrate this with the calculation for a set of 3 elements with equal probability in example (4)

- (4) a. If all 3 elements have the same probability of occurrence, then p of each element is 0,333. If we insert this value in the formula $-\log_2(p)$ we get:
- $$\begin{aligned}
 -\log_2(p) &= -\log_2(0,333) \\
 &= -\log_2(1/3) \\
 &= -\log_2(3^{-1}) \\
 &= -1 * (-\log_2(3)) \\
 &= \log_2(3) \\
 &= 1,585 \text{ bits of information}
 \end{aligned}$$
- b. If $M = 3$, then: $\log_2(M) = \log_2(3) = 1,585$ bits of information

If the elements differ in their probability of occurrence, then the elements with a high frequency of occurrence (the probability of occurrence, p , is high) will have a low informative value. Suppose that we have an element with a probability of occurrence of 0.9. If we insert this value into the formula $-\log_2(p)$ we will find:

⁵⁹ Assuming that both elements have the same probability of occurrence

$$\begin{aligned}
 (5) \quad -\log_2(p) &= -\log_2(0,9) \\
 &= -\log_2(1/1,111) \\
 &= -\log_2(1,111^{-1}) \\
 &= -1 * (-\log_2(1,111)) \\
 &= \log_2(1,111) \\
 &= 0,152 \text{ (bits of information)}
 \end{aligned}$$

On the other hand, an element with a low probability of occurrence (p is low) will have a high informative value. Suppose that the probability of an element is 0.1. This leads to the following calculation:

$$\begin{aligned}
 (6) \quad -\log_2(p) &= -\log_2(0,1) \\
 &= -\log_2(1/10) \\
 &= -\log_2(10^{-1}) \\
 &= -1 * (-\log_2(10)) \\
 &= \log_2(10) \\
 &= 3,321 \text{ (bits of information)}
 \end{aligned}$$

Hence, the amount of information (I) of an individual element (x) in a set, expressed in bits of information, equals the logarithm of the probability that that specific element will appear, and expresses the decrease in uncertainty that will be achieved if the element becomes available.

Expressed in a general formula this argumentation gives us:

Formula (1) Informative value individual element:

$$I(x) = -\log_2 p(x)$$

in which:

$I(x)$ expresses the informative value of a specific single element, x , out of the set.

Tribus (1961) called this measure the 'surprisal' as it measures the surprise that we get when we see the symbol. If a specific element is a very 'common' element, the surprise will be 'low', if it is a 'rare' element the surprise will be 'high'.

Let us now proceed to a second important notion in Information Theory, the uncertainty of the set, or the *average* information per element of the set. The value of the uncertainty of a *set* is the *average* of the surprisal value of the elements in the set.

Formula (2) (Absolute) Entropy of a set

$$H = - \sum p_i \log_2 p_i \text{ (bits per symbol)}$$

This formula is Shannon's general formula for uncertainty of a set and is called the *entropy (H) of the set*. As the formula shows the entropy consists of the sum of the informative values of the single elements, in which each of these values is weighed for its share in the total frequency. It expresses *the average information per element of a set*, in other words the *average surprisal* of a set of N elements with corresponding (not necessarily equal) probabilities $p_1, p_2, p_3 \dots p_N$.

The entropy of a set reaches its maximum level if all elements have the same probability of occurrence. Let me illustrate this with an example of a set of three elements. In order to make the example less abstract, let us take the set of Dutch articles 'de', 'het' and 'een'.

Suppose all articles have the same probability of occurrence:

$$p_{de} = 0,333$$

$$p_{het} = 0,333$$

$$p_{een} = 0,333$$

These have surprisals or informative values (calculated using formula (1) $I = - \log_2 P_i$):⁶⁰

$$I_{de} = 1,585 \text{ bits}$$

$$I_{het} = 1,585 \text{ bits}$$

$$I_{een} = 1,585 \text{ bits}$$

All articles have the same probability of being selected ($p = 0,333$), and the average uncertainty is:

$$H = - \sum p_i \log_2 p_i$$

$$H = 0,333 * 1,585 + 0,333 * 1,585 + 0,333 * 1,585 = 1,585 \text{ bits per symbol.}$$

This is the maximum level of uncertainty in the set, *the maximum entropy level (H max)* of a set of three elements.⁶¹

⁶⁰ $-\log_2 (p)$ = $-\log_2 (0,333)$
 = $-\log_2 (1/3)$
 = $-\log_2 (3^{-1})$
 = $-1 * (-\log_2 (3))$
 = $\log_2 (3)$
 = 1,585 bits of information

If the articles do not have equal probability of occurrence, but differ in their probability distribution, we will find a lower entropy value. Let us calculate the entropy based on the actual probability distribution rate of articles in Dutch (in conformity with Corpus Gesproken Nederlands)

$$\begin{aligned} p_{de} &= 0,479 \\ p_{het} &= 0,182 \\ p_{een} &= 0,339 \end{aligned}$$

These have surprisals or informative values (formula 1: $I = -\log_2 P_i$)

$$\begin{aligned} I_{de} &= 1,062 \text{ bits} \\ I_{het} &= 2,466 \text{ bits} \\ I_{een} &= 1,561 \text{ bits} \end{aligned}$$

The average uncertainty is:

$$\begin{aligned} H &= -\sum p_i \log_2 p_i \\ H &= 0,479*1,062 + 0,182*2,466 + 0,339 *1,561 = 1,485 \text{ bits per} \\ &\text{symbol.} \end{aligned}$$

The level of uncertainty of a set (average information per symbol), based on the actual probability distribution, is called the *absolute entropy level* (= H).

The higher the differences in probability distribution between the elements, the lower the absolute entropy level. It is not difficult to see that absolute entropy reaches its minimum level when one element has probability 1, and all the rest have probability 0. There is no uncertainty in such a case. We can be completely sure about which element will be selected, hence $H = 0$.⁶²

We can say that the more unequal the probability distribution, the less uncertainty there is in the set and the lower the entropy value will be. In the natural sciences the entropy associated with a situation is a measure of the degree of randomness in the situation. If a situation has a high degree of randomness, the entropy will be high, if there is a low degree of randomness, the entropy will be low.

⁶¹ The maximal entropy of a set of n elements is the absolute entropy of a set with n elements with equal probabilities.

⁶² If we take for example a set of three elements, A, B, and C with the (hypothetical) probability distribution: $p(A) = 1$, $p(B) = 0$, $p(C) = 0$ we get the absolute entropy value $H = -\log_2 (P_i) = -\log_2 (1) = 0$

If we divide the *actual* level of uncertainty of the set (the *absolute entropy*, H) by the *maximum* level of uncertainty (H_{\max}) of a set with n elements, we find the *relative entropy* value (H_r).

$$\text{Formula (3)} \qquad \qquad \qquad \text{Relative Entropy}$$

$$H_r = \qquad \qquad \qquad H / H_{\max}$$

In the example of the Dutch article set, introduced in the preceding discussion, we find:

$$H_r = \qquad \qquad H / H_{\max} = 1.485 / 1.585 = 0.937.$$

What does a relative entropy value of 0.937 mean? It is a measure, an index, of the degree of uncertainty of a set. If we have a set with a number of n elements, then the entropy reaches its maximum level if all elements have the same probability of occurrence (in our example maximum entropy = 1.585). If the actual probability distribution of elements is different, and if elements do not have equal probability of occurrence, the actual entropy (the absolute entropy, H), will be lower than the maximum entropy. Relative entropy measures how far away the absolute entropy is from the maximum entropy, and thus, it is a measure of the degree of uncertainty of the set. It is an index for the level of uncertainty. The lower the relative entropy value, the more the absolute entropy value deviates from the maximum entropy level, and hence, the lower the degree of uncertainty of the set.

Then the question arises: why do we need *relative* entropy? After all, absolute entropy already expresses the level of uncertainty of the system. Why then do we need an additional measure? The reason is that if we want to compare two sets with an unequal number of elements, we cannot just compare the absolute entropy values of the two sets. The absolute (as well as the maximum) entropy is calculated on the basis of the sum of the informative values of the elements. Therefore it is sensitive to the number of elements in the set, and cannot simply be used to compare two sets with unequal numbers of elements. The sum of the informative values of the individual elements is very likely to be higher in a set of for example 18 elements than in a set of three elements. It can be expected that the set with the larger number of elements will have the higher absolute entropy value. The more elements a set has, the more choices there are. Hence, there is more ‘information’

(uncertainty, freedom of choice) if one selects freely from a set that contains 18 elements than if one selects from a set of three elements.

Two sets with an *equal* number of elements with non-uniform probability may or may not have an equal (absolute) entropy value. But two sets with an *unequal* number of elements with non-uniform probability will (most probably) not have equal (absolute) entropy values. If we wish to meaningfully compare the entropy values of sets with unequal numbers of elements, we will have to calculate their relative entropy values. Relative Entropy is the ratio of absolute and maximum entropy. We calculate the absolute entropy of the set, then we calculate what the entropy of the set would have been if all elements had the same probability of occurrence (maximal entropy, H_{\max}). We finally divide the absolute entropy of the set by the maximal entropy, and get the value of Relative Entropy.

Relative Entropy is, we could say, an index of homogeneity of probability distribution. The higher the relative entropy value, the more the actual entropy value of the set approaches the maximum level of entropy, and hence, the more the uncertainty level of the set approaches the maximum level of uncertainty. Another way to formulate it is by saying that the relative entropy value reflects the level of the complexity of the selection of an element from the set. The higher the relative entropy, the more homogeneous the probability distribution of the elements is, and the more difficult it is to select an element out of the set.

4.2.4 From theory to practice

I will illustrate the argumentation in the previous section with some examples. Let me start with two figures, representing two sets with an *equal* number of elements that differ with respect to the probability distribution of the elements within the sets. In set A we find only small differences in the probability distribution of elements, in set B we find large differences.

Figure 1 *Set A: set of 5 elements with small differences in the probability distribution of elements.*

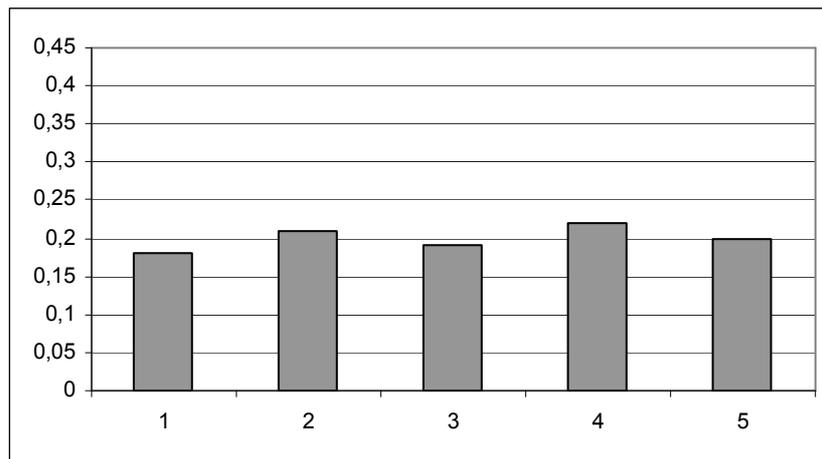
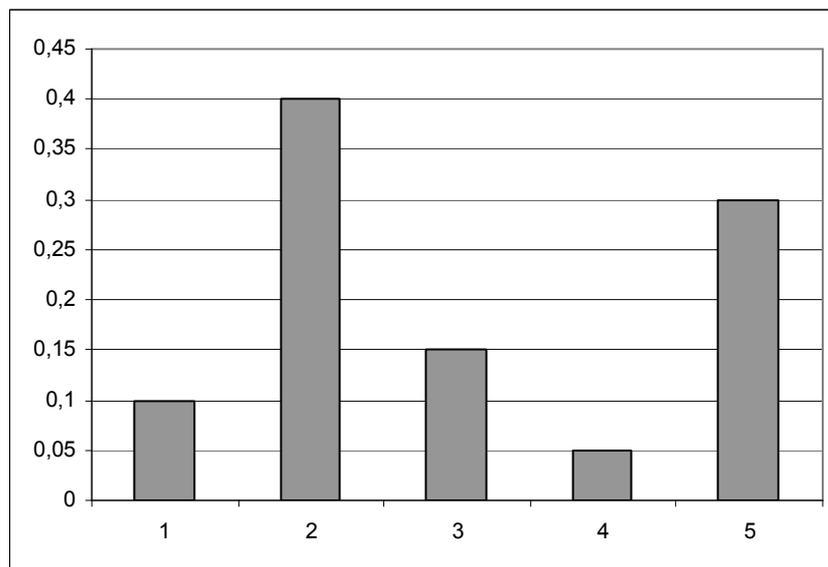


Figure 2 *Set B: set of 5 elements with large differences in the probability distribution of elements.*



Just by looking at the graphs we can see that the elements in set B are more conspicuous. They can be more easily distinguished from the other elements within the set. Thus the uncertainty is higher in set A than in set B, because in set A the elements have a more equal probability of occurrence. How should we measure now the uncertainty levels? Let us start by looking at the absolute entropy values of the sets (formula 2 in the previous section). Table 1 shows the calculation of the absolute entropy of set A.

Table 1 Calculation Absolute Entropy set A (picture 1).

	p	log p	p log p	$H = - \sum p \log p$
element	probability	individual informative load	average informative value	absolute entropy
1	0,18	2,473931	0,445308	
2	0,21	2,251539	0,472823	
3	0,19	2,395929	0,455226	
4	0,22	2,184425	0,480573	
5	0,2	2,321928	0,464386	
	1		Sum: 2,318316	2,318316

Let me explain this calculation in more detail. As a first step, in order to calculate entropy we need to know the probability distributions of the elements in the article sets. The probabilities of the elements of set A are indicated in the second column (p) of the table. Once we know the probability of each element we can calculate the individual informative load, using formula (1), discussed in the previous section: $I(x) = -\log_2 p(x)$. These values are calculated in the third column. Next, in order to calculate the (absolute) entropy of the article set we have to calculate the *average* informative value of the elements in the set, using formula (2): $H = - \sum p_i \log_2 p_i$ (bits per symbol). As the tables show the sum of these values gives us absolute entropy value: 2,318316.

Next we make the same calculations for the elements in set B, and we find the absolute entropy value 2,008695, see Table 2.

Table 2 *Calculation Absolute Entropy set B (picture 2).*

	P	log p	p log p	$H = - \sum p \log p$
element	probability	individual informative load	average informative value	absolute entropy
1	0,1	3,321928	0,332193	
2	0,4	1,321928	0,528771	
3	0,15	2,736966	0,410545	
4	0,05	4,321928	0,216096	
5	0,3	1,736966	0,521090	
	1		Sum: 2,008695	2,008695

We find a higher absolute entropy value for set A. Set A is also the set that, as we saw by looking at the probability distribution of the elements illustrated in Figures 1 and 2, has the higher uncertainty and complexity level. Hence, at first sight absolute entropy seems to reflect the complexity level. But recall that here we are looking at two sets with an equal number of elements.

Let us now take a look at the absolute entropy values of two sets with unequal number of elements.

Figures 3 and 4 illustrate two sets with *unequal* numbers of elements.

Figure 3 Set C: Set of three elements with small differences in probability distribution

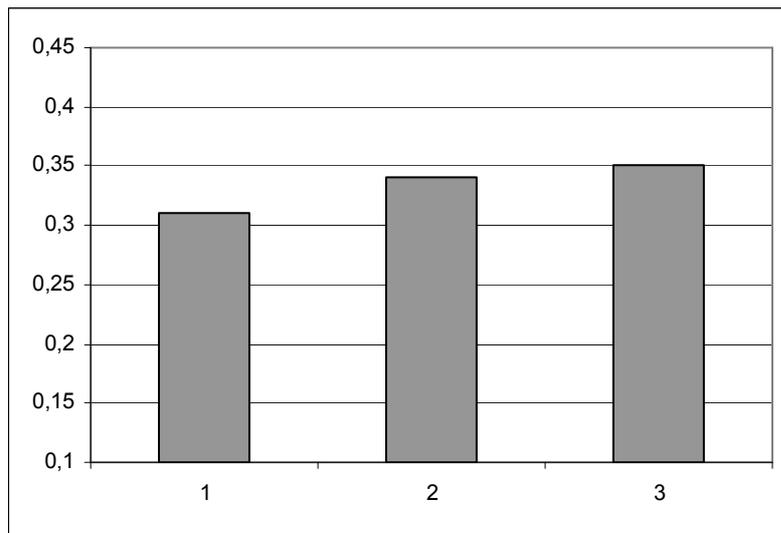
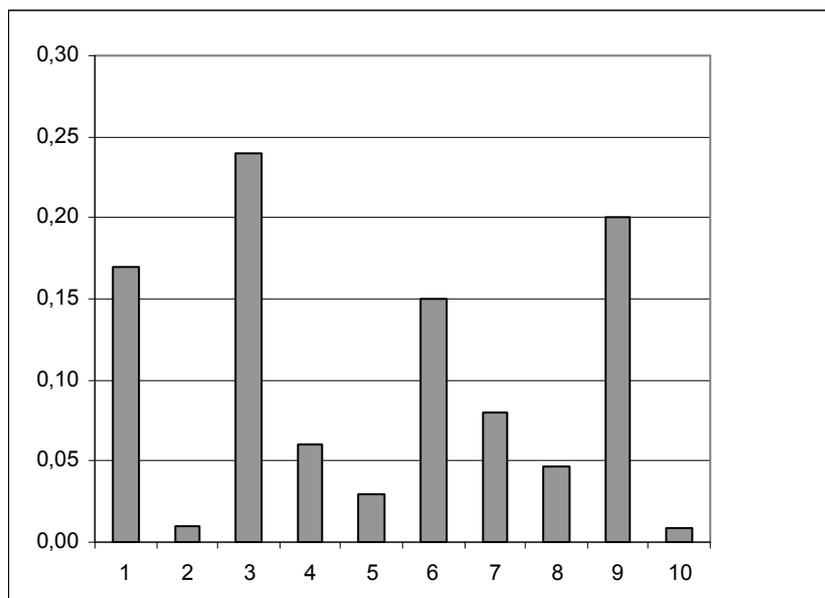


Figure 4 Set D: Set of ten elements with large differences in probability distribution.



Looking at the graphs, we see that the set D illustrated in Figure 4 is the set with the lower level of uncertainty, as the differences in probability distribution are higher than in set C. The elements in set D are more conspicuous, can be distinguished more easily within the set. Let us now calculate the absolute entropy values for sets C and D, in the same way we did for sets A and B. We can then see whether also in the case of different numbers of elements the value of absolute entropy is a correct predictor of the complexity level. For set C, with three elements, we find an absolute entropy value of 1,583069 bits, and for set D, the set with 10 elements, we find an absolute entropy value of 2,821975 bits. The calculations are shown in Tables 3 and 4.

Table 3 *Calculation Absolute Entropy set C.*

	P	log p	p log p	$H = - \sum p \log p$
element	probability	individual informative load	average informative value	absolute entropy
1	0,31	1,68966	0,523795	
2	0,34	1,556393	0,529174	
3	0,35	1,514573	0,530101	
sum	1		Sum: 1,583069	1,583069

Table 4 Calculation Absolute Entropy Set D.

	p	log p	p log p	$H = - \sum p \log p$
element	probability	individual informative load	average informative value	absolute entropy
1	0,17	2,556393	0,434587	
2	0,01	6,643856	0,066439	
3	0,24	2,058894	0,494134	
4	0,06	4,044614	0,245091	
5	0,03	5,058894	0,151767	
6	0,15	2,736966	0,410545	
7	0,08	3,643856	0,291508	
8	0,05	4,433603	0,205168	
9	0,20	2,321928	0,464386	
10	0,01	6,881914	0,058351	
sum	1		Sum: 2,821975	2,821975

We can now see that the absolute entropy value does not make the right predictions for the level of complexity if we want to use it to compare sets with an unequal number of elements. If we look at the graphs we see that in graph D the elements are more conspicuous, and therefore, the uncertainty level in set C is higher than in set D. But set (D) has the higher absolute entropy value. Therefore absolute entropy is not a reliable predictor of the complexity level of the set.

The fact that the (absolute) entropy is higher in set D is not really surprising. Since absolute entropy is based on the *sum* of the informative values of the elements, it can be expected that the set with the larger number of elements will have the higher *absolute* entropy value. But it does not tell us anything about the differences in the probability distribution of the elements in the set, and therefore it does not tell us anything about differences in complexity level between the two sets.

In order to face the problem of sensible comparison of entropy values of sets with unequal number of elements, we will have to use a

measure that is insensitive to the number of elements. This is the measure called Relative Entropy, H_r , which can be calculated using formula (3): Relative Entropy: $H_r = H / H_{\max}$

We already calculated H , absolute entropy, and we now need to calculate H_{\max} , the maximal entropy. As we saw in the previous section, the maximal entropy is the absolute entropy of a set in which the elements have an equal probability of occurrence. For a set of three elements we find the maximum entropy value of 1,584962.⁶³ The calculation is illustrated in Table 5.

Table 5 *Calculation maximum entropy of a set of three elements.*

	p	$\log p$	$p \log p$	$H = - \sum p \log p$
element	equal probability	individual informative load	average informative value	absolute entropy = here maximum entropy
1	0,333333	1,584964	0,528321	
2	0,333333	1,584964	0,528321	
3	0,333333	1,584964	0,528321	
			Sum: 1,584962	1,584962

For a set with ten elements the calculation of maximal entropy is 3,321928 bits is illustrated in Table 6:

⁶³ As we saw earlier this section, in a set with equal probability distribution both the individual informative load of the elements and the absolute entropy value of the set equal the logarithm of the number of elements. Hence, in a set with 3 elements the absolute entropy value can also be calculated as follows: $\log(M) = \log(3) = 1,584962$ (see example 4 for details).

Table 6 Calculation maximum entropy of a set of ten elements

	p	log p	p log p	H = - ∑ p log p
Element	equal probability	individual informative load	average informative value	absolute entropy= here maximum entropy
1	0,1	3,321928	0,332193	
2	0,1	3,321928	0,332193	
3	0,1	3,321928	0,332193	
4	0,1	3,321928	0,332193	
5	0,1	3,321928	0,332193	
6	0,1	3,321928	0,332193	
7	0,1	3,321928	0,332193	
8	0,1	3,321928	0,332193	
9	0,1	3,321928	0,332193	
10	0,1	3,321928	0,332193	
			Sum: 3,321928	3,321928

As the next step we can now calculate the Relative Entropy values of both sets. These calculations are shown in example 7

(7) Set C: Absolute Entropy: H = 1,583069
 Maximal Entropy: H_{max} = 1,584962
 Relative Entropy: H_{rel set C} = H / Hmax = 0,9988

Set D: Absolute Entropy: H = 2,821975
 Maximal Entropy: H_{max} = 3,321928
 Relative Entropy: H_{rel set D} = H / Hmax = 0,8494

The relative entropy value is higher in the set with the higher uncertainty, set C. This calculation shows that the complexity level is reflected in the relative entropy value and not in the absolute entropy value. The set with

the lower complexity level is the set with the lower relative entropy value (but the higher absolute entropy value). So in order to derive a measure, an index, for the complexity of a set, relative entropy is a more reliable measure than absolute entropy. This is because, differently from absolute entropy, relative entropy is not sensitive to the number of elements in a set.

Summary

This concludes the introduction to some basic notions of Information Theory. I will briefly summarize the notions:

- the informative value of a single element from a set:
 $I(x) = -\log_2 p(x)$.
 This measure indicates the surprise we experience when we see element x appear. The higher the probability that x will appear (the more common x is), the lower its informative value will be.
- the *average* informative value or the *average* surprisal of the elements in the set. This is called the entropy or uncertainty of a system: $H = -\sum p_i \log_2 p_i$ (bits per symbol).
 The entropy will reach its maximum level (H_{\max}) if all elements have equal probability of occurrence. The more unequal the probability distribution of elements in the set, the lower the entropy value will be.
- The *relative* entropy of the set: $H_r = H / H_{\max}$
 This measure indicates how far away the actual entropy (uncertainty level) of the set is from the maximum level of uncertainty of a set with the same number of elements, and is therefore a measure, or an index, of the ‘complexity’ of the set. We need this measure if we want to compare sets with unequal numbers of elements, because the freedom of choice, and therefore the uncertainty level of a set, is influenced by the number of elements. The more elements a set has, the more likely it is that the absolute entropy will be higher. By itself this does not necessarily mean that the set is more ‘complex’ than a set with a lower number of elements, since complexity depends on the homogeneity of probability distribution of elements: the more equal the probability distribution, the higher the

complexity. Only by expressing the given entropy of a set of n elements as a proportion of the maximal entropy of a set of n elements can we compare the complexity of sets with different numbers of elements.

4.2.5 Application of Information Theory to Processing of Inflected Morphology in Serbian.

In this section I will illustrate the approach of Kostić, who showed in a convincing way that human language processing brains are sensitive to information theoretical notions. Later in this study I will apply his approach to account for the observed patterns of article omission in Dutch and Italian.

Kostić (Kostić & Katz 1987; Kostić 2004; Moscoso del Prado et al. 2004) conducted a series of studies on the processing of inflected morphology in Serbian. In particular, his studies focused on the question of how the amount of information of the inflected forms, defined along the proposals put forward by information theory, influenced the processing of inflected morphology. The experiments to determine this were lexical decision tasks in which the reaction time was measured of subjects reading words (inflected nouns) and non words (pseudo-inflected nouns).

Serbian is a highly inflected and to a great extent free word order language. Open class words, as well as some types of closed class words, consist of a base morpheme to which an inflectional suffix is attached. Serbian has a complex case system. Each Serbian noun and adjective can appear in seven cases, in both singular and plural. The surface form of Serbian nouns can be ambiguous. One particular surface form may represent more than one case (for example, the feminine noun 'vod-e' could be the genitive singular, the nominative plural or the accusative plural of the noun meaning ('water').) A further complication in the Serbian case system is that not only the noun, but also the inflectional suffix can be ambiguous (for example the suffix 'i' attached to feminine nouns specifies the dative and locative singular).

We find two opposing positions in the traditional accounts of the processing of inflected words:

- Full listing hypothesis: both affixed and non-affixed words are stored in the lexicon and therefore there is no additional

processing for affixed words (Manelis & Tharp, 1977). If there is a processing difference between words it should be ascribed to frequency differences between the two words

- Decomposition hypothesis: this view assumes obligatory morphological parsing where an affixed word is decomposed into its constituents (base form and affix), followed by a search for the base morpheme in the lexicon and then by an evaluation of the validity of their combination (Taft 1979, 1981, 1994). This implies that processing of an affixed word may take longer than processing of a non-affixed word. Frequency plays a role here too, and it has been suggested that two types of frequencies have to be distinguished here: surface frequency (the frequency of the presented form) and cumulative frequency (the sum of the frequencies of all affixed forms that share a particular base morpheme). It has been shown that processing latency for words of the same surface frequency varies as a function of their cumulative frequency (Baayen, Dijkstra & Schreuder, 1997; Cole, Segui & Taft, 1997). Items of higher cumulative frequency tend to be processed faster than those of lower cumulative frequency. Hence, in order to predict processing latency for a morphologically complex word, not only its own frequency but also the frequency of all other words that share the same base morpheme must be considered.

Previous studies on Serbian nouns (see Kostić 2004 for references) do not seem to support either the full listing hypothesis, nor the decomposition model. Neither of the two crucial factors, i.e. decomposition and frequency can capture the patterns of decision latencies for inflected Serbian noun forms. Kostić has put forward an alternative factor that reflects linguistically relevant aspects of inflected forms: the number of functions and meanings of a specific form which can be seen as a measure of the complexity of a form. What is meant by the number of functions and meanings of a specific form? Kostić argues that grammatical case represents the set of potential relations that a noun can enter into, and that these relations are realized in terms of functions or meanings.⁶⁴ When we say that a specific noun case in

⁶⁴ Functions are well-defined grammatical relations a noun can enter into, such as direct object or agent. Meanings are other semantic relations that can be expressed by a

Serbian can be ambiguous, we are saying that it can have multiple functions or meanings. To give an example, nominative case in Serbian represents both subject and predicate roles as in the following sentences:

- (8) *Prijatelj je došao.* (The friend has come.)
Petar je učitelj. (Peter is a teacher).⁶⁵

On the other hand, accusative can, in addition to its most common object role, represent a vast number of different kind of relations such as

- *time:* *Zoru je proveo čekajući ga.* (He spent the morning waiting for him.)
- *place:* *Popeo se na planinu.* (He climbed the mountain.)
- *cause:* *On je odgovoran za njihovu nesreću.* (He is responsible for their tragedy).

Basing himself on six standard grammar books of Serbian, Đorđe Kostić compiled a taxonomy of functions and meanings for all Serbian cases (Kostić, Đ. 1965). This taxonomy includes a list with the number of functions and meanings for each Serbian case, specified for gender and number. For example, it shows that the masculine singular nominative case in Serbian (for example *konj*, 'horse') has three functions and meanings, and the masculine singular dative case (for example *konju*) has 22 functions and meanings. The list also gives an example for each case function and meaning. Further, since some noun forms can correspond to more than one case, the taxonomy also includes a listing of the *cumulative* number of functions and meanings encompassed by a particular noun form. For example, the noun form '*konju*' with inflectional suffix -u encompasses 43 functions/meanings, as it can be used in masculine singular dative case with 22 functions/meanings and in masculine singular locative case with 21 functions and meanings.

Kostić suggested that the absolute number of case functions and meanings should be taken as tentative. However, what is more relevant is the *proportion* of functions and meanings encompassed by a particular case *relative to* other cases. This proportion is an index of the case's relative complexity. The more functions and meanings a form encompasses, the more complex it is, and, consequently, its processing

noun form, such as place, time, purpose, cause, direction, among others (Blake, 1994)

⁶⁵ Nominative particle is zero - the dictionary form for the noun is the nominative form. Accusative is -u (planin-u, zor-u, nesrec-u)

latency should be longer. He thus predicted that processing time should be influenced by ‘form complexity’, defined as the number of associated functions and meanings. However, he found no significant correlation between form complexity and decision latencies. Hence, neither frequency nor complexity of the form can by themselves account for decision latencies. But Kostić showed that the combination of the factors frequency and form complexity can.

The combination of frequency and number of functions and meanings of a specific form gives the ‘average frequency per function/meaning for a particular noun form’. If we express this as a proportion, relative to the sum of average proportions per function/meaning for all noun forms of a given gender, we can calculate the amount of information carried by a particular noun form, using the following adjusted version of formula (1), defined above in the introduction to Information Theory.⁶⁶

(9)

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

What this informative value tells us is how big the surprise is to find a specific inflected form used in a particular function. The bigger the surprise, the higher the informative value.

Let me explain this notion and its calculation step by step, using the example of a masculine noun with inflectional suffix –u, like *konju*, of which I discussed the frequency data earlier in this section.

Table 7 shows the calculation of the informative value for masculine nouns in Serbian.

⁶⁶ In formula (1) the log was calculated on the probability that a specific signal/element occurs in a system. In the adjusted version the log is calculated on the probability that a specific element (inflected form) occurs in a specific function.

Table 7 The calculation of the informative value for masculine nouns in Serbian.

	Case / number	F_m %	R_m	F_m / R_m	$(F_m / R_m) / \sum (F_{mj} / R_{mj})$	$I = -\log ((F_m / R_m) / \sum (F_{mj} / R_{mj}))$
konj	sg.nominative	12,830	3	4,276667	0,7404835	0,433
konja	sg.genitive + sg.accusative, pl. genitive,	18,007	109	0,165202	0,0286039	5,128
konju	sg.dative + sg.locative,	4,635	43	0,107791	0,0186634	5,744
konjem	sg.instrumental	1,895	32	0,059219	0,0102534	6,608
konji	pl.nominative,	3,326	3	1,108667	0,1919601	2,381
konje	pl. accusative	2,211	58	0,038121	0,0066004	7,234
konjima	pl.dative, pl.instrumental, pl. locative	1,488	75	0,01984	0,0034352	8,185

F_m : indicates the frequency of a specific masculine noun form expressed as a percentage of the total number of (all)inflected nouns. For example, the inflected masculine noun form *konju* is used in masculine singular dative and locative case. The frequency of nouns used in masculine singular dative and locative case expressed as a percentage of the total number of inflected nouns is 4,635 % (frequency data from Kostić, D.,1965). Thus, $F_m = 4,635$.

R_m : the number of functions/meanings the inflectional suffix can appear in. Earlier this section we saw for example that a masculine noun with inflectional suffix *-u* can be used in singular dative case with 22 functions/meanings and in masculine singular locative case with 21 functions and meanings. Thus a masculine noun with the inflectional suffix *-u*, like *konju* has 43 functions/meanings, $R_m = 43$.

F_m / R_m : indicates what is the frequency with which the specific inflected noun form appears in a certain meaning. From the perspective of language processing this is an important relation. Every time a specific form appears with a certain meaning it creates a memory trace in the language processing brains. The less ambiguous a form is, the stronger the memory trace will be (given a certain frequency value). For example,

the average frequency with which a masculine singular noun with inflectional suffix *-u* (*konju*) appears in one of its 43 functions/meanings is $4,635 / 43 = 0,107791$. However, this value by itself does not give us all the necessary information with respect to the processing cost of this specific form. After all, there is a competition between the elements in the paradigm, they all fight to be selected. Therefore knowing what is the strength of the memory trace of an *individual* element is not enough. We need to know the strength of the memory trace of the *other* elements in the paradigm, the ‘competitors’. An element with a low ambiguity, and consequently a strong memory trace, will be more difficult to retrieve when it is surrounded by other elements with a low ambiguity value (and strong memory traces) than when it is surrounded by elements with a high ambiguity (low memory traces). Therefore we need to know whether the value 0,107791 we obtained means that the noun form *konju* has a high or a low surprisal value. And this depends on the surprisal value of the other elements in the paradigm. We need to know how ambiguous the other forms in the paradigm are. Hence, the ambiguity of the other elements in the set matters. We therefore need to ‘weigh’ the value F/R . We need to know what the average frequency is with which the other inflectional suffixes in the same paradigm appear in a certain meaning. In other words we need to know the average surprise value of the other elements in the paradigm. We thus need to calculate:

$\Sigma F_{mj} / R_{mj}$. We calculate the F/R rate for the other inflected forms of the paradigm, and add them up.

Next we divide the individual F/R rates of the inflectional affixes by the sum:

$((F_m / R_m) / \Sigma F_{mj} / R_{mj})$: and find the values in column 6. Of these values we calculate the logarithm, and we find the Informative Values of the individual elements. This informative value indicates how big the surprise (or the cognitive effect) is to find a specific inflected form used in a specific function. The bigger the surprise, the higher the informative value.

Kostić found that I_m , the amount of information of a particular noun form, could account for the decision latencies, not only in his own experiments but also in a replicated version of the experiments of Lukatela et al. (1987). Kostić conducted four experiments each with a group of nouns from one particular gender: masculine, neuter, regular and irregular feminine nouns, and in all experiments he found a very

strong correlation between the amount of information carried by a specific form and response latencies: the higher the amount of information, the longer the reaction time was. Figure 5 shows, for example, the relation between the amount of information of a specific form and the response time for regular feminine nouns.

Figure 5 *The relation between response latencies to six forms of regular feminine nouns and the amount of information (I) carried by these forms (from Kostić, 2004.)*

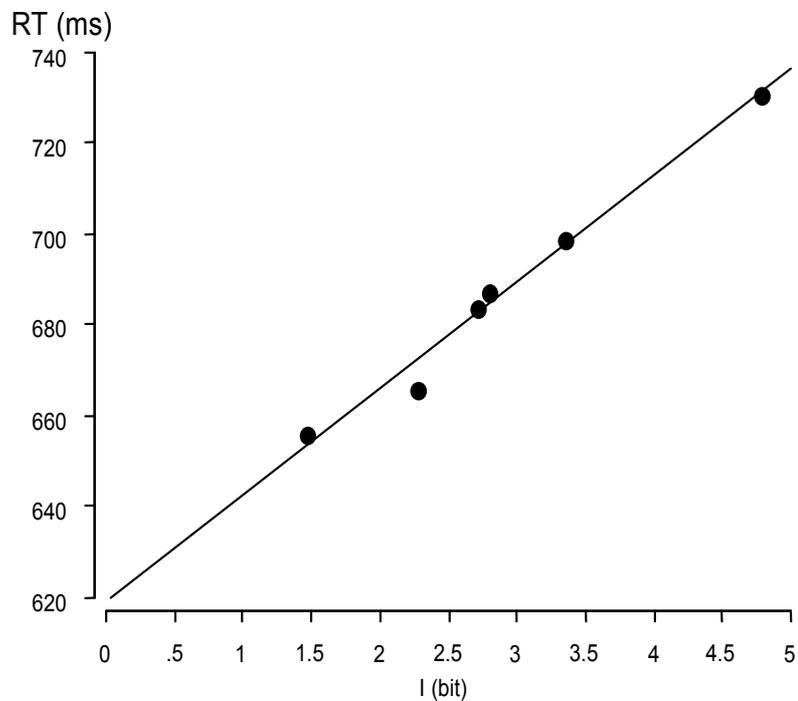


Figure 5 shows an almost perfect correlation between processing latencies and the amount of information (correlation coefficient: 0,98). In other words, Kostić showed that frequency all by itself is not enough to account for the reaction times. What a form can do is relevant as well. The amount of information of a specific form expresses the relation between these two values, the frequency and the number of functions of the form. And the amount of information, calculated with the formula shown in 9, is a reliable predictor of the response times of the forms

within a paradigm. The observation ‘within a paradigm’ leads us to another important factor influencing response time. The amount of information carried by a form is not the only factor that influences response latencies. The slope of the regression line in Figure 5 indicates the amount of time necessary to process one bit of information. Interestingly, a further inspection of the results of Kostić’ experiments revealed that the processing time per bit of information varied systematically with the *average* amount of information in the experiment. Let us follow his argumentation, starting with the question how to calculate this *average* amount of information in the experiment. Recall that the experimental conditions were divided into four groups, depending on gender. Therefore, the *total* amount of information in an experiment was the sum of the informative values of the noun forms within each paradigm. If we divide this sum by the number of forms of the paradigm, we find the *average* amount of information in the experiment. Let me illustrate this with an example. Table 7 presented the calculation of the informative values of masculine inflected noun forms, Table 8 reports these values:

Table 8 *Sum of the informative values of masculine inflected noun forms.*

Form	Informative Value
konj	0,433
konja	5,128
konju	5,744
konjem	6,608
konji	2,381
konje	7,243
konjima	8,185
Sum of Informative Values	35,722

If we divide this sum by the number of forms we find the *average* amount of information, within the paradigm: $35,722/7 = 5,1$. If we make the same calculation for the other gender paradigms we find the following *average* informative values:

Table 9 *Average amount of information (global information load) in the four experiments.*

Experiment	Average amount of information = global information load
Masculine nouns	5,1
Feminine nouns, regular	2,89
Feminine nouns, irregular	2,48
Neuter nouns	2,81

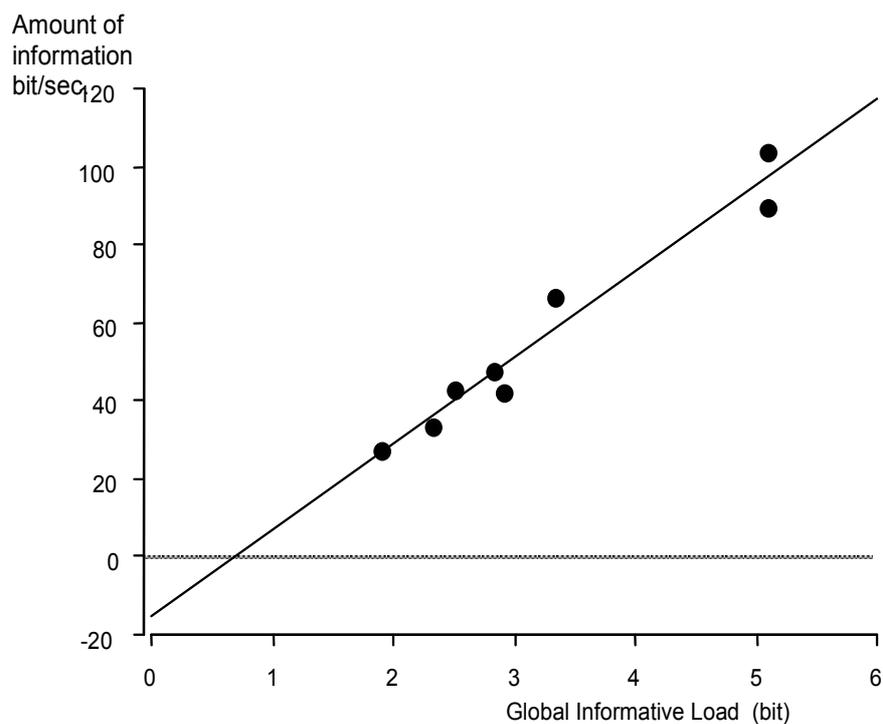
Kostić called this average amount of information ‘global information load’, to distinguish it from the ‘individual amount of information’ of a specific form. The experimental results showed that the processing speed per unit of information varied across the experiments. Formulated differently the amount of information that was processed per unit of time (second) varied in the different experiments, as is illustrated in Table 10.

Table 10 *Amount of information processed per unit of time (bit/sec) in the four different experiments.*

Experiment	Global information load	Amount of information processed per unit of time (bit/sec)
Masculine nouns	5,1	104
Feminine nouns, regular	2,89	41,79
Feminine nouns, irregular	2,48	42,39
Neuter nouns	2,81	46,95

Interestingly Kostić observed that the relation between the global information load and the amount of information that was processed per unit of time was systematic. As Figure 6 shows an increase in the global information load is paralleled by an increase in the amount of information processed per unit of time (bit/sec).

Figure 6 *Relation between information load (global informative load) and the amount of information processed per unit of time (bits/ second).⁶⁷*



In other words, if a particular form is presented in the context of other forms with high informational values (in a context with a high global information load), its processing time per unit of information (bit) will be shorter than if it is presented in the context of forms of low informational value. Hence, the processing time for a particular form varies as a function of the informational values of the other forms presented in the experiment.

⁶⁷ Regression analysis showed that 96% of the variation in processing speed was accounted for by the information load in the experiment ($Y = -15.159 + 22.127 * X$; $R^2 = .957$, see Kostić, 2004). The figure contains not only the data points of the four experiments discussed in this section, but also the data points of the studies of Todorović (1988) and Kostić and Katz (1987) and of replicated versions of studies of Lukatela et al (1987).

Summarizing, Kostić found that two factors influence recognition latencies for inflected noun forms in Serbian:

- at the level of the individual form the processing time was directly proportional to the amount of information carried by that specific form : the higher the amount of information carried by the specific form, the longer the reaction time.
- at the global level, where information load was defined in terms of the average amount of information in an experiment, higher global information loads were associated with faster processing time per bit of information. In other words; the higher the global information load of the experiment, the higher the processing speed per bit of information.

The processing time of an element not only relies on its individual informational value, it also depends on the average informational value of the system in which the element is presented (in Kostić' study the different experiments). An element in context A with an informational value of 2.5 bits will not necessarily have the same processing time as an element with the same informational value in context B. If context B has a higher average informational value than context A, the element with informational value 2.5 will be processed faster in context B than in context A.

4.3 Application of information theory to processing of articles

In the previous section we saw how individual and average informative values influence language processing. In this section I will show how we can apply these findings to account for the differences in the complexity level of the article selection process in Dutch and Italian.

Differently from Kostić, we will not look at inflected forms, but at sets of function words, in particular articles. And, differently from Kostić, our first interest is to compare the levels of complexity of article selection of two different sets, namely Dutch and in Italian, rather than the processing difficulties related to specific forms (articles) within one paradigm or set (these will be discussed later).

The Italian and Dutch article sets are two sets with an unequal number of elements. The Dutch article set contains 3 elements, the

Italian article set contains 18 elements. As we saw in our introduction to Information Theory, if we want to make a sensible comparison between the levels of complexity of two sets with an unequal number of elements, we will have to calculate the *relative entropies* of the two sets. Tables 11 and 12 illustrate the calculations of relative entropies for the Dutch and Italian article sets.

Table 11 *Calculation Relative Entropy Dutch article set.*

RELATIVE ENTROPY DUTCH ARTICLES (based on freq. 'Corpus Gesproken Nederlands)				
	freq.	p = rel freq	log p	p*log p
de	253210	0,478969311	-1,061994874	-0,508662953
het	96327	0,182211116	-2,456317116	-0,447568284
een	179119	0,338819573	-1,561410877	-0,529036566
	528656	1		-1,485267803
			Absolute Entropy H =	1,485267803
H MAX: if all articles have some probability:				
de		0,333333333	-1,584962501	-0,528320834
het		0,333333333	-1,584962501	-0,528320834
een		0,333333333	-1,584962501	-0,528320834
		1		-1,584962501
			Max. Entropy H max =	1,584962501
RELATIVE ENTROPY Dutch articles = H/Hmax				
H rel =	H/H max =			0,937099649

Table 12 Calculation Relative Entropy Italian article set.

RELATIVE ENTROPY ITALIAN ARTICLES (freq.data corpus De Mauro,1993)				
	freq.	p = rel freq	log p	p*log p
il	7111	0,19534641	-2,355893356	-0,460215308
la	7254	0,199274765	-2,327169067	-0,463746069
le	2536	0,069666502	-3,843391067	-0,267755611
lo	425	0,011675183	-6,420411066	-0,074959472
i	2637	0,072441075	-3,787048241	-0,274337844
l'	3120	0,085709576	-3,544399783	-0,303789004
gli	820	0,022526235	-5,472249997	-0,123269188
un	6624	0,181968024	-2,458243139	-0,447321646
un'	800	0,021976814	-5,507873907	-0,121045523
una	3838	0,105433767	-3,245591101	-0,342194897
uno	227	0,006235921	-7,325181609	-0,045679255
degli	111	0,003049283	-8,35731423	-0,025483816
dei	394	0,010823581	-6,529678277	-0,070674503
del	18	0,000494478	-10,9818051	-0,005430265
della	15	0,000412065	-11,2448395	-0,004633608
delle	470	0,012911378	-6,27521315	-0,081021652
dello	1	2,7471E-05	-15,1517301	-0,000416233
dell'	1	2,7471E-05	-15,1517301	-0,000416233
sum	36402	1		-3,112390128
			Absolute Entropy H =	3,112390128
H MAX: if all determiners have equal probability:				
il		0,055555556	-4,169925001	-0,2316625
la		0,055555556	-4,169925001	-0,2316625
le		0,055555556	-4,169925001	-0,2316625
lo		0,055555556	-4,169925001	-0,2316625
i		0,055555556	-4,169925001	-0,2316625
l'		0,055555556	-4,169925001	-0,2316625
gli		0,055555556	-4,169925001	-0,2316625
un		0,055555556	-4,169925001	-0,2316625
un'		0,055555556	-4,169925001	-0,2316625
una		0,055555556	-4,169925001	-0,2316625
uno		0,055555556	-4,169925001	-0,2316625
degli		0,055555556	-4,169925001	-0,2316625
dei		0,055555556	-4,169925001	-0,2316625

del		0,055555556		-4,169925001	-0,2316625
della		0,055555556		-4,169925001	-0,2316625
delle		0,055555556		-4,169925001	-0,2316625
dello		0,055555556		-4,169925001	-0,2316625
dell'		0,055555556		-4,169925001	-0,2316625
sum		1			-4,169925001
				Max Entropy H max=	4,169925001
RELATIVE ENTROPY Italian articles = H/Hmax					
H rel	=	H/Hmax	=	0,746389954	

The calculation of relative entropy of a set was discussed in detail in section 4.2.4. I will here summarize the calculation briefly. As a first step, in order to calculate relative entropy, we need to know the frequency distributions of the elements in the article sets. I used corpora of spoken Dutch (Corpus Gesproken Nederlands, Nederlandse Taalunie, 2004) and spoken Italian (De Mauro, 1993) to calculate the frequencies of the different article forms in both languages. On the basis of the frequency distributions we can calculate the probability that each article will appear ($= p = \text{rel.freq.}$ in Tables 11 and 12) Once we know the probability of each article form we can calculate the individual informative load, using formula (1) from the introduction: $I(x) = -\log_2 P(x)$. In order to calculate the entropy of the article set as a whole we have to calculate the *average* informative value of the elements in the set, $H = -\sum p_i \log_2 p_i$ (bits per symbol). As the tables show this gives us the following values:

$$\text{Dutch: } H_{\text{dutch}} = 1,485267803$$

$$\text{Italian: } H_{\text{italian}} = 3,112390128$$

We see that in Italian the value of the absolute entropy is higher. In the previous section we saw that a proper comparison of entropy values of systems with unequal number of elements requires the calculation of the Relative Entropy: $H_r = H / H_{\text{max}}$

Dutch:	H_{dutch}	= 1,485267803
	$H_{\text{max dutch}}$	= 1,584962501
	$H_{\text{rel dutch}}$	= 0,937099649
Italian:	H_{italian}	= 3,112390128
	$H_{\text{max italian}}$	= 4,169925001
	$H_{\text{rel italian}}$	= 0,746389954

Figures 7 and 8 illustrate the probability distribution of the articles in the Dutch and Italian article sets. The figures show that the sets differ with respect to the homogeneity of the probability of the elements. The probability differences are more conspicuous in the Italian article set, as is confirmed by the calculation of Relative Entropy.

Figure 7 Probability distribution of the articles in the Dutch article set.

Dutch article set: Absolute Entropy = 1,49 bits, Relative Entropy = 0.94

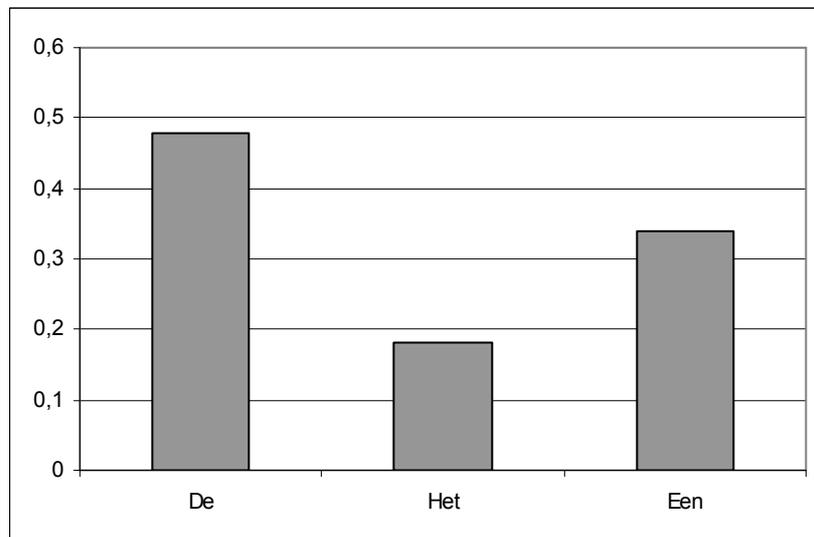
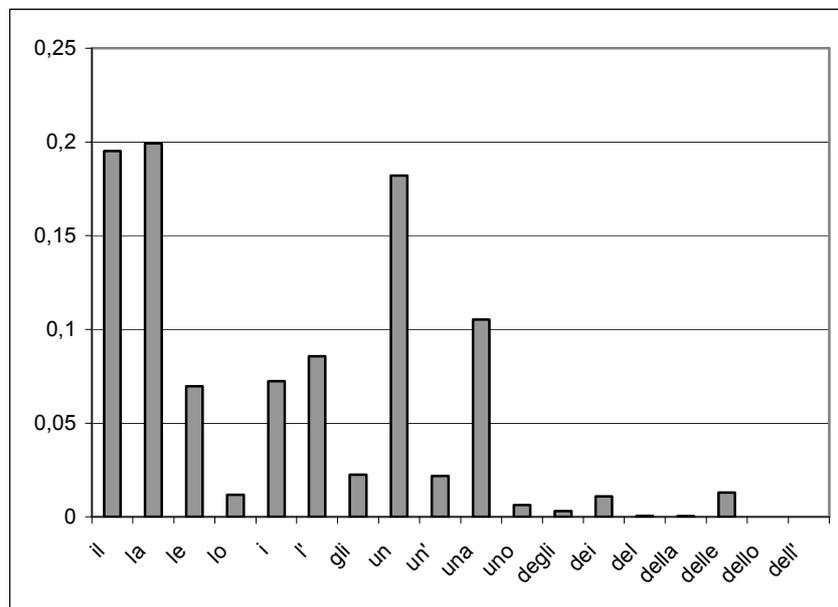


Figure 8 *Probability distribution of the articles in the Italian article set.*

Italian article set: Absolute Entropy = 3,11 bits, Relative Entropy, Hr = 0.75



Let me now explain what the data we have found tell us about the process of article selection in Dutch and Italian. The relative entropy values show that the degree of uncertainty in the Dutch determiner selection process is higher than in Italian. This is because the Dutch article set has a more equal distribution of probabilities of the different articles. The actual entropy value of the Dutch article set is closer to the maximum entropy level than the Italian entropy value. Stated differently, the contrasts between the probability values of the individual elements is bigger in Italian than in Dutch, and this leads to a smaller relative entropy in Italian. The overall probability differences in Italian are, one could say, more conspicuous. That is why the degree of uncertainty is lower in the Italian determiner selection process. The larger the probability differences between the elements, the lower the relative

entropy and, as a consequence, the less cognitive effort to process an element of a paradigm. Hence, due to the fact that the degree of uncertainty of the Italian article selection process is smaller, it requires less cognitive effort to select a specific element from the set.

In the next section, I will combine these findings based on an information theoretical approach to article selection with the model of article selection I proposed in chapter 3, and show how we can use this model to account for the observed omissions in child speech and headlines.

4.4 Putting the pieces together

4.4.1 Model of article selection

In chapter 2 I proposed a model for article selection, and suggested how this model could account for the patterns of omission in child speech, headlines and for the crosslinguistic differences. The only missing element in my proposal was the calculation of the complexity of the selection process of articles in Dutch and Italian. After the conclusion in the preceding section we can now put the pieces together, and complete the model.

Focusing on the production of a noun with an article (a DP), we saw in chapter 2 how the communicative intention was passed to syntax, where the grammatical encoding takes place. The grammatical encoding of a noun involves a so-called functional noun procedure, which is stored in our mental lexicon and triggers the creation of a syntactic structure that is appropriate for the noun lemma. A functional noun procedure contains the rules that specify the article frame, defining which slots have to be filled in. The functional noun procedure also contains the functional elements that are required to implement these rules. These elements will be selected from the article set on the basis of the information contained in the article slots. I showed that the slots of the article frame can be filled in by either information that has been encoded in the communicative intention (like for example the accessibility index), or by information inherent to the selected noun (like number, gender, but also semantic features like inherent uniqueness of the noun). Following Caramazza I argued that an article can only be

selected if all article slots are filled in. But this does not imply that an article will always be selected if all article slots are filled in. Only in a normal speech situation with a normal speaker is a filled article slot a necessary and sufficient condition for retrieval of an article from the article set. In 'special' contexts, or with 'special' speakers an article can be omitted in spite of the fact that the frame is filled in.

In chapter 2 I speculated that these omissions were related to the processing resources and time necessary to select an element from the article set: the more processing resources and time are necessary, the more likely we will find omissions. I speculated that the necessary time and processing resources would be influenced by the level of complexity of the article selection process, and that this level of complexity depended on the competition level between activated articles. The stronger the competition, the more difficult and the more costly in terms of required processing resources the article selection process would be. We can now replace my speculations with the calculations and notions we have available after the first sections of this chapter and complete the model. In the following part of this section I will show how we can account for the fact that, in special situations, an article may not be selected because of the high relative entropy value of the set.

As we saw, in chapter 2, the communicative intention is passed to Syntax which aims to transform the information contained in the communicative intention into a linguistic form. I argued that the processing resources that are available for this transformation process at the Syntax level are limited, and that therefore the output necessarily has to be compressed. I argued, following Levelt and Avrutin that the processes taking place at the Syntax level are automatic and non-intentional. But, I also emphasized, that the fact that the processes are automatic does not necessarily imply that the output is fully predetermined on the basis of the input. Of course, it is not possible to influence the internal structure of the automatized process, but there are factors within the level of Syntax that influence the realization of the process and hence influence the output. These factors are:

- the processing resources that are required to carry out the automatized processes.
- the time speakers have at their disposal: this factor follows automatically from the previous one: if the number of informational units that can be processed *in a given time span* is restricted, then the total number of informational units that can

be processed depends on the available processing time. The more time available, the more units of information can be processed.

According to the model proposed by Shannon and Weaver (1949) error free transmission of information is possible as long as the entropy level does not exceed the channel capacity. The channel capacity is the maximum amount of information per unit of time that can be processed by the channel. ($C = I_{\max} / t$). Switching back to my model: as long as the entropy level (or: the amount of information the syntactic channel has to process) does not exceed the resources that are available for the syntactic channel (the channel capacity), the output will be a 'normal' grammatical utterance.

Only in a normal situation, with a normal speaker under normal discourse circumstances, is the output fully predetermined on the basis of the input. Using Shannon and Weaver's terminology, in such a normal situation the actual amount of information that is sent through the channel per unit of time ($R = I/t$) does not exceed the capacity of the syntactic channel.⁶⁸ In a normal discourse situation with a normal adult speaker the available processing resources and the available time will suffice to translate the communicative intention of '*the weird scientist*' into a linguistic form. This means that the grammatical encoder will be able to retrieve the lexical and functional elements that constitute the linguistic message from the lexicon. Errors can occur, according to Shannon and Weaver's model, if the capacity level of the channel is too low to process the signal. The capacity level of the channel is the amount of information that can be processed in a certain time unit.

⁶⁸ De Roo, Kolk & Hofstede (2003) conducted an interesting experiment in which normal adults were tested in a 'non-normal' discourse situation. De Roo et al. asked normal and agrammatic adults to describe pictures. The normal adults were instructed to use sentences no longer than two or three words to describe pictures. Comparison of the results showed similar structural properties for the syntactically reduced speech of both categories of speakers (especially in the two-word condition for normal adults: omission of finiteness, use of intransitive prepositions (e.g. '*boven*' (on top)), use of non-verbal predicative adjectives (e.g. '*vierkant blauw*' (square blue)). The authors hypothesize that these constructions are structurally less complex, and therefore require less computational steps and less time to be computed. In other words, they require less processing capacity of the speaker. In Shannon's terminology: they require less channel capacity.

4.4.1.1 Article selection in child speech

As I discussed in chapter 2, children's brains are not fully matured. Their brain capacity is not yet fully developed nor structured in an adultlike way. Therefore the available processing resources and the available processing time are not always allocated in such a way that the brain regions involved in language processing can cope with the language-specific requirements for the selection of elements from the lexicon. In Shannon's terminology the maximal capacity of the child's syntactic channel is lower than the maximal capacity of the adult's syntactic channel. Thus, what happens is that elements that require too much cognitive effort and processing time will be omitted (because the child's channel capacity is too low to process these signals). Now why does this lead to omission of functional categories in general? Differently from lexical elements, functional elements all come from a relatively small set of closed-class elements. In chapter 2 I proposed that functional elements are stored in the mental lexicon together with the rules that have to select them, so we have a separate article store, a store with inflections, etc. Lexical elements, which are open class elements come from a far larger set of elements, hence, in information-theoretical terms, their surprise (or informational) value is by definition larger than the surprise value of a functional element. Making an exact calculation of the informative value of lexical elements is beyond the scope of this study. In order to calculate them one would first have to decide on a specific model of storage of lexical elements, then about a way to treat affixes, etc. It is, however, reasonable to assume that the informational value will be higher than the informational value of functional elements. In addition, the differences in probability distributions between lexical elements will be more pronounced than the differences in probability distributions between functional elements. After all, we will not find many functional elements in the lowest ranks of frequency distributions of words. Almost all functional elements can be found in the highest ranks of frequency distribution. This is not the case for lexical elements. They can be found in the highest as well as in the lowest ranks of frequency distributions. Following this reasoning we may hypothesize that the relative entropy of lexical elements will undoubtedly be lower than the relative entropy of functional elements. Recall that the higher the relative entropy, the more processing effort is necessary to select an element from a set. Hence, the fact that children with a lower channel

capacity than adults omit functional elements rather than lexical elements is not surprising.

Why do Dutch children omit more articles than Italian children? The answer to the question should follow immediately, after the calculations performed in the previous section. The relative entropy of the article set in Dutch is higher than in Italian. I hypothesize that children do not have enough processing resources to process this high relative entropy value, and therefore we find more omissions, and during a longer period of time, in Dutch than in Italian.

At this point a question may arise. If children do not have enough processing resources to process the high relative entropy value, then how can we account for the *optionality* in the omission pattern? How can we explain that sometimes, in spite of their limited processing resources, they are able to select an article from the article set? There are a number of reasons for this optional article use. In the first place, it is important to note that the question whether or not the resources in a given situation are too restricted to process the relative entropy depends on more factors than the entropy value alone. For example, factors like the child's fatigue, distraction, (lack of) concentration, etc. may all influence the amount of processing resources available for language processing at a specific moment, and hence, affect the omission pattern in a specific utterance. As I argued in chapter 2, the fact that the child's brain is not mature does not only mean that its capacity is more limited, but also that the available resources are not always distributed in an adult-like way. Hence, the child's speech production may be far more vulnerable to resource-consuming non-linguistic activities and other such circumstances.

There may also be another reason for why the child can still, optionally, produce articles. It may be the case that the child works with an article set with a different probability distribution among its elements than the adult article set. Children can work with the same article set as adults, but may use them statistically differently from what adults do. Therefore, the relative entropy value of the child's article set is lower, and this makes it possible for the child to sometimes produce an article. I will discuss this option in more detail in the next section.

4.4.1.2 Relative entropy in children's files: additional evidence for the relation between omission in child speech and relative entropy

Relative entropy is a measure of uncertainty. It reflects the level of uncertainty our processing resources have to deal with in order to select an element from a set. This implies that relative entropy also reflects the (minimal) capacity of our processing resources. For example, if in spoken adult Dutch the relative entropy of the article set is 0,94, then, since Dutch adults are normally capable of producing articles without errors, this means that humans are able to deal with an uncertainty level of at least 0,94.

Now, in the same vein we could argue that the articles that children do produce reflect the level of uncertainty they can handle. They cannot handle the adult level of uncertainty, therefore their production of articles is not adultlike, but they do not always omit articles either, so they are able to retrieve articles from a set. In the previous section I already alluded to the assumption that this is because they work with the same article set as adults, but with a different probability distribution. This same idea was pursued by Voinescu (1967) for phoneme production in Rumanian aphasia. She showed that aphasic patients do not lose certain phonemes, but use them statistically differently from normal adults.

Let us assume that children do not work with a different article set, but that they too produce them in a statistically different way and that they produce them with a more unequal probability distribution than normal adults do. This would imply that children would work with a set with lower entropy, an 'easier' set. As the child's brain develops, her ability to process a higher level of uncertainty will increase, and this will be reflected in a higher entropy value of the actually produced article set. The probability distribution of the elements in the set will become more adultlike. And, at the same time, as the child's ability to process higher uncertainty increases she will omit less. We can further hypothesize that, if it is correct to argue that omission is caused by the fact that the child does not have the processing resources necessary to process the high relative entropy values, we should find a relation between the child's development of article use and the child's developing capacity to handle higher entropy values. In other words, we should find a relation between omission rate and the relative entropy of the actually produced article set. The more the child's statistical use of articles (the probability distribution of articles used by the child) develops towards the adult's statistical use of articles, the higher will be the capacity of the child to handle relative

entropy values and the less omissions we should find. Of course, these are all assumptions, but we can test these assumptions with the child speech data we have.

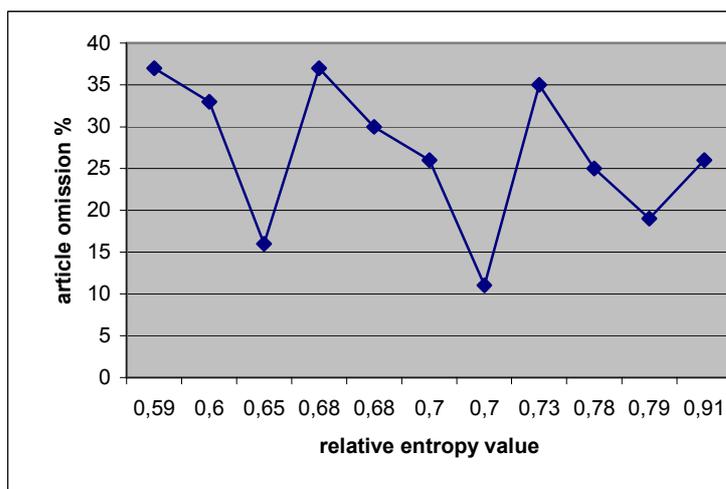
I have therefore calculated the relative entropy value of the set of articles children produce, based on their actual article production. This calculation is based on files with omission rates below 50%. The reason for this is that a reliable calculation of relative entropy of the files with omission rates above 50% is not possible, as in these files the articles are omitted in the majority of the contexts. If we focus on the files with omission rates below 50%, and calculate the relative entropy values of the child's actual production we find the relative entropy values shown in Table 13. For this calculation we made for each individual childfile a calculation of relative entropy in the same way as it was illustrated in Table 11 and discussed in section 4.3 for Dutch adult language. The only difference was that the frequency data were not the frequencies found in the Corpus of Spoken Language, but the frequencies found in the individual CHILDES files.

Table 13 Calculation relative entropy in individual childfiles, datapoints with omission% <50%.

File	Omission in %obl.contexts	Relative entropy children's determiners
P207	18	0,77
P208	26	0,90
T209	33	0,60
T210	37	0,59
T310	26	0,70
A210	37	0,68
A211	30	0,68
A301	25	0,78
S30	35	0,73
S36	11	0,70
S39	16	0,65

If we now plot a graph indicating the relation between incorrect article use and relative entropy in the set of articles realized by the child we will find:

Figure 9 *Relation between relative entropy values and the omission rate observed in the individual files of the children, (correlation coefficient $r = -0,281494286$)*



Let me explain this graph. Relative entropy is a measure of uncertainty, it reflects the level of uncertainty our processing resources have to deal with in order to select an element from a set. This implies that if a speaker is able to retrieve in a correct way elements from the set, the relative entropy of that set reflects the (minimal) capacity of the speaker's processing resources. As argued earlier, the fact that for example Dutch adults are under normal circumstances capable of producing articles without errors, implies that humans are able to deal with an uncertainty level of at least 0,94 (the relative entropy value of the Dutch article set). This means that we can take the relative entropy value for a measure of the (minimally) available processing resources. If children omit because of a lack of processing resources, we should find less omissions as their processing resources develop. Now if we use the relative entropy value of the child's article set as a measure of their available processing resources, we should find that as the value of relative entropy increases, omissions decrease.

However, the high fluctuations and the low correlation coefficient show that, contrary to our expectation, there is no relation between the

relative entropy value of the article sets the children produce and the omission rates. What causes these strong fluctuations?

The problem is that the files differ in their complexity: if a child, for example, is playing a card game with its mother ('kwartetten' *Happy Families-game*) in which she constantly uses utterances like: *Can I have the....? NO, I don't have the....* It will be far easier for the child to produce a high number of correct determiners than in the case of a freer conversation. In other words, the articles the child actually produces are not only influenced by the level of entropy of the article set, but also by the level of 'difficulty' that can be reached in the file. What I mean by this is, that if only 'de' is needed, in a particular file, then even an adult cannot reach the normal adult level of entropy (of 0.94, see Table 11) in that specific file.

The question to ask at this point is how we can measure the complexity of the file as far as article use is concerned? The answer is that we can do this by looking at *all* the children's utterances, so the ones in which they use the articles as well as the ones where articles have been omitted, and calculate the '*theoretical* relative entropy of the file'. This figure indicates what the relative entropy value would have been if the child had produced all the articles. If we do this, we find that indeed a file with the card game described above has a much lower theoretical relative entropy value than a file with a 'normal conversation'. To be more precise, the files with a 'kwartet-game' (like S36 and S39) have a theoretical entropy value of about 0,72, while a file with a more normal, natural conversation have a theoretical entropy value of about 0,90 (see Table 14).⁶⁹

These figures show that in an 'easy' file not even an adult speaker could come to an article production with a relative entropy value of more than 0,72. The specific discourse context somehow does not allow

⁶⁹ The fact that a 'normal' child file has a theoretical relative entropy value of about 0.9 is not surprising, as this is close to the relative entropy value found in the calculation on the basis of the Corpus Geproken Nederlands (Spoken Dutch Corpus), see table11 in this section. If we had found a different theoretical relative entropy value in all of the observed Childes files this would imply that children use 'de'- 'het'- 'een'- article requiring nouns in another probability distribution than normal adults do, and I see no reason for this. Of course this should not be confused with the *actual* realization of the articles *de*, *het* and *een*. In the actual realization of articles by children the probability distribution differs from the probability distribution we find in adult speech, because children make omission or substitution errors.

for more. But this also means that looking at the relative entropy of the determiner set the child realizes in a specific file is simply not enough. The data need to be adjusted for ‘the complexity of the file’. The child’s actually realized entropy has to be related to the theoretical entropy of the file. We should thus derive an ‘adjusted, weighed relative entropy’. After all, realizing all determiners in a file with a theoretical entropy of 0,70 is far easier than realizing them in a file with a theoretical entropy of 0,90.

Let me give an example of two files of Peter:

In P208 Peter’s actual use of the three articles is (in absolute numbers):

de 43
het 14
een 21

If we calculate the relative entropy of this particular set we find a relative entropy value of 0,90 (see Table 13).

Then, if we look at all the utterances of Peter in file 208, and count the article-requiring contexts we find:

de 57
het 35
een 29

The relative entropy value of this set is 0,96, and this is the ‘theoretical’ entropy value of the file, the value we would have found if Peter had produced all the necessary articles in an adultlike way.

Let us now look at another file: in P207 Peter’s actual use of the three articles is:

de 52
het 7
een 23

If we calculate the relative entropy of this particular set we find a relative entropy value of 0,77.

Then, if we look at all the utterances of Peter in file 208, and count the article-requiring contexts we find:

de 83
het 14
een 29

The relative entropy value of this set is 0,78, and this is the ‘theoretical’ entropy value of the file, the value we would have found if Peter had produced all the necessary articles in an adultlike way.

This shows that in this particular file Peter is able to deal with an entropy value of 0,77, but that this value is very close to the maximum 'score' he could have obtained in this file (0,78). It was a 'relatively' easy file. Hence, we cannot just say that in file 207 Peter was able to deal with an entropy value of 0,77 and in file 208 he was able to deal with an entropy value of 0,90 and that this means that he had less processing resources available in file 207. We have to weigh the values for the maximum complexity level of the files, and then we find for file 208: $0,90/0,96 = 0,94$ and for file 207: $0,77 / 0,78 = 0,98$. These adjusted values reflect the processing resources the child has available in those specific files.

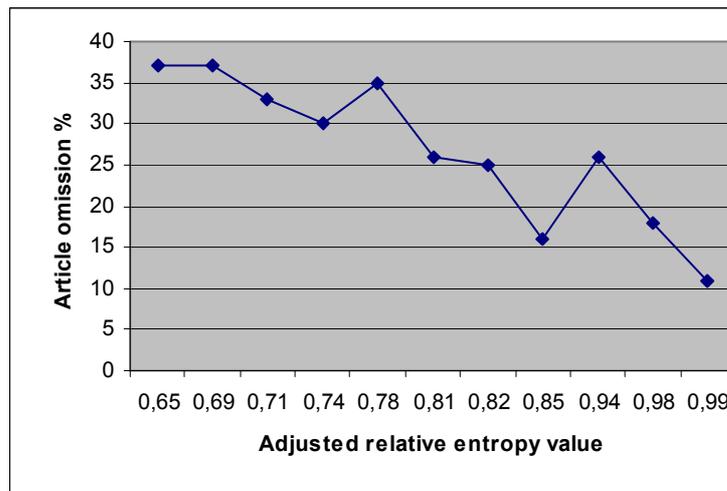
Adjusting the data in the manner indicated for all the files of the children, with omission rates < 50%, we get the following table:

Table 14 *Percentage of incorrect article use in the individual child files, the entropy in the actually produced article set, the theoretically possible entropy value of the files, and the weighed entropy value.*

File	Article Omission in %obl.contexts	Relative entropy children's determiners (A)	Relative entropy contexts in file (theoretical entropy) (B)	Adjusted Rel.Entr. (= A/B)
P207	18	0,77	0,78	0,98
P208	26	0,90	0,96	0,94
T209	33	0,60	0,84	0,71
T210	37	0,59	0,91	0,65
T310	26	0,70	0,86	0,81
A210	37	0,68	0,99	0,69
A211	30	0,68	0,91	0,74
A310	25	0,78	0,95	0,82
S30	35	0,73	0,95	0,78
S36	11	0,70	0,71	0,99
S39	16	0,65	0,76	0,85

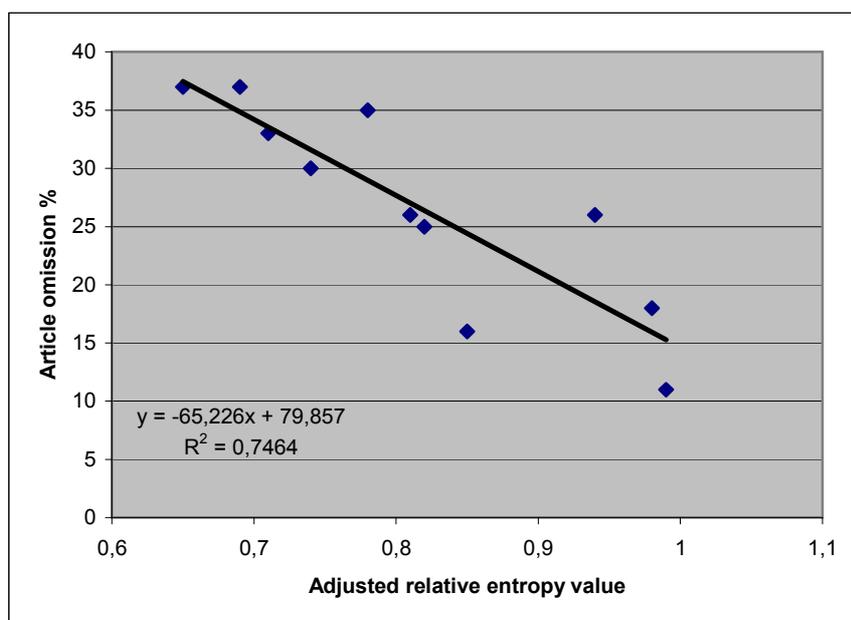
If we now plot a graph, with omission rates and adjusted relative entropy, we find:

Figure 10 *Relation between adjusted relative entropy values and the omission rate observed in the individual files of the children*



The calculation of the correlation between the adjusted relative entropy and article omission indicates a highly significant negative correlation ($r = -0,864$). The graph shows that there is a correlation between omission of articles and the entropy level a child can deal with. The regression analysis in Figure 11 shows the relation that can be established between the child's processing capacity and the omission rate.

Figure 11 Regression analysis of the relation between adjusted relative entropy and article omission %.



The interpretation of the regression analysis shown in Figure 11 is as follows:

- a. 74,64% of the omission variability can be accounted for by the adjusted entropy value variation.
- b. The slope of the linear function indicates that an increase by 0,01 of relative entropy value is paralleled by 0,65% decrease of article omission rate.⁷⁰

Hence, the amount of available processing resources of the child, which is reflected in the entropy level a child can deal with, is an important predictor of the omission percentage we will find in a particular file. The question may arise now whether this is a circular line of reasoning. One could argue:

⁷⁰ An interesting observation that can be made concerns the fact that the intercept indicates that the omission starting point is 79,86%. Kostić (p.c.) suggests that the implication of this fact may either be trivial, or that it may indirectly offer some insights into the default value of the child's initial processing capacity.

$$\begin{aligned}
 \text{omission rate} &= \text{omitted articles} / \text{obligatory contexts,} \\
 \text{production} &= \text{produced articles} / \text{obligatory contexts} \\
 \text{and thus: production} &= (\text{contexts} - \text{omitted articles}) / \text{obligatory} \\
 &\quad \text{contexts}
 \end{aligned}$$

But, this is not what my argumentation is about. Of course it is obvious that children produce more, when they omit less. But that is not what is indicated in the graph. The graph shows what we find if we look at a deeper level of the articles produced. We see that the probability distribution among the articles produced changes. Important for my argumentation is what these differences in probability distribution tell us. There is, in my view, no a priori reason why the probability distribution among the articles produced should change, when more articles are produced. But, the distribution *does* change, and it changes in one direction: the probability differences increase, the uncertainty increases. Moreover, there is a strong relation between these changes in probability distribution and the omission rate of articles. The adjusted entropy value is not a reflection of the number of articles produced, but of the probability distribution among the produced articles. And this probability distribution of articles actually produced shows the entropy level a child can actually deal with, and thus it reflects the child's processing capacity.

These data show that the better the child is able to process higher entropy levels, the less the child will omit articles. Hence, as the child's ability increases to use articles in a probability distribution that becomes more and more similar to the probability distribution of the adult article set, the omission rate decreases. And this shows, again, that omission of articles is caused by the child's inability to process high relative entropy values.

4.4.1.3 Article selection in headlines

In chapter 2 I argued that omissions in headlines can not be caused by limited processing resources. What is limited in headlines is not the channel capacity of the headline writers or readers (at least, it is not more limited than it is under normal circumstances), but the *time* readers want to spend on reading the headlines. They want to receive as much information as possible in a limited amount of time. A headline has to be a 'news-flash' with, preferably, a high surprisal and informative value. How can a headline writer reach his goal? In chapter 3 I argued that headline writers aim at providing the reader with the highest informational value for their cognitive effort. Dor (2003) expressed this

process of optimal headline construction (or ‘headlines encoding’) in a formula for ‘*relevance optimization*’:

- (10) Maximize $R = C/E$, where
 R = the relevance of the story for the reader
 C = the informational value
 E = the cognitive effort.

But we also noted that Dor could not give exact specifications, only intuitive suggestions about the values of these specific elements (see chapter 3, section 3.2.3).

Shannon and Weaver (1949) suggest that the optimal rate of encoding of information, R, is a rate that comes arbitrarily close to the channel capacity, C, but does not exceed it.⁷¹ The rate of encoding, R, is the *actual* amount of information that is sent through the channel per unit of time. The channel capacity, C, is the *maximal* amount of information that can be sent through the channel per unit of time. If the amount of information that is sent through the channel per unit of time (R) exceeds the channel capacity C, errors will occur.

Summarizing:

‘*encoding optimization*’ in Shannon’s model means:

- strive for a rate of encoding such that $R \leq C$, in which $C = I_{\max}/t$
- the closer R comes to C, the more effective the encoding is.

Let us now combine Dor’s intuitive approach with Shannon’s information-theoretical approach. The fact that in headlines not the available channel capacity, but the available time is reduced forces headline writers to make highly effective use of the channel in the available time. Therefore they will aim at maximizing the transmission rate (Shannon’s R), the number of informational units that are sent through the channel per unit of time. The closer R gets to C the more effective the use of the channel is, or: the higher the amount of information that is sent per unit of time. If time is limited, but one succeeds in transmitting more information per unit of time one can compensate for the loss of time. Why does the aim to maximize the transmission rate lead to omission of articles?

⁷¹ Note that Shannon’s R refers to the amount of information that is sent through the channel per unit of time. This should not be confused with Dor’s R, which indicates the ‘relevance of the story’.

What headline writers and readers are extremely sensitive to, because they want to maximize the effectiveness of channel use, is the processing time *per unit of information*. Writers do know that their ‘audience’, the reader, wants to receive the best informative value for their cognitive effort. The processing time per unit of information is influenced by the informative load of the element and the time necessary to select an element from the set ($= I/t$). In the previous section I argued that the informative load of articles, when compared to lexical elements, is low, in both, Italian and Dutch. Hence, we expect to find omissions, in both languages. And we do find omissions, but more in Dutch than in Italian. This crosslinguistic difference is caused by differences in the entropy of the two sets. Selecting an article from the Italian article set costs less processing time than selecting an article from the Dutch article set. Therefore, use of an article in Dutch leads to a stronger increase of the processing time per unit of information in the headline than use of an article in an Italian headline. Formulated differently, headline writers strive to minimize the processing time per unit of information (the encoding rate R), and use of an article in Dutch leads to a more perceptible ‘drop’ in processing time per unit of information than use of an article in Italian. That is why we find more omissions in Dutch headlines.

4.4.1.4 Conclusion

We have seen that in Dutch article selection is a more costly process than in Italian.⁷² This is caused by the fact that the degree of uncertainty in the Dutch article set, reflected in the value of relative entropy, is higher. The probability differences between the elements are lower in Dutch than in Italian. In the previous chapter we saw that besides an overall higher rate of omission (in child speech and the database headlines) and higher preference for omission (in the headlines experiment) the omission patterns also showed a number of particularities. They were influenced by finiteness, the sentence position of the article-requiring noun, and by the type of article: ‘de’ or ‘het’ in Dutch, ‘definite’ or ‘indefinite’ in Italian. In the following sections I will discuss these specific characteristics in the omission pattern.

⁷² In the remainder of this study I will shorten the notion ‘less costly process in terms of required processing resources and time’ to ‘less costly process’

4.4.2 How to account for finiteness and position effect?

Why are we more likely to find omissions in child speech in non-finite sentences than in finite sentences? Why are we more likely to find omissions in headlines in non-finite sentences? Why is the preference for omission in headlines higher in non-finite sentences? Why do we find more omissions in sentence-initial position? At first sight, the calculated information-theoretical values do not predict these patterns. The relative entropy value of the article set does not change depending on the sentence position of the article-requiring noun or due to the finiteness of the verb. How can we account for these patterns? Is it possible to account for them by a structural account, or do we need a processing account?

I will start with a brief overview of the data on the finiteness and position effects presented in the previous chapter:

Finiteness

In Dutch and Italian headlines we find more articles used or a higher preference to use the article if the headline contains a finite verb. In Dutch and Italian child speech we find that use of finiteness and use of articles are related over time: as children's use of finiteness matures, their use of articles matures as well. However, we saw that if we divide the child speech data over developmental stages it is not the case that children omit significantly more articles in non-finite than in finite sentences.

Position effect

-Position effect in child speech:

In Dutch child speech we find more omissions from sentence-initial position than from sentence-internal position. However, this difference reached significance only in sentences with finite verbs (and only in stage 2 and in the data of all stages mapped together).

In Italian child speech there was a tendency to omit less from sentence-initial position, but the difference did not reach significance.

- Position effect in the Headlines database:

The headlines database for Dutch showed the same effect we found for Dutch child speech: in sentences with finite verbs we find more

omissions from sentence-initial than sentence-internal position. Because of the low absolute number of article-requiring contexts in sentence-internal position in non-finite headlines, it is not possible to derive strong conclusions about the difference between omissions from sentence-initial and sentence-internal position in non-finite headlines. In the headlines database for Italian an effect of position was found in finite and non-finite sentences.

- Position effect in the Headlines experiment:

In Dutch in the experimental conditions with finite sentences there was a position effect. The preference for omission was higher for nouns in sentence-initial position. In the Italian experiment an effect of position was found in finite sentences, but only with indefinite determiners, not with definite determiners. There were no test conditions in which the position effect was tested in non-finite sentences.

In chapter 3 I already discussed my concerns with respect to a structural account of these observations. An alternative account of the position effect in Dutch child speech is the prosodic or metrical approach as it has been developed by Gerken (1991).⁷³ Although metrical accounts may play a role in the explanation of determiner omission in child language, a purely metrical account fails to explain why we find the same differences in headlines too.

In the remainder of this chapter I will try to show that it is possible to come to a unified account of this intricate cluster of findings. An

⁷³ Gerken argues that children frequently omit weakly stressed syllables from iambic (weak-strong) feet. Typically, determiners of sentence-initial subject DPs and sentence-initial subject pronouns appear in such iambic feet (see examples a and b). Sentence-internal object pronouns and determiners of sentence internal object DPs on the other hand often appear in trochaic (strong-weak) feet (see examples c and d).

- (i) a. He (weak) *kissed* (strong) Jane
 b. The(weak) *bear* (strong) *kissed* Jane
 c. *Pete* *kissed* (strong) her (weak)
 d. *Pete* *kissed* (strong) the (weak) lamb

Gerken suggests that children tend to apply trochaic templates to the feet they form during the speech planning process, given that trochaic feet are the most common foot-type in English (like in Dutch). An attractive aspect of Gerken's proposal is that it not only tries to explain why children omit more determiners from sentence-initial position than from sentence-internal position, but also why subject (pronoun) omission is far more frequent in English acquisition than object (pronoun) omission.

important observation we can make on the basis of the data presented in the previous chapter is that in both Dutch and Italian child speech and headlines, use of finiteness is a predictor of article use. If we first focus on child speech, the observed relation over time between use of finiteness and use of articles suggests that the difficulties children have with the realization of determiners and finiteness may be related to the children's reduced processing resources. I will explain this in more detail, using the model of article selection I proposed in chapter 2. This model describes the process of article selection of a simple NP like, for example 'an apple'. In many situations, however, the produced utterance is more complex and consists not only of a NP, but also a VP and possibly other NPs, PPs, etc., for example 'John eats an apple'. Of course, it is not the case that in order to formulate this utterance, we use three separate communicative intentions, one for the VP, and two for the NPs. The communicative intention is conceived for the complete 'event' (see Levelt, 1989, for a detailed discussion), and, in the case of 'John eats an apple' the grammatical encoder will build a surface structure that consists of a VP and two NPs.⁷⁴

In the model I proposed for the processing of a simple NP we saw that the grammatical encoding of a noun involves a so-called functional noun procedure which is stored in our mental lexicon and triggers the creation of a syntactic structure that is appropriate for the noun lemma. Every grammatical type of lemma has its own functional procedures, hence in a more complex utterance like 'John eats an apple' at the level of grammatical encoding we not only find a functional noun procedure, but also a functional verb procedure. This functional verb procedure will create the syntactic structure that is appropriate for the verb lemma, containing for example slots that have to be filled in with information necessary to select the inflection of the verb. Inflection, like articles, are functional categories, closed class words. As argued earlier in this chapter, the relative entropy of sets of closed words is higher than the relative entropy of open class words, and selection of a closed class word, like inflection is, therefore, a more costly process than selection of

⁷⁴ Whether the utterance consists of a simple NP, or whether it is a more complex 'event', like the VP given as example earlier: if we talk about the 'processing resources that are available' this availability concerns the complete NP or VP, and will therefore influence the processing of the complete processing unit, this is a simple and obvious observation, but I will return to this observation later.

lexical elements. If a child does not have the processing resources that are necessary to select the article from the article set, it means that the child does not have enough processing resources to process the (high) relative entropy value of articles. The effects of this limitation in processing resources, however, are very likely to influence not only the selection of articles, but also the selection of other elements coming from a set with high relative entropy. Thus, if we see the processing of an utterance like 'John eats an apple' as one processing unit, then it comes as no surprise that if the child has problems with producing the article, it is also more likely to have problems with verb inflection in the same processing unit. After all, also verb inflection is selected from a closed class set with a high entropy value. And, the other way round, if the child succeeds in processing the relative entropy value of the article set, and produces the article, it is more likely to be able to produce the verb inflection as well. Let me emphasize that these relations are described in terms allowing for optionality, like 'higher probability' and 'more likely'. There is no strong constraint that states that if a child cannot produce an article in an utterance, it will not be able to produce inflection either. It is only the case that, because of the limitations in processing resources, the likelihood of omission of articles is larger in the case of omission of inflection than in the case of use of inflection.

One could argue that a reverse line of reasoning is also possible. One could argue that leaving out inflection should increase the probability of occurrence of the article because of the fact that leaving out inflection 'saves' processing resources, and these can then be used for the selection of the article. This would imply however, that the processing resources that are available but not used in a particular stage of the processing of a sentence could be 'saved' for later use in the next processing step. I do not consider this to be a plausible option. The processes that take place at the level of syntax take place in a highly unconscious and automatized way. This implies that the speaker cannot intentionally or subintentionally control the use of the available syntactic processing capacity. Capacity that is not used will be lost. But since the language processor works very efficiently, this is not likely to occur. Thus, if the processing capacity necessary for inflection is available, inflection will be produced. And if inflection is not produced, this implies that not enough processing resources were necessary to process inflection. Thus we have two possibilities:

A. Inflection is used.

For use of inflection a higher level of activation is necessary, which

means that more processing capacity is available. This higher level of activation will not immediately drop down after the processing of inflection, and this increases the probability that an article can be selected.

B. Inflection is not used.

This means that the level of activation is too low to process inflection (otherwise inflection would have been used, capacity will not be ‘wasted’ in an automatized process). Thus, the probability that the available capacity that is necessary to select an article will be present is lower.

If this account is on the right track, two further questions have to be asked:

- 1. why do we find an effect of position: why do we find more omissions from a NP in sentence-initial position than in sentence-internal position? Since both the noun that precedes the verb and the noun that follows the verb belong to the same processing unit, we would not expect differences. A further question is why we find this position effect in Dutch only with finite verbs?
- 2. Why do we find an effect of finiteness in headlines? Headline writers do not have limited processing resources, and have no problems with processing high entropy values. Still we find a relation between article use and finiteness also in headlines.

Let me first answer the first question:

Why do we find a position effect in Dutch child speech only in finite sentences?

Language processing is an incremental process, applying from left to right as much as possible (see Philips, 2003, for a syntactic account, Levelt, 1989, for a processing account, for evidence from eye-tracking studies: Kamide et al. (2003), Traxler et al. (1997, 1996)). As I argued before, processing an inflected verb implies selecting an inflected form from the inflection set, with a high entropy. This requires a high level of brain activation in the child. Therefore, processing inflection forces the child’s brains to increase the level of activation, to work ‘faster’. It is plausible to assume that this extra brain activation will not immediately dissipate after the inflected verb has been selected. It is reasonable to argue that the brain activation level directly after the processing of an inflected verb is higher than it is at sentence-initial position.⁷⁵ This

⁷⁵ In the same vein we could hypothesize that the presence of case/agreement features (inflectional features) on the first NP will provide a facilitation effect for the processing

higher brain activation level directly after an inflected verb provides a facilitation effect on the processing of the element directly following the inflected form. In other words: the brains are already in a state of higher activation, because they were ‘triggered’ to work faster in order to process inflection, and this higher activation level makes it easier for the brains to process articles on a noun following an inflected verb. The threshold of activation necessary for the activation of the next element can be reached faster because of the higher activation level, higher processing speed and reduced processing time per unit of information.

Evidence that the level of syntactic encoding influences the level of brain activation comes from a study by Indefrey et al. (2001). With the use of positron emission tomography (PET) the cortical activations (cerebral blood flow) that are related to the syntactic encoding process during normal spoken language production were investigated. Indefrey et al. used scene descriptions to elicit utterances varying in the complexity of syntactic encoding. Three levels of syntactic complexity were distinguished:

- full sentences, like for example:
 - (11) **Das rote Viereck stösst die blaue Ellipse weg.**
The red square launches the blue ellipse.
- non-finite clauses with a phrasal structure, like for example:
 - (12) **Rotes Viereck Blaue Ellipse wegstossen.**
Red square, blue ellipse, launch.
- non-finite clauses without phrasal structure: sequences of single words, having no syntactic relationship, like for example:
 - (13) **Viereck, rot, Ellipse, blau, wegstossen.**
Square, red, Ellipse, blue, launch.

of inflection on the following verb. That is, for child speech we would predict that if case/agreement features are realized on the first NP, the probability that the following verb is a finite verb increases (and thus, the probability that the noun following the (finite) verb contains an article increases

The response types differed in the degree of syntactic encoding and the corresponding application of grammatical markers (printed in bold) in German. Local gender agreement marking on the adjective was required in noun phrase and sentence responses, but not in the single-word-responses. Only in the sentence condition (example 11) did syntactic relations across several words have to be expressed by means of word order and inflection of the main verb. The results showed a neural correlate of syntactic encoding during speech. Interestingly this syntactic activation showed a graded response related to the syntactic complexity: most activation was measured during the production of full sentences, less activation was measured during the production of non-finite clauses with phrasal structure and the least activation was measured during the production of the sequences of single words, without phrasal structure.

This study shows that it is plausible to argue that the processing of inflection leads to higher brain activation. This activation level will not decrease directly after the processing of the finite verb: there is a gradual decay of the activation level of the brain. Therefore, the brain activation level directly after the processing of finiteness will be higher than at for example sentence-initial position, and this makes it easier for the speaker to reach the activation level required to retrieve the article. In fact, this finding is also compatible with the findings of Kostić, discussed in section 4.2. Recall that he found a relation between the informative load of the context (the global informative load) and the processing time per unit of information. In his experiments the processing time per unit of information decreased as the global informative load increased. If the context is more 'complex', a higher level of brain activation is necessary to process this higher complexity level. Once the brain has reached this state of higher activation, this will not immediately drop down. It can therefore be expected that the processing time per unit of information decreases in a context with a higher global informative load where the brain is already in a state of higher activation.

If this argumentation is correct we should expect to find no effect of position if the verb in the utterance is non-finite (no inflected form is selected there, hence, no higher brain activation). And this is what we found in the analysis of Dutch child speech. The fact that we did not find a significant effect in Italian child speech also fits in this approach. We saw that selecting an article out from the Italian article set is a less costly process in terms of processing resources, than in Dutch. Hence, in Italian the facilitation effect offered by the selection of an

inflected form prior to the selection of an article is not necessary, and therefore, does not result in (visible) effects.

We now turn to the second question: why do we also find a finiteness effect in headlines? And why is this effect stronger in Italian than in Dutch?

It is not possible to account for this effect with a structural approach. The problem with structural approaches is that they predict a strong relation between the use of finiteness and the use of the article.⁷⁶ They predict that if finiteness is used, the preferred version of the headlines should be the one with the article present, if finiteness is omitted, the preferred version should be the one with the article omitted. This is the pattern we find in Italian, but not in Dutch. In Dutch the preferences for omission do change in the presence of a finite verb, but still the preferred headline is the one with the article omitted. Hence, structural accounts make the right predictions for Italian, but not for Dutch.

Instead I want to propose an account analogous to the account proposed for child speech, but not with processing resources, but with processing time as inhibitory factor. Let us take a look at what happens with the use of the channel capacity when a finite verb is processed. Recall that in Shannon's model the actual use of channel capacity, R , could be measured as $R = I / t$ (the actual amount of information that is sent through the channel per unit of time). The maximal channel capacity, the maximum amount of information per time unit that can be processed by the channel = $C = I_{\max} / t$. The closer R comes to C , the more effectively the channel is used. Processing inflection is a costly process. For normal adults this is not problematic since they do have the brain capacity to process the level of uncertainty, but there is a price that has to be paid for use of articles, and use of inflection. The effectiveness of the channel use, measured as the actual amount of information that is sent through the channel per unit of time, decreases. Let us now look at the headlines used in the test conditions:

- (14) A. VREDESPLAN VOOR HAÏTI VERWORPEN
Peaceplan for Haiti rejected
HET VREDESPLAN VOOR HAÏTI VERWORPEN
The peaceplan for Haiti rejected

⁷⁶ For example Landau (2004) Guéron & Hoekstra (1995) Enç (1987), could eventually be applied to the use of tense in headlines. See De Lange et al., 2006, for a discussion of an approach based on Landau (2004).

- (15) B. VREDESPLAN VOOR HAÏTI IS VERWORPEN
 The peaceplan for Haiti is rejected.
 HET VREDESPLAN VOOR HAÏTI IS
 VERWORPEN
 The peaceplan for Haiti rejected

As the examples show, there were two versions of the test sentences. One version (see example 14, the ‘A –versions’) in which the finite auxiliary verb was omitted (and the article present in one sentence and omitted in the other), and one (see example 15, the ‘B-versions’) in which the finite auxiliary verb was used (with again the article present in one sentence and omitted in the other). The two versions A and B were minimal pairs, except for the presence/omission of the finite auxiliary.

As I argued above, use of the auxiliary leads to a decrease in processing speed, since it has to be selected from a set of high relative entropy (functional) elements. Therefore we can say, even without giving concrete details about the exact values, that the average processing speed per unit of information of the A version of the headlines (without the finite auxiliary) will be higher than the average processing speed of the B version of the headlines (with finite auxiliary). There is another way to arrive at the same conclusion. We can give an accurate estimate of the relation between the *average* informative load of the headlines in version A and version B. In principle the *average* informative load of a headline can be calculated by summing up the informational loads of the individual elements in the sentences and dividing this number by the sum of elements (see section 4.2.5, and Kostić, 2004). Now, even without knowing the exact informative values of the individual elements in the A version and the B version, we can easily see that whatever these values are, the average informative value of the A versions will be higher than the average informative value of the B versions, because A and B are minimal pairs, which only differ in one element: the presence/absence of a (low informative, functional) element, the auxiliary. The informative load of an auxiliary (coming from a class of closed class elements) is lower than the informative load of lexical elements. Hence, the average informational value of the A version sentences is higher than the average informational value of the B version sentences⁷⁷. Kostić has

⁷⁷ In an abstract formula: suppose
 informative load ‘vredesplan’ (peaceplan) = A

shown that the processing time per element in the sentence not only depends on the informative load of the element itself, but also on the average informative load of the context ('global informative load'). This global informative load is higher in the A versions than in the B versions. Therefore in the A version the processing speed per unit of information is higher.

Thus, the actual channel use, expressed as $(R) = I / t$, is closer to the maximum capacity level in the A versions than in the B versions. This means that in the B versions there is more channel capacity left to process the article than in the A versions, and therefore the fact that in the B versions the preference for omission is less strong than in the A versions is not surprising.

In other words, in the A versions the channel is used very effectively, with an actual channel use $(R) = I/t$ close to the maximum capacity level. In the B versions the channel is used less effectively than in the A versions. The use of the channel in the B versions is further away from the maximum capacity level than in the A versions. Therefore in the A versions use of an article leads to a stronger, and hence more perceptible, decrease in the effectiveness of channel use (processing time per unit of information) than in the B version, in which the channel use had already diverged from the maximum capacity level. Consequently we find a stronger preference for omission of the article in the A versions.

Of course it is true that use of an article leads to a less effective use of the channel, and hence to a decrease in the rate of encoding in both versions. This is reflected in a preference for the headline in which the article is omitted in both versions (in Dutch for all article types, and in Italian for indefinite articles). But, the preference for the headline in

informative load 'voor' (for) = B

informative load 'Haïti' = C

informative load 'verworpen' (rejected) = D

informative load auxiliary 'is' = E

Then average informative load of the version without auxiliary 'VREDESPLAN VOOR HAITI VERWORPEN' = $(A+B+C+D) / 4$.

Average informative load of the version with auxiliary: 'VREDESPLAN VOOR HAITI IS VERWORPEN' = $(A+B+C+D+E) / 5$.

$(A+B+C+D) / 4 > (A+B+C+D+E) / 5$ if $5(A+B+C+D) > 4(A+B+C+D+E)$, hence, if: $(A+B+C+D) > E$.

Since E is an auxiliary, coming from a set of functional elements with a lower informative value than the lexical elements A, C and D it is highly plausible to assume that the sum of A, B, C and D will be higher than E.

which the article is omitted is stronger in the version in which the auxiliary is omitted than in the version in which the auxiliary is present.

Formulated differently, if it is correct to assume that the available processing time is the limitation that influences omission in headlines, we can now say that if the speaker has enough time available to process the finiteness, then the probability increases that he will have enough time available to process the article. And this argumentation is analogous to the one used to account for the relation between finiteness and omission in child speech, where not time but available processing resources were the bottleneck: if the child has enough processing resources to process finiteness, then the probability increases that there will be enough processing resources to process the article.

In this same vein we can also account for the preposition effect. Like finite auxiliaries, prepositions are closed-class elements. Hence, elements with a low informative value, which will slow down the processing speed (processing time per unit of information) of the headline. This will lead to a less strong preference for omission of the article directly after a preposition.

One question remains: why is the effect of finiteness (and preposition) stronger in Italian? Why does use of finiteness in Italian lead to a more drastic change in preference? If finiteness is used, the version with the article used is preferred, if finiteness is not used, the version without the article is preferred. The answer should follow straightforwardly by now. In Italian, as in Dutch, use of finiteness leads to a decrease of the average processing speed, and to a rate of encoding that is further away from the maximum capacity level. Thus, more channel capacity becomes available to process the article, like in Dutch. But processing an article in Italian costs less processing time and effort than processing an article in Dutch. Therefore, the channel capacity that has become available because of the decrease in processing time per unit of information by using finiteness is enough in Italian to allow for selection of the article. In Dutch, in spite of the fact that channel capacity has become available, selection of the article still costs too much time and processing resources. Therefore, the general preference remains the preference to omit, but it becomes less strong.

In the same vein, it is now possible to account for another effect observed in the results of the headlines experiment, namely the 'order of omission' effect we tested. Following Stowell's observation on omission in English headlines we tested whether participants, when forced to

make a choice between a headline in which the article was used before the subject and omitted from the object, or a headline in which the article was omitted from the subject and used before the object would reject the one with omission from the object, as Stowell predicted.

Example (16) illustrates the test items. The prediction made on the basis of Stowell's observation for English was that participants would strongly reject (a).

- (16) a. DE KRAB VEROVERT NOORDZEE
The crab conquers Northsea
b. KRAB VEROVERT DE NOORDZEE
Crab conquers the Northsea

In chapter 3 we saw that Stowell's predictions were not confirmed. What we found was at most a stronger preference for the B version. And in Italian the participants did not even have a preference.

Note that the participants were forced to make a choice between an article-containing NP *before* an inflected verb and one *directly after* an inflected verb. The Dutch participants would undoubtedly have preferred a headline with no article at all. But, forced to make a choice, the Dutch participants preferred the B-version of the headlines. This means that they preferred the version with the article directly following the inflected verb above the version with the article at the beginning of the sentence. In my view, the reason for this preference lies in the fact that, as I argued before, at sentence-internal position, after the inflected verb, the processing speed per unit of information has decreased because of the processing of inflection. This has reduced the effectiveness of the use of the channel. Thus, after the processing of inflection the actual use of the channel is further away from the maximum channel capacity. In other words, at the point where the processor works on the (bare) noun that precedes the inflected verb (CRAB, in example 16 b) the channel is used very effectively, with an actual channel use $(R) = I/t$ close to the maximum capacity level. This is what headline readers expect: process as much information as possible as fast as possible. Subsequently, because of the processing of the inflected verb, the channel is used less effectively. This means that the use of the channel after the processing of inflection is further away from the maximum capacity level than it was before the processing of inflection. Therefore use of an article before a noun directly *following* the inflected verb leads to a less strong, and hence less perceptible, decrease in the effectiveness of channel use (processing

time per unit of information) than use of an article before a noun that *precedes* the inflected verb. Therefore, if an article *has* to be present in a headline - in other words, if the channel *has* to be used in a less effective way - then preferably before the noun that follows the inflected verb, since at that point the channel is already in a state of less effective use. Processing inflection drives the channel use away from the maximum capacity level. Therefore there is more channel capacity available for processing the article before a noun that follows the inflected verb than before a noun that precedes the inflected verb. The results would suggest that the effect of position appears to be a side-effect of finiteness.

4.4.3 Differences between article types

The results of the analysis in the previous chapter showed that not all articles have an equal probability of being omitted. There are intriguing differences between articles, in particular between omission of ‘*de*’ and ‘*het*’ in Dutch child speech, and between omission of *definite* and *indefinite* articles in Italian headlines. In this section I will show how we can account for these differences in the model adopted here.

4.4.3.1 The difference between ‘*de*’ and ‘*het*’ in Dutch

The analysis of the Dutch child speech data confirmed the results of previous studies on the differences between omission of ‘*de*’ and ‘*het*’ in Dutch child language: Dutch children omit ‘*het*’ significantly more often than ‘*de*’ (see chapter 3, section 3.5.3.4). As Table 15 shows there is also a difference in the age of first use of ‘*de*’ and ‘*het*’. All children start to use ‘*het*’ later than ‘*de*’.

Table 15 *Age of first use of the different articles*

	DE	HET	EEN
TOM	2;02-01	2;05-07	2;02-01
ABEL	2;01-02	2;10-00	2;01-02
PETER	2;02-03	2;04-19	2;00-07
SARA	1;11-01	2;02-28	1;11-01

I will explain how we can account for this in my model. It is important to note that we are talking now about the processing of different elements within the *same* set. We are not, as we did in the previous section, comparing elements coming from two different sets with different numbers of elements. We can thus directly follow Kostić' approach, discussed in section 4.2.5. Kostić focused on the processing of inflectional elements coming from one and the same set. He compared the time necessary to process these elements, and found that, within the set, the processing time was related to the informative load of each element. The higher the informative loads of the element, the longer the processing time. Let us follow his approach and compare the informative load of the different articles in Dutch. We saw in section 4.2.5 that the informative load can be calculated with the formula repeated in (17):

(17) Calculation Informative Load:

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

Important in this calculation is the number of functions and meanings. Looking at the Dutch article set, we can distinguish the following number of functions and meanings:⁷⁸

- DE =
1. sg/common/def.
 2. pl/common/def.
 3. pl/neuter/def.

⁷⁸ The definition of the number of function and meanings, as Kostic already suggested for the definition of functions and meanings in Serbian, will remain an arbitrary issue as it depends on one's theoretical viewpoint. However, what is important is not the *absolute* number of functions and meanings, but the proportion of functions and meaning encompassed by a specific form, relative to other forms, and this proportion is less sensitive to the definition of functions and meanings

- 4. dim/pl/common/def
- 5. dim/pl/neuter/def

HET =
 1. sg/neuter/def
 2. dim/neuter/def
 3. dim/common/def

EEN =
 1. sg/common/indef.
 2. sg./neuter/indef.
 3 dim/common/indef
 4. dim/neuter/indef

If we calculate the informative loads of the individual articles, taking into account the number of functions and meanings listed above, we derive the results in Table 16

Table 16 Calculation informative load individual articles, based on the numbers of functions and meanings.

	Frequency (F) (corpus 'Gesproken Nederlands)	number of functions/ meanings (R)	av.freq.per funct/m.= F/R	(F/R det) / (sum F/R paradigm) (p)	$I = -\log_2 p$
De	253210	5	50642	0,397096	1,332439
Het	96327	3	32109	0,251775	1,989796
Een	179119	4	44780	0,351129	1,509927

The table shows that the informative load of 'het' is higher than the informative load of the other Dutch articles. In his study Kostić found that a higher informative load of an element implies longer processing time. The higher the informative load, the longer it takes to process the element. It is therefore not surprising that children, because of their limited processing resources, are sensitive to the higher informative load of 'het', and that this leads to more omissions and a later age of first use.

A few additional questions arise. The table not only shows a difference between the informative loads of *'de'* and *'het'*, but also between the informative loads of *'de'* and *'een'*. Doesn't this difference lead to differences in omission in Dutch child speech? And don't we find differences between the informative load of individual elements in Italian as well? Should this not also lead to differences in acquisition age and omission pattern of the different articles?

I will start with the question on the difference in informative load between *'de'* and *'een'* in Dutch. It turns out to be problematic to make a proper comparison of the use and omission of definite and indefinite articles in child speech on the basis of the spontaneous data in CHILDES. It is not always possible to determine whether a definite or indefinite article has been omitted in spontaneous speech data. It is easier to see that an article has been omitted (regardless of the type) than to decide which type of article had to be used.

It is more useful to investigate the correct use of definite and indefinite articles by experimental tasks. A large number of experimental studies on a variety of languages have shown that children have problems with the correct pragmatic use of indefinite articles and overuse definite articles in contexts where indefinite articles would be appropriate (see Schaeffer and Matthewson, 2005, and references therein). However, attributing this overextension of definite articles to differences in informative load between definite and indefinite articles by arguing that children overuse the definite article *'de'* because of its lower informative load, and therefore, lower processing cost, would be a wrong application of the model I propose. As I already argued in chapter 2, discourse pragmatic problems with definiteness have to do with another level of the communicative process than my model for article selection. They relate to the level of the conception of the communicative intention. My model is concerned with the problems that arise at the level of syntax, when the selection of elements from the article set takes place, on the basis of the information in the slots of the article frame. It is not concerned with the question whether the information in the communicative intention was 'correct' from an adult point of view. Selection of *'de'* and *'het'* are based on the same discourse pragmatic situation. Both are definite determiners, so the same communicative intention will serve as input for the Syntactic Formulator. Thus, the formulation of the communicative intention of *'de'* and *'het'* require the same discourse pragmatic abilities. In other words: if the child is able to formulate the communicative intention that will lead to selection of *'de'*,

she will also be able to formulate the communicative intention that will lead to selection of *'bet'*. Therefore, differences in omission patterns between *'de'* and *'bet'* (more omissions of *'bet'*) cannot be caused by problems with the formulation of the communicative intention in the case of *'bet'*, and have to be captured at the level where articles are selected from the article set, where differences in informative load cause a higher omission of *'bet'*.

However, if we compare the selection of the definite article *'de'* and the indefinite article *'een'*, we are comparing articles that are used in different discourse pragmatic situations. Distinguishing between these situations is known to be problematic for children, they have particular problems with the formulation of the adult-like communicative intention for indefinite determiners. Hence, the formulations of the communicative intention of *'de'* and *'een'* require different discourse pragmatic abilities, in particular the input for *'een'* may be defective. Therefore a difference in omission pattern between *'de'* and *'een'* may, but need not necessarily be caused by the differences in informative load. There are other differences in the process leading to selection of the article that may influence the selection process.

What about the informative load (and differences therein) of the individual articles in Italian? What are the consequences of the differences in informative load between the individual articles in Italian? Table 17 shows the calculation of the informative loads of the individual articles in Italian, based on the assumption that they have one function and meaning.⁷⁹

⁷⁹ It could be argued that *'expletive'* should be seen as a separate meaning of definite articles, but this does not lead to qualitative differences in informative load, as it does not lead to large differences in the relative proportion of functions and meanings of the articles, see the table in Appendix B.

Table 17 *Calculation informative load individual articles Italian.*

	Frequency (=F) (corpus De Mauro)	number of functions/ meanings (=R)	av.freq. per funct/m. = F/R	(F/R det)/ (sum F/R paradigm) (p)	I = -log ₂ p
il	7111	1	7111	0,195351776	-2,3558537
la	7254	1	7254	0,19928024	-2,3271294
le	2536	1	2536	0,069668416	-3,8433514
lo	425	1	425	0,011675503	-6,4203714
i	2637	1	2637	0,072443065	-3,7870086
l'	3120	1	3120	0,085711931	-3,5443602
gli	820	1	820	0,022526854	-5,4722104
un	6624	1	6624	0,181973023	-2,4582035
un'	800	1	800	0,021977418	-5,5078343
una	3838	1	3838	0,105436664	-3,2455515
uno	227	1	227	0,006236092	-7,325142
degli	111	1	111	0,003049367	-8,3572746
dei	394	1	394	0,010823878	-6,5296386
del	18	1	18	0,000494492	-10,981765
della	15	1	15	0,000412077	-11,2448
delle	470	1	470	0,012911733	-6,2751735
dello	1	1	1	2,74718E-05	-15,15169
dell'	1	1	1	2,74718E-05	-15,15169

The table above shows that there are in fact considerable differences between the informative loads of individual articles in Italian. Let me now compare these differences with the results of the analysis of the Italian child speech data, based on the CHILDES files discussed in the previous chapter.

I will start with the articles that in the overview of informative load values have the highest informative load: the partitive articles (*del, della, dello, etc.*). If the value of the informative load were the only predictor for the omission of articles, we would expect higher omission rates for these articles than for the other Italian articles. Table 18 shows the omission rates of partitive articles, compared to the omission rates of all articles, in all stages, and in the stages 1 and 2 separately.

Table 18 Omission rate of partitive articles in Italian child speech, compared to omission rate all articles in Italian child speech, in all stages, and in the stages separately.

Developmental period	Omission rate partitive articles	Omission rate all articles	
Stage 1	66,7	58,3	Fisher's exact: p=0.6546
Stage 2	16,7	16,9	Fisher's exact: p=0.7058
All stages collapsed	26,7	38,2	Fisher's exact: p = 0.5542

The results in Table 18 show that we find no significant differences between children's omission of partitive articles and their omission rate of all articles. This is not what is expected on the basis of the informative load values of the partitive articles. To explain this it is important to note first of all that there is only a very small number of partitive contexts of use in the files used in our study (total number of partitive article contexts: all stages together 13, stage 1: 3, stage 2: 10). This small number makes it very difficult to derive strong conclusions about children's use and omission of partitive articles. However, there is a reason for this small number of contexts in which partitives are used. In my explanation for Dutch I argued that definite and indefinite articles cannot be compared by using only informative load values, as they differ in the communicative intention that serves as their input. The same holds for partitive determiners in Italian. The formulation of the communicative intention of a partitive determiner requires other discourse pragmatic abilities of children than the formulation of the communicative intention of other articles.⁸⁰ It is therefore not surprising that children use them less frequently than other articles. However, again, it would be wrong to attribute these differences in use to differences in informative load. Differences in informative load play a

⁸⁰ As argued in the introduction, a partitive article is used with mass nouns to indicate an unspecified quantity or part of the whole denoted by the noun, as in for example: *C'è dell'acqua dentro la bottiglia* (There is (some) water in the bottle).

In the plural the partitive article can indicate an unspecified quantity or part of the whole denoted by the plural noun: *Ci sono delle mosche dentro la bottiglia* (There are (some) flies in the bottle) (Maiden and Robustelli, 2000).

role at the level of article selection, on the basis of the filled-in article slots. Children do not use partitive determiners often, especially not in the first stage of their development. When their cognitive ability grows to use the pragmatic contexts that require the use of a partitive determiner, they have already reached the stage where they can process the higher informative loads of the articles. Hence, at the stage where they can produce the communicative intention necessary for the use of a partitive determiner, they are also at a stage where they can select the article from the set.

There is another group of articles with relatively higher informative values than the other articles, namely the articles that are used with special phonological contexts, like ‘*lo*’, ‘*gli*’ and ‘*uno*’. Use and omission of the indefinite article ‘*uno*’ will be influenced by other factors than informative load alone. ‘*Lo*’ and ‘*gli*’, however, are both definite articles. Table 19 shows the omission of ‘*lo*’ and ‘*gli*’ in comparison with the omission of all articles.

Table 19 *Omission of ‘lo’ and ‘gli’ in comparison with the omission of all articles.*

Developmental period	Omission rate <i>lo, gli</i>	Omission rate all articles	
Stage 1	85,7	58,3	Fisher’s exact p=.2403
Stage 2	35,7	16,9	Fisher’s exact: p=.0618
All stages collapsed	52,4	38,2	$\chi^2 = 7.846$, p =.0051

These results show that children do make more errors with ‘*lo*’ and ‘*gli*’ than with the other articles, as is predicted by the differences in informative load. Hence, in Italian too, children are sensitive to the differences in informative load of the individual articles.

4.4.3.2 The difference between definite and indefinite articles in Italian headlines

In the previous chapter we saw that in the headlines experiment a difference was found in the preferences for omission in conditions with definite and indefinite articles. In Italian the preference for omission was higher in the case of indefinite articles. In Dutch there was a slight tendency towards a stronger preference of the indefinite article, but the

difference did not reach significance. Normal adults, like headline writers, do not have problems with the formulation of the communicative intention of indefinite articles in normal circumstances, and it does not seem plausible to argue that this process will somehow suddenly become problematic in headlines. Therefore, the intriguing question remains why we find a stronger preference for omission of indefinite articles in headlines.

Let me start with an important observation. In the condition with definite articles the nouns were inherently unique nouns, like '(the) Italian government', (the) Queen, (the) Dutch drug-policy'. Inherently unique nouns are associated more often with a definite determiner than common, not-inherently unique nouns. The probability of a definite determiner with a noun like 'prime minister' is higher than the probability of a definite or indefinite determiner with a noun like 'journalist'. Therefore, the strength of association between 'prime minister' and a definite determiner is stronger than the strength of association between 'journalist' and an (in)definite determiner, and this strength will lead to faster retrieval of the definite determiner in the case of an inherently unique noun.⁸¹ Formulated in terms of the model I propose, using an inherently unique noun increases the probability of a definite determiner. It influences the probability distribution of the elements in the article set associated with the noun. There is less uncertainty about the type of article that has to be used, and we will thus find that the relative entropy of the article set, in the specific context of an inherently unique noun, decreases. Hence, what in Anderson's (see footnote 25) proposal is called 'increase of strength of association' is, formulated in the terminology of the model I propose, an increase of probability (or decrease of uncertainty), which will lead to a faster retrieval time of the definite article in the case of an inherently unique

⁸¹ See Anderson & Reder (1999) and references therein for a series of experiments on the relation between strength of association and reaction time. They found that the higher the probability that in the past fact i occurred when concept j was present, hence, the stronger the strength of association between fact i and concept j , the faster the reaction time was. Translated into our account of articles: the probability that fact i (read: definite article, 'the') occurs when concept j (read: inherently unique noun, for example 'prime minister') is present is higher than the probability that an indefinite article occurs when a noun like for example 'journalist' is present. Thus, the strength of association between 'prime minister' and 'the' is higher than the strength of association between 'journalist' and 'a', leading to a faster retrieval time of 'the' with 'prime minister'.

noun. The goal of a headline is to provide as much information as possible in a limited amount of time.

I argued that retrieval of a definite article in the case of an inherently unique noun is faster than retrieval of an indefinite article. Hence, in the formula that expresses the actual amount of information that is sent through the channel per unit of time ($R = I/t$) t will be lower in the case of a definite article of an inherently unique noun than in the case of an indefinite article. If we assume that the amount of information that is added by the use of an article is the same for definite and indefinite articles, this means that the increase of I in the formula ($R = I/t$) is the same for both types of articles.⁸² Therefore a definite article, if used, will have less effect (compared to indefinite articles) on moving R (the *actual* channel use) away from the maximal channel use. Therefore we find a higher preference for omission of indefinite articles than definite articles.

Why do we find crosslinguistic differences? With an inherently unique noun the relative entropy of the article set associated with the noun will change because of the different probability distribution of the articles. I assume that the effect of inherent uniqueness of the noun on the probability distribution of the articles is the same in Dutch and Italian (I see no reason to assume that there are differences between the probability with which 'de' ('the') occurs before 'zon' ('sun') in Dutch and 'il' before 'sole' in Italian). Therefore the decrease in entropy will be the same in both languages. But since the relative entropy (in the article set in general) is higher in Dutch than in Italian, it will also be higher after the decrease caused by the difference in probability distribution in the case of an inherently unique noun. Thus, selecting an article still takes longer in Dutch than in Italian. Hence, we will still find that the

⁸² It may even be the case that a definite article in the case of an inherently unique noun like 'the sun' has a lower informative value than an indefinite article in for example 'a journalist'. After all, the definite article in 'the sun' adds, because of its 'predictability', less information than an indefinite article. Thus, the reduction of uncertainty offered by the use of a definite article is less in the case of 'the sun' than in the case of 'a journalist'. Future research may enable us to calculate in a precise way in how far the informative value of an article depends on the specific noun with which it is associated, for example by using a measure like 'conditional entropy'. The conditional entropy of a random variable, given another random variable, shows how the second affects the uncertainty (and thus, the required processing resources/time) of the first (see for example Clark, 2001). What is important for my argumentation is that it is not plausible to assume that the informative value of an indefinite article before for example 'journalist' will be *lower* than the informative value of a definite article before an inherently unique noun like 'sun'.

retrieval time t of an article in Italian will be less than the retrieval time t of an article in Dutch. A headline writer strives to maximize $R = I/t$. Use of a definite article leads to an increase of I (the information that is added by the article), and to an increase of t (the time necessary to retrieve the article). The small increase of I in Italian may be compatible with the resulting increase in t , and therefore we find a very low preference for omission. In Dutch the increase in t may be disproportionately higher than the increase in I . Formulated differently: the increase in t may be higher than the increase in I . This will lead to a reduction of the amount of information that is sent through the channel, expressed as I/t , and therefore lead to a less optimal channel use. This explains why we find a preference for omission in Dutch, but not in Italian.

Let me illustrate this with a simplified example of the NP *la regina / de koningin* ('the queen') using fictitious data. Let us assume that (in Dutch and Italian) the noun '*regina/koningin*' contains 12 bits of information ($=I$) and can be processed in 2 ms ($=t$). Hence, the amount of information sent through the channel per unit of time (I/t) when the noun is produced is 6 bits of information. Then, let us assume that use of a definite article leads to an increase of I with 3 bits of information, in both languages. Suppose that retrieval of a definite article costs 0,40 ms in Italian and 1,00 ms in Dutch. Thus, the amount of information sent through the channel per unit of time in the Italian NP '*la regina*' is then: $(12 + 3) / ((2 + 0,4) = 6,25$. Hence, the amount of information sent through the channel per unit of time increases, and for the Italian headline writer striving to maximize the effectiveness of channel use there is no reason to omit the article. In the Dutch NP '*de koningin*' we find that the amount of information sent through the channel per unit of time will be $(12+3)/(2+ 1) = 5$. Hence, the amount of information sent through the channel per unit of time decreases. That is, for the Dutch headline writer there is every reason to prefer the version without the article. Of course, these are fictitious data. The actual data may very well be less pronounced. It is, for example, imaginable that the outcome will be that for Italian the amount of information that is sent through the channel per unit of time, if the article is used, will be 5,9. Hence, a little bit lower than in the case of use of only the bare noun (6). However, the value found for Italian will always be higher than the value found for Dutch, and closer to the maximal channel capacity than the value found for Dutch. The question then of course is: if the amount of information sent through the channel per unit of time is always lower when the

article is used, even in Italian (even if it is just 0,1 bit lower), why then don't Italian headline writers omit the articles more often? And this brings us to another question: why don't we always omit the articles, also in normal speech? If we can use the channel more effectively by omitting articles, then why don't we just omit them? I will answer these questions in the concluding section of this chapter.

4.5 Conclusion: Why we do not always omit the articles.

I argued in the preceding sections that the amount of information that is sent through the channel *per unit of time* is always lower when an article is used. This means that the channel is always used in a less effective way if an article is used. Then, instead of asking ourselves the question: '*why do we omit articles?*' shouldn't we better ask ourselves the question: '*why do we still prefer to use the article more often than to omit it?*'.

Let me start with the question I asked in the preceding section. Even in Italian headlines the use of an article may lead to a 'slower' rate of information processing. Still, Italian headline writers and readers prefer the use of the article more often than their Dutch colleagues. Why?

Headline writers will only omit the article if it leads to a *perceptible* drop in processing cost. Now, of course, it is not possible to define the boundaries of what exactly is a perceptible drop in processing cost and what is not. It depends on the situation, on the condition of the speaker, etc. First of all, this, of course, explains the optionality we find in the judgements. We can compare it with lifting weights in a fitness school. We will definitely feel the difference between dumb-bells of 5 -10 -15 kg. And probably also between dumb-bells of 7.5 -10 – 12,5 kg. Probably not between dumb-bells of 9,9 – 10 – 10,1 kg, though this may be a personal matter. A very weak or a very tired person may in fact feel the difference, while a very strong person may not. Hence, the judgements will be optional. However, and this is crucial: the bigger the difference between the weights of the dumb-bells, the larger the probability that people will notice it and be sensitive to it. And that's how it is for the processing cost of articles too. The bigger the difference between the processing cost per unit of information when the article is used and the processing cost per unit of information when the article is omitted, the larger the probability that a headline writer/reader will perceive it, and, if

he strives for minimalization of the processing cost per unit of information, will prefer to omit it.

Let us now face the question why we do not always omit articles. Why do headline writers choose the version *with* the article present if there is no strong difference in processing cost between the version with and without the article, and why do normal speakers in normal contexts use the version with the article present?

A plausible answer could be that if there is no reason to omit, we use the article, because that will always be the automatic outcome of the syntactic process if there is enough channel capacity available. Only if the channel capacity is limited will the article be omitted, because it cannot be selected then. This answer, however, needs further explanation. After all, as I argued, if we do not use the article we use the syntactic channel in a more effective way. We process information at a higher processing rate, with a lower processing time per unit of information. This means that in the same amount of time we can process more information, and maximize the effectiveness of channel use. But then, why don't we always try to maximize the effectiveness of channel use? Why do we accept and even prefer a less effective channel use than would be maximally possible?

The reason for this is that a channel that is used at its maximal capacity is more vulnerable to influences from outside that affect the processes taking place within the channel, just like a car that is driving too fast. If for some reason the speaker/listener is distracted, this will affect the communication process. This already happens if the speaker is speaking in a normal way, and uses the channel at a less than maximally possible processing rate. If the channel is used at (or more close to) its maximal capacity, the risk of 'damage' to the communication process will even be higher.⁸³ We do not want to needlessly hamper the communication process, and therefore, in normal situations, we tune in to a less effective channel use with a longer processing time per unit of information. That is why in normal speech functional elements are produced, even if this leads to a decrease in processing speed per unit of information. And the same goes for headline writers if they do not feel a

⁸³ An interesting study in this respect is the study of Stuart et al. (2002). This study investigated the effect of short and long auditory feedback delays at two speech rates with normal speakers. Seventeen participants spoke under delayed auditory feedback (DAF) at 0, 25, 50, and 200 ms at normal and fast rates of speech. There were significantly more dysfluencies observed at the fast rate of speech ($p = 0.028$).

perceptible difference in processing cost of a noun with or without an article. If the use of the article does not have perceptible effects on the processing time, headline writers will use it.

This means that in my view use of functional elements, like articles, leads to a higher processing time per unit of information, but it makes the communication process more resistant to influences from outside the actual communication process that may interfere with the transmission of information. And that is why, when we are in a situation where we have enough time, we tune in to a slower than maximally possible transmission rate of information through the syntactic channel. This makes the use of the syntactic channel somewhat less effective, but more 'shock-proof'. Children cannot yet fully benefit from the use of the syntactic channel, as they do not have the processing resources that are necessary to retrieve the (low informative) functional elements from the closed class sets of functional elements (with high relative entropy values). Headline writers strive for a minimal processing time per unit of information, and this leads to omission of functional elements.

Chapter 5 Application of the model to another language: German

5.1 Introduction

This chapter focuses on article omission in German. Investigating this phenomenon in German offers us the possibility to examine whether the pattern observed in Dutch and Italian is related to differences in properties of Germanic and Romance languages more generally. It has been argued, for example, that Germanic and Romance languages differ with respect to the structure of DP's (Longobardi, 2001) or with respect to parameter setting (Chierchia, 1998).^{84,85} If differences in omission patterns of articles between Dutch and Italian are related to these properties we would expect German to be similar to Dutch with respect to article omission. We would thus predict: *omissions in German = omissions in Dutch > omissions in Italian*.

Further, there are differences between the morphological paradigms in the three languages, as is shown in Tables 1 to 3. For ease of reference I will repeat the tables of the article paradigms in Dutch and Italian presented in the first chapter.

⁸⁴ Longobardi (2001) proposed a hierarchy of languages, depending on the possibility of having bare nouns functioning as arguments. In this hierarchy French is the most restricted language, as it does not permit bare nouns in argument position. The other Romance languages are less permissive than French, but more restricted than Germanic languages, as they do allow the use of bare nouns in argumental position only in a restricted range of syntactic contexts. The Germanic languages are the most permissive, as in Germanic languages bare (plural) nouns can occur with either an existential or a generic interpretation, and mass nouns can be used without an article in all positions.

⁸⁵ According to Chierchia's (1998) model of nominal parameter mapping, nouns are universally specified for [+arg] or [+pred]. If a language is specified for [+arg], as is the case in for example Chinese, no article is necessary for a noun to function as an argument. If a language is specified for [+pred], an article is required for a noun to function as an argument. Romance languages are specified for [+pred], and therefore in general bare nouns are not permitted in Romance languages. The specification of Germanic languages is [+arg, +pred], which explains why in Germanic languages we find both bare nouns as well as nouns with an article functioning as arguments.

Table 1 *Morphosyntactic forms of Dutch articles*

	Definite		Indefinite	
	Common gender	Neuter	Common Gender	Neuter
Singular	de	het	een	een
Plural	de	de	-	-

The Dutch article paradigm has two ‘gender’ types, neuter and common gender. As the table shows, there is an overlap in the grammatical contexts in which a specific article form can be used. The article ‘*de*’ is used before singular nouns of common gender, and before plural nouns of all gender types. The Dutch indefinite article form ‘*een*’ is used with neuter as well as common gender singular nouns. Thus, we can conclude that the Dutch article paradigm is ambiguous in the mapping of the grammatical feature combinations ‘*gender/number*’ on article forms.

Table 2 *Morphosyntactic forms of Italian articles*

	Definite		Indefinite		Partitive	
	Mas-culine	Femi-nine	Mas-culine	Femini-ne	Mas-culine	Femi-nine
Singu-lar	il lo l’	la l’	un uno	una un’	del dello dell’	della dell’
Plural	i gli	le	dei degli	delle	dei degli	delle

Italian has a so-called ‘full’ paradigm of articles because it consists of definite, indefinite *and* partitive articles. The Italian article paradigm distinguishes between two gender types (masculine and feminine), and these types have different article forms for singular and plural. The morphological form of the article depends on the phonological context. Differently from Dutch, in Italian each grammatical feature combination (‘*gender/number*’) is reflected by a specific form.

Table 3 *Morphosyntactic forms of German articles*

	Definite			Indefinite		
	Mas- culine	Femi- nine	Neuter	Mas- culine	Femi- nine	Neuter
Singular						
Nominative	der	die	das	ein	eine	ein
Genitive	des	der	des	eines	einer	eines
Dative	dem	der	dem	einem	einer	einem
Accusative	den	die	das	einen	eine	ein
Plural						
Nominative	die	die	die			
Genitive	der	der	der			
Dative	der	der	der			
Accusative	die	die	die			

Differently from Dutch and Italian the German article paradigm distinguishes between three gender types and makes case distinctions. If we compare the three article systems, we see that the number of grammatical distinctions encoded in the articles is higher in German than in the other two languages. In addition, we find a higher amount of ambiguity in the mapping of grammatical feature combinations to specific article forms in German than in Dutch and Italian. For example the form ‘*der*’ encodes seven different grammatical specifications.

- (1)
- | | |
|----------------------------------|-------------------|
| a. masculine singular nominative | <i>der Mann</i> |
| b. feminine singular genitive | <i>der Frau</i> |
| c. feminine singular dative | <i>der Frau</i> |
| d. masculine plural genitive | <i>der Männer</i> |
| e. feminine plural genitive | <i>der Frauen</i> |
| f. neuter plural genitive | <i>der Kinder</i> |

It could therefore be argued that from the viewpoint of morphology German is a more ‘complex’ language than Dutch and Italian.⁸⁶ If differences in omission pattern of articles are related to this alleged level

⁸⁶ See also Kupisch, 2007, for a similar observation in a crosslinguistic comparison of German, Italian and French

of morphological complexity, we would expect to find more omissions in German than in Dutch and Italian and we would thus predict:
omissions in German > omissions in Dutch and Italian.

What are the predictions that can be made on the basis of the model I proposed in chapter 4? I showed that the measure of relative entropy makes the right predictions for omission of articles in Dutch and Italian. I have calculated the relative entropy of the article set in German, using the corpus data of the Tübingen Treebank of Spoken German (TüBa-D/S, formerly Verbmobil), Table 4 shows the results of this relative entropy calculation.

Table 4 Calculation Relative Entropy German Article Set

RELATIVE ENTROPY GERMAN ARTICLES (freq.data TüBa-D/S, formerly Verbmobil)				
	freq.	p = rel freq	log p	p*log p
der	7380	0,185721116	-2,428790237	-0,451077634
die	6103	0,15358482	-2,702892463	-0,415123253
das	3638	0,091551954	-3,44926551	-0,315786998
des	334	0,008405265	-6,894491046	-0,057950022
dem	3975	0,100032715	-3,321456193	-0,332254281
den	6406	0,161209955	-2,632987255	-0,424463758
ein	4423	0,111306842	-3,167385811	-0,352551714
eine	3111	0,078289755	-3,675032658	-0,287717407
einer	360	0,009059567	-6,786342242	-0,06148132
eines	57	0,001434431	-9,445305324	-0,013548642
einem	632	0,015904573	-5,97441459	-0,09502051
einen	3318	0,083499006	-3,582097167	-0,299101553
	39737	1		-3,106077091
			Absolute Entropy H =	3,106077091
H MAX: if all determiners have same probability:				
der		0,083333333	-3,584962501	-0,298746875
die		0,083333333	-3,584962501	-0,298746875
das		0,083333333	-3,584962501	-0,298746875
des		0,083333333	-3,584962501	-0,298746875
dem		0,083333333	-3,584962501	-0,298746875
den		0,083333333	-3,584962501	-0,298746875
ein		0,083333333	-3,584962501	-0,298746875
eine		0,083333333	-3,584962501	-0,298746875
einer		0,083333333	-3,584962501	-0,298746875
eines		0,083333333	-3,584962501	-0,298746875
einem		0,083333333	-3,584962501	-0,298746875
einen		0,083333333	-3,584962501	-0,298746875
		1		-3,584962501
			Max Entropy H max =	3,584962501
RELATIVE ENTROPY German articles = H/Hmax				
H r	=	0,866418293		

The table shows that the entropy value of the German article system is 0,87. This value is lower than the value found for Dutch (0,94) and higher than the value found for Italian (0,75).

In chapter 5 I argued that relative entropy reflects the complexity of the article selection process. We saw in section 4.4.1.2 for Dutch child speech data that there is a strong negative correlation between the relative entropy value of the article set the child actually produces and the article omission rate of children. The better the child is able to process higher entropy levels (this ability is reflected in the adjusted relative entropy value of the article set the child produced), the less frequently the child will omit articles. This shows that omission of articles is related to the child's inability to process high relative entropy values. Moreover, it shows that relative entropy is a very strong measure of the processing resources that are required to select an article from the set and, consequently, a strong measure of the complexity level of the article set. The higher the relative entropy level of the set, the more processing resources are required for the selection of an element from the set, and the more omissions we find in child speech.

Adding the data of a third language can provide us with more evidence for the strength of relative entropy as a measure of the complexity of a set. The fact that German has a relative entropy value that is lower than the Dutch and higher than the Italian entropy value enables us to examine how precisely the relative entropy value reflects the differences in complexity level. Differences in complexity level will lead to differences in processing cost required for the selection of an element from the set. These differences will cause differences in omission of the elements by adults in 'special' time-constrained contexts and in article productions of children. Particularly, if relative entropy is a reliable measure of the complexity level of the article set, then, since relative entropy is higher in German than in Italian, we should find more omissions in German than in Italian in adult's special registers and child speech. And, in the same vein, since relative entropy is higher in Dutch than in German we should find more omissions in Dutch than in German adults' special registers and child speech.

Summarizing: If omissions of articles are related to the level of relative entropy (H_r), we should expect to find more omissions of articles in Dutch than in German, and more omissions of articles in German than in Italian.

$H_r \text{Dutch} > H_r \text{German} > H_r \text{Italian} \rightarrow$
omissions Dutch > omissions German > omissions Italian.

To test my predictions I created and analyzed a database of German headlines, I conducted an experiment with German newspaper readers, and analyzed German child speech, using files from the Childes Database. In this chapter I will present a summary of the results.

5.2 Headlines database

The database of German headlines which I created was analyzed in the same way as the Dutch and Italian databases (see chapter 3.3).⁸⁷ Let us start with a look at the overall rate of omission in so-called ‘obligatory’ contexts in phrasal headlines, i.e., the contexts that in the adult grammar would have required the use of an article. Examples are given in 2.

- (2) FEUERBALL ERSCHRECKT DUTZENDE SPANIER
Fireball frightens thousands of Spaniards
RUPPRATH SUCHT NEUEN TRAINER
Rupprath is looking for new coach

Table 5 shows the results for German and compares these results with those found for Dutch and Italian and discussed in chapter 3 (see Table 1 in chapter 3).

⁸⁷ For the German database I used headlines from *the Franfurter Rundschau*, *Frankfurter Allgemeine*, *Westdeutsche Zeitung* and *Westdeutsche Allgemeine*. Like the Dutch and Italian databases, the German database consists of 1000 headlines.

Table 5 *Omission rates of articles in obligatory phrasal contexts in German, Dutch and Italian.*

Language	Non-standard omission in phrasal contexts
<i>Dutch</i>	89,9
<i>German</i>	58,8
<i>Italian</i>	5,3

Statistical analysis shows that the rate in German is truly different from Dutch and Italian, (Pearson chi-square German – Dutch: $\chi^2 = 300.781$, $p < .0001$; German – Italian: $\chi^2 = 472.595$, $p < .0001$). Hence in this first analysis based on overall omission rates, the observed results in the three languages confirm the predictions made on the basis of the Relative Entropy level:

$$\text{omissions Dutch} > \text{omissions German} > \text{omissions Italian.}$$

Let us now examine the omission of articles in different types of linguistic context and in different sentence positions. I will compare

- (a) omissions in sentences containing a finite verb versus omissions in sentences containing no finite verb
- (b) omissions before nouns directly following a preposition versus omissions before nouns not following a preposition
- (c) omissions in sentence-initial position versus omissions in sentence-internal position

Effect of finiteness and preposition

Let me start with the effect of finiteness and preposition. I will first compare the omission rate found in phrasal headlines containing a finite verb with the omission rate found in phrasal headlines with no finite verb. Since in Dutch and Italian a difference was found in the omission rates observed between nouns directly following a preposition and nouns not following a preposition (see chapter 3, section 3.3.2), I will make the same distinction in preposition versus non-preposition contexts for German. Table 6 shows the results of this analysis, for all three languages.

Table 6 Omission percentages in article-requiring noun contexts in headlines with finite verb and headlines with no-finite verbs; nouns following a preposition (columns 1 and 2) and nouns not following a preposition (columns 3 and 4) analyzed separately.

	Nouns not following preposition		Nouns directly after preposition	
	Finite verb (1)	No finite verb (2)	Finite verb (3)	No finite verb (4)
Dutch	92,9	98,5	81,0	79,8
German	70,5	87,5	21,5	28,0
Italian	2,1	33,3	0,0	0,0

Table 6 shows a number of interesting effects. First of all in non-preposition contexts in German we find significantly less omissions if the headline contains a finite verb.⁸⁸ Further, if we compare the results in the first two columns for German with those found for Dutch and Italian, we find that German is different from Dutch: both in finite and non-finite phrasal headlines we find less omissions in German.⁸⁹ German is also different from Italian: both in finite and non-finite phrasal headlines we find less omissions in Italian.⁹⁰ Hence, both in phrasal headlines with a finite verb and in phrasal headlines with no finite verb we find more omissions in German than in Italian, and less omissions in German than in Dutch.

A further observation is that in German, as was the case with Dutch and Italian, the omission rate is influenced by the presence of a preposition before the noun. In German too, we find significantly less omissions with nouns directly after a preposition.⁹¹ In the preposition contexts too the differences between the three languages are significant.⁹²

⁸⁸ Pearson chi-square $\chi^2 = 4.221$, $p = .0399$.

⁸⁹ Pearson chi-square finite verbs: $\chi^2 = 117.282$, $p < .0001$, non-finite verbs: Fisher's exact $p = .0035$.

⁹⁰ Pearson chi-square finite verbs: $\chi^2 = 629.892$, $p < .0001$, non-finite verbs: $\chi^2 = 34.373$ $p < .0001$.

⁹¹ Difference in finite verb headlines (column 1 versus column 3) $\chi^2 = 166.878$, $p < .0001$, no-finite verb headlines (column 2 versus column 4): $\chi^2 = 27.967$ $p < .0001$.

⁹² In finite verbal headlines, with nouns following a preposition, German is different from Dutch: $\chi^2 = 173.388$ $p < .0001$, and from Italian: $\chi^2 = 69.415$ $p < .0001$. In the non-finite verbal headlines, with nouns following a preposition German is different from Dutch: $\chi^2 = 26.810$ $p < .0001$, and from Italian: $\chi^2 = 35.791$ $p < .0001$.

Note that, although presence of a preposition leads to significantly less omissions in all three languages, the concrete consequences in the three languages differ. In Italian we hardly find any omissions of articles if a preposition is present. As I argued in chapter 3 this finding is in accordance with the predictions of structural theories that relate omissions of articles to the absence of a case assigner. However, we also saw in chapter 3 that in Dutch presence of a preposition only leads to a less strong preference for omission, but omission is still preferred. German is different from Dutch and Italian. Different from Italian in that omission of the article after the preposition is still an option. In fact it is the preferred option in 21, 5 % and 28% of the cases; different from Dutch in that in German presence of a preposition leads to a switch in preference. In Dutch the preference for omission becomes less strong when a preposition is present, but omission of the article is still preferred over use of the article. In German the preference switches: if no preposition is present the version without article is preferred, if a preposition is present the version with article is preferred (though, differently from Italian, omission is still possible). This observed optionality in preposition contexts in both German and Dutch shows that the observed effect cannot be accounted for by a theory that relates omission of articles to the presence of a case assigner. We need a different account.

Effect of position

Table 7 *Omission percentages in article-requiring-noun contexts in initial and internal position (non-preposition contexts), in headlines with finite verb and headlines with no finite verbs.*

	Finite verbs		No finite verbs	
	Initial	Internal (non- prep.)	Initial	Internal (non- prep.)
Dutch	95,2	89,2	98,9	92,9
German	79,1	61,2	90,0	78,6
Italian	4,0	0,6	48,8	13,4

The results in Table 7 show that in German in sentences with finite verbs we find significantly more omissions from sentence-initial

position.⁹³ In Dutch and Italian the differences between omissions from initial and internal position in headlines with finite verbs were significant too (see chapter 3).

In German, as for Dutch, in headlines with no finite verbs the very low absolute number of internal contexts makes it difficult to draw a strong conclusion. The results show that there is a tendency to omit more from sentence-initial than from sentence-internal position, also in non-finite verbal headlines. However, possibly due to the low number of contexts (14 in German, 14 in Dutch) the difference does not reach significance ($\chi^2 = 0.470$, $p = .4929$). In Italian the difference was significant, as we saw in chapter 3. Hence, in all three languages we find an effect of position. In verbal headlines we find a higher omission rate from sentence-initial than sentence-internal position. Crucially, in no language is there an opposite effect, that is, in no language do we find more omissions from sentence-internal than sentence-initial position.

If we compare the omissions found in German with the omissions found in Dutch and Italian we observe, again, that in German we find less omissions than in Dutch, and more omissions than in Italian.⁹⁴

Let me summarize the results of the comparison of the omission rates of articles in German headlines with the omission rates found in Dutch and Italian headlines:

- In all contexts collapsed together the overall omission rate in German is higher than the omission rate found in Italian headlines, and lower than the omission rate found in Dutch headlines:
omissions Dutch > omissions German > omissions Italian
- We find that both in finite and non-finite phrasal headlines the omission rate in German is higher than the omission rate in Italian, and lower than the omission rate in Dutch:
omissions Dutch > omissions German > omissions Italian

⁹³ Pearson chi-square $\chi^2 = 24.038$ df 1, $p < .0001$.

⁹⁴ In finite headlines with respect to omissions from initial position German is different from Dutch: $\chi^2 = 47.363$, $p < .0001$; and from Italian: $\chi^2 = 340.143$, $p < .0001$; with respect to omissions from internal position in finite headlines German is different from Dutch $\chi^2 = 56.686$, $p < .0001$; and from Italian: $\chi^2 = 285.122$, $p < .0001$. In non-finite headlines with respect to omissions from initial position German is different from Dutch $\chi^2 = 8.025$, $p < .005$; and from Italian: $\chi^2 = 21.492$, $p < .0001$; with respect to omissions from internal position in non-finite headlines German is different from Italian: $\chi^2 = 23.037$, $p < .0001$, but, most likely due to the small sample size, not different from Dutch $\chi^2 = 0.292$, $p = .5892$

- Distinguishing between sentence-initial and sentence-internal position, the omission rate in German is higher than the omission rate in Italian, and lower than the omission rate in Dutch:
omissions Dutch > omissions German > omissions Italian
- In all three languages we find less omissions if the headline contains a finite verb.
- In all three languages find less omissions from sentence-internal position than from sentence-initial position
- In all three languages we find significantly less omissions if the article-requiring noun is preceded by a preposition. However, presence of a preposition has different consequences in the three languages: *omissions Dutch > omissions German > omissions Italian*

5.3 Headlines Experiment

I also conducted an experiment with German headline readers. The experimental conditions were similar to the conditions tested in Dutch and Italian. I tested 36 participants, all students of the University of Oldenburg.⁹⁵ In this section I will present the results of the German experiment, and then compare them with the results found in Dutch and Italian.

5.3.1 Finiteness effect

In this section I will discuss the results of the conditions that tested whether the preference for omission was influenced by the absence or presence of a finite auxiliary verb. There were two versions of each headline, tested in separate conditions:

- one version in which the finite auxiliary verb was omitted, *with* and *without* an article, see examples given below

(3) DER MORD AN ANGELA B. AUFGEKLÄRT

⁹⁵ I acknowledge Esther Ruijgendijk for her help in conducting the German headlines experiment.

MORD AN ANGELA B. AUFGEKLÄRT
(The) murder of Angela B. solved

- one version in which the finite auxiliary verb was present, *with* and *without* an article, see examples given below

- (4) DER MORD AN ANGELA B. IST AUFGEKLÄRT
MORD AN ANGELA B. IST AUFGEKLÄRT
(The) murder of Angela B. has been solved

Table 8 Preferences for omission in headlines with either the finite auxiliary present or omitted.

Condition	Dutch	German	Italian
No-AUX	87,6	84,9	82,9
YES- AUX	62,6	66,9	12,9

The difference between the preferences for omission in German headlines with finite auxiliary and headlines without finite auxiliary is significant.⁹⁶ In chapter 3 we saw that in Dutch and Italian too the difference between the preferences for omission in headlines with finite auxiliary and headlines without finite auxiliary was significant.

Let us now compare the omission rates of the three languages. In the condition with finite auxiliary verb (the second row in Table 8) German is different from Italian, but not from Dutch.⁹⁷ We saw in chapter 4 that Dutch is different from Italian, hence: *omissions Dutch* = *omissions German* > *omissions Italian*.

In the condition without finite auxiliary verb (the first row in Table 8) the differences between the three languages were not significant.⁹⁸
omissions Dutch = *omissions German* = *omissions Italian*.

⁹⁶ In the condition with definite articles (Wilks'Lambda F (35,1) = 25.415, p < .0001), in the condition with indefinite articles (Wilks'Lambda F (35,1) = 9.445, p = .004)

⁹⁷ Oneway Anova showed that German is different from Italian (F (1,68) = 73.126, p < .001), but not from Dutch (F (1, 84) = 0.806 p = .546).

⁹⁸ Oneway between group Anova: F (2, 117) = 2.255 p = .109), Oneway Anova showed that German is not different from Dutch (F (1, 84) = .002 p = .962), nor from Italian (F (1,68) = 3.825, p = .055), Dutch is not different from Italian (F (1,82) = 3.442, p = .067).

5.3.2 Effect of position(1): omissions from sentence-initial and sentence-internal position and definiteness effect

Let us now look at the results of the headlines with finite verbs in which a determiner was present before the sentence-initial subject or sentence-internal object and the second noun (if there was one) was a noun that did not require an article. The following examples illustrate the test-headlines in German:

IniSubDef (yes/no definite determiner in subject):

- (5) (DER) ÖLPREIS STEIGT KRÄFTIG
(The) price of oil rises sharply

IntObjDef (yes/no definite determiner in object):

- (6) WASHINGTON WILL (DIE) BOSNISCHE REGIERUNG
STÄRKEN
Washington wants to reinforce (the) Bosnian government

IniSubIndef (yes/no indefinite determiner in subject):

- (7) (EIN) BULLTERRIER VERLETZT BRITNEY SPEARS
(A) pitbull terrier attacks Britney Spears

IntObjIndef (yes/no indefinite determiner in object):

- (8) RUPPRATH SUCHT (EINEN) NEUEN TRAINER
Rupprath is looking for (a) new coach.

Table 9 Preferences for omission of the articles in the conditions *IniSubDef*, *IntObjDef*, *IniSubIndef*, *IntObjIndef* in Dutch, Italian and German

	Dutch	German	Italian
I. <i>IniSubDef</i> : definite article in sentence-initial subject	88,8	58,3	21,2
II. <i>IntObjDef</i> : definite article in sentence-internal object	79,6	56,7	18,8
III. <i>IniSubIndef</i> : indefinite article in sentence-initial subject	91,7	85,9	80,6
IV. <i>IntObjIndef</i> : Indefinite article in sentence-internal object	84,4	77,8	45,9

Let us start with an important observation on the crosslinguistic comparison of the omission rates. In all four conditions we find that the omission rate in German is ‘in-between’ the Dutch and Italian omission rates: *omissions Dutch* > *omissions German* > *omissions Italian*.

Crucially: in no condition is there an opposite effect, that is, in no condition do we find more omissions in Italian than in German, or more omissions in German than in Dutch.⁹⁹

Further, if we look at the effect of definiteness, in German the differences between the preferences for omission before subject and object nouns in the condition with *definite* articles differ significantly from the preferences for omission before subject and object nouns in the conditions with *indefinite* articles.¹⁰⁰ In this respect German patterns like Italian, where a similar difference was found, and is different from

⁹⁹ *IniSubDef* (definite article on subject): One way Anova German -Dutch (F (1,84) = 27.369, p <.0001), German-Italian (F (1,68) = 29.183, p <.0001).

IntObjDef (definite article on object) One way Anova German - Dutch (F (1,84) = 14,097, p <.0001), German-Italian (F (1,68) = 31.561, p <.0001)

IniSubIndef (indefinite article on subject) One way Anova German -Dutch (F (1,84) = 2.915, p = .091), German- Italian (F (1,68) = .434, p = .512). Since the difference between Dutch and Italian is significant (see chapter 4 for details) this means that German is ‘in-between’ Dutch and Italian with respect to omission of indefinite articles from subject position.

IntObjIndef (indefinite article on object) One way Anova German -Italian (F (1,68) = 20.454, p <.0001), German- Dutch (F (1,84) = 1.459, p = .231).

¹⁰⁰ Subject condition: (F (35,1) = 21.160, p <.001, object condition: (F (35,1) = 15.618, p <.0001)

Dutch, where a tendency for a higher preference of omission of indefinite articles was found, but the difference did not reach significance.

5.3.3 Effect of position(2): order of omission

In this section I will discuss the results of the conditions in which the headline contained two article-requiring nouns. In one version of the headline the article was used before the subject and omitted before the object (see example 8, the A-version), in the other version of the headline the article was omitted before the subject and used before the object (see example 9, the B-version).

- (9) A: EIN POLIZIST ERSCHIEßT BANKRÄUBER
 (10) B POLIZIST ERSCHIEßT EINEN BANKRÄUBER
 Police officer shoots (a) bankrobber

Tested were definite and indefinite articles.

Table 10 *Preferences for article used in sentence-initial subject or sentence-internal object noun in German.*

	Definite articles		Indefinite articles	
	Article preferred in initial subject	Article preferred in internal object	Article preferred in initial subject	Article preferred in internal object
German	28,9	61,7	16,1	69,4

In chapter 3 we saw Stowell's (1999) proposal that article omission from the object would be impossible if the article had not been omitted from the subject, hence headlines of the type 9 A should not be possible. In this condition this proposal was tested. The difference between the preferences for the version with the article in initial subject and the version with the article in internal object were significant.¹⁰¹ At a first glance this means that German patterns like Dutch and seems to confirm

¹⁰¹ In the condition with definite articles (repeated measures ANOVA, Willks'Lambda $F(35,1) = 12.248$, $p < .001$) and in the condition with indefinite articles (repeated measures ANOVA, Willks'Lambda $F(35,1) = 56.000$, $p < .001$).

the prediction made by Stowell. However, it is important to note that Stowell's account makes a very strong claim. It predicts that omission from the object should be *impossible* (because of structural constraints) if the article has not been omitted from the subject. This is not what I found. Although in German (as in Dutch, as we saw in chapter 3) there is a stronger preference for the version in which the article is omitted from the subject, the results do also show a form of optionality. In the condition with definite articles the version in which the article is omitted from the subject is preferred in 'only' 61,7 % of the test-items in German and 72,4 % of the test-items in Dutch. In German we find a preference of 28,9 % for the version in which the article is omitted from the object. This is too much for an absolute rule-based account. In the conditions with indefinite articles the preferences for the version in which the article is omitted from the subject are 'only' 69,4% for German and 65,2% for Dutch. In my view it is hard to explain these responses by a structural account. Larger differences between the preferences for the two conditions would have been expected. In chapter 3 (section 3.4.3.3.) we saw that the results in Italian, where the difference in the preferences for omission from the subject or the object did not even reach significance, are even less expected on such a structural account.

5.4 German Child Speech

In chapter 3 we compared article omission by Dutch and Italian adults in special registers with the omission of articles by language-acquiring children in these languages. In order to complete the compilation of comparative data we need to look at article omission in German childspeech as well. The obvious question is: Do German children, like German adults in a special register like Headlines, omit less articles than Dutch children, and do they omit more articles than their Italian counterparts?

To my knowledge, this study is the first one that compares omission of articles in German and Dutch child speech. There have been previous studies on the comparison of article omission in child speech in German and Italian, especially in language acquisition of bilingual children (Kupisch, 2006).

Table 11 *Chilides Files German used in this study*

	Files	Age range	MLU range	Utterances with article-requiring nouns
Lisa	105,109,110,202,204,206,210,301	1;5-3;0	1;2-3.2	423
Kerstin	110, 201,203,206,207,209,210,302	1;10-3;2	1.8-3.5	401
Falco	109,110,200,201,202,204,206,209	1;9-2;9	1.1-3.5	375
Simone	110,202,204,205,207,208,211,302	1;10-3;2	1.5-3.4	396

The age of the children during the period investigated is roughly the same as in the other two languages (Mann Whitney, $z = -.901$, $p = .367$; age in months: German $M = 28.2$, age range 17 to 38, Italian $M = 27.7$, age range: 19 to 36, Dutch: $M = 28.8$, age range: 19 to 36). Differences could therefore not be attributed to the fact that some children are older than others.

Table 12 *Overall omission mean and standard deviation in whole period in all obligatory context types collapsed together.*

	Overall Omission	
	M(=Mean)	SD(=Stand.Dev)
German	41,2	34.02
Dutch	62.71	29.63
Italian	38.17	31.85

In our analysis of German child speech we made the same distinction in context-types as was done for Dutch and Italian (see chapter 3, section 3.5.2) Table 12 shows the overall omission rates of articles in obligatory contexts: the contexts that in normal adult grammar would have required the use of an article.

The difference between omissions in child speech in German and Dutch is significant, we find more omissions in Dutch.¹⁰² The difference

¹⁰² German-Dutch: Mann Whitney, $z = -2.558$, $p = .011$

between German and Italian is not significant.¹⁰³ As discussed in chapter 3, the difference between Italian and Dutch is significant.¹⁰⁴

Summarizing the findings:

omissions Dutch > omissions German = omissions Italian.

Let us now look at the differences in omission pattern in the two different stages of linguistic development. Similarly to what we did for the data of Dutch and Italian child speech, we used the rate of verbal utterances (VU) as a measure of linguistic development (see chapter 3, section 3.5.3.1). Table 13 illustrates the results.

Table 13 Omission in different periods of linguistic development as measured by rate of VU, in all obligatory context types collapsed together.

	Stage 1 0.03-0.29 VU		Stage 2 0.3-0.6 VU		Mann-Whitney st.1 versus st. 2
	M	SD	M	SD	
Dutch	87.05	14.8	41.6	21.8	$z = -4.372, p < .0001$
German	72.1	32.7	19.16	16.8	$z = -3.544, p < .0001$
Italian	58.29	29.5	16.9	17.1	$z = -4.097, p < .0001$

In the first stage German is significantly different from Dutch and Italian.^{105, 106}

Stage 1: *omissions Dutch > omissions German > omissions Italian.*

In the second stage German is, again, different from Dutch, but German is not different from Italian.^{107, 108}

Stage 2: *omissions Dutch > omissions German = omissions Italian.*

Hence, in stage 1 we find a difference between German and Italian, and between German and Dutch. In stage 2, however, the article omission rate of the German children is not different from the omission rate of

¹⁰³ German-Italian: Mann Whitney, $z = -.195, p = .846$

¹⁰⁴ Dutch-Italian: Mann Whitney $z = -3.012, p = .003$

¹⁰⁵ German-Dutch: Mann Whitney, $z = -2.033, p = .045$

¹⁰⁶ German-Italian: Mann Whitney, $z = -2.001, p = .045$

¹⁰⁷ German-Dutch: Mann Whitney, $z = -3.024, p = .002$

¹⁰⁸ German-Italian: Mann Whitney, $z = -.198, p = .845$

their Italian counterparts. Both groups, however, the Italian and the German children, omit less articles than the Dutch children. This implies that the article production of the German children shows a rapid development in the second stage. As a result of this fast development of the German children's article production in stage 2 the omission rate in German is not different any longer from the Italian omission rate in stage 2. Interestingly, a similar observation has been reported by Lléo & Demuth (1999) who compared the article production of German and Spanish children. They found that Spanish children tend to show a gradual increase in the use of articles over the developmental stages. For German children, however, the article production was particularly problematic in the early developmental stages (between age 1;5 and 1;10), but in later stages (from age 1;10 to 2;3) Lléo & Demuth observed a fast development of article production in German child speech towards the level of article use of Spanish children.¹⁰⁹

5.5 An information-theoretical account of crosslinguistic findings

In the first section of this chapter I argued that if omission of articles is related to the value of relative entropy, we can expect to find more omissions of articles in Dutch than in German, and more omissions in German than in Italian. In this section I will discuss whether this prediction is borne out.

Let us first take a look at the *child speech data*. In the first stage we found that the omission rate in German was in-between the Dutch and Italian omission rates, and that in the second stage the omission rate in German was like the omission rate in Italian, and different from Dutch. If we assume that children omit articles because they do not have the processing resources necessary to process the high relative entropy value

¹⁰⁹ Although Spanish is in many respects similar to Italian, we cannot straightforwardly claim that the development of article use in Spanish child speech is similar to the development of article use in Italian speech. Important, however, in this respect is a study on the comparison of article omission by Catalan and Italian speaking children (Guasti et al. 2007) which revealed no differences in omission pattern between these two groups of children. Admittedly, Catalan is also a different language from Spanish, but the findings from the study of Guasti et al. strengthen the assumption that the development of article use in Spanish child speech will not deviate much from the development of article use in Italian child speech.

of these elements, the order found in stage 1 is to be expected. We find most omissions in the language with the highest relative entropy. The lower the entropy value, the lower the omission rate.

In stage 2 we found for Italian that children produced articles almost adult-like, and, as I argued in chapter 4, this indicates that children's brains can process *at least* a relative entropy value of 0,75, the relative entropy value of Italian articles. But this does not exclude the possibility that children are capable of processing an even higher relative entropy value. The fact that they still omit articles in Dutch at this stage means that the relative entropy value for Dutch articles is still too high. More specifically, 0.94 must be too high. The value of German is 0,87, and apparently children at this more advanced stage of development are able to process not only 0,75, but even 0.87. That is why we find an almost adult-like article production not only in Italian, but also in German child speech in stage 2.

Headlines:

If differences in omissions are related to the amount of processing effort, and thus to the differences in relative entropy between the three languages, the prediction is that in German we will find an omission rate (or preference for omission) that lies in between the percentages found for Dutch and Italian.

- In the *database of headlines*, when we look at all contexts collapsed together, as well as in the finite phrasal and non-finite contexts examined separately, we find that the omission rates in German confirm the prediction; they are between the observed omission rates in Dutch and Italian.

The observed effects of finiteness and preposition fit in the account I propose in chapter 4, section 4.4.2. The goal of a headline (from the perspective of language processing) is to maximize the effectiveness of the use of the information channel, in other words: to transfer as much information as possible in a limited amount of time. Now, both, finite auxiliary verbs and prepositions are closed-class elements. Hence, elements with a low informative value, which will slow down the processing speed of the headline (see chapter 4, section 4.4.2 for a detailed argumentation) Thus, use of a finite auxiliary verb and use of a preposition will lead to a less effective use of the channel. If the channel is already used less effectively, use of other closed class elements, like for example articles, leads to a less perceptible decrease in the effectiveness

of the channel use than would be the case with use of closed class elements at the point where the channel is used optimally. Hence, we could say that a less effective use of the channel makes the ‘syntactic route’ of encoding information, using closed-class, functional elements become available more easily. Whether or not the articles will actually be selected depends on the processing resources necessary to select them from the lexicon. Because of differences in relative entropy of the article set in different languages, this amount of processing resources will differ between languages. Most articles will be selected in the language with the lowest relative entropy, and we will find the highest omission in the language with the highest relative entropy. This is the effect we observe in the omissions in headlines. Thus, in all languages use of a preposition or use of finiteness facilitates accessing the article set, but the effect is different, depending on the relative entropy value of the article set.

- In the *experiment*, specifically in the conditions: *IniSubDef*, *IntObjDef* and *IniSubIndef* we saw: *Preferences for omission Dutch > Preferences for omission German > Preferences for omission Italian*. In the condition *IniObjIndef* we observed more omissions in Dutch and German than in Italian, and a higher tendency to omit in Dutch than in German. In no condition did we find an opposite effect (e.g. more omissions in Italian than in German, more omissions in German than in Dutch). The difference in the value of relative entropy makes the right predictions in these conditions.

- The fact that in the conditions *IniSubDef*, *IntObjDef*, *IniSubIndef* and *IniObjIndef* we find differences between the preferences for omission of *definite* and *indefinite* articles in German fits in the account I proposed in chapter 4, section 4.4.3.2 for a similar difference in these conditions in Italian. There I showed that the nouns we used in the conditions with definite articles were inherently unique nouns. I argued that the retrieval time of an indefinite article is longer than the retrieval time of a definite article before an inherently unique noun, and that this explains why we find a higher preference for omission of indefinite articles.¹¹⁰

¹¹⁰ The crosslinguistic differences in the definiteness effect can be explained by the account proposed in chapter 4, section 4.4.3.2 The relative entropy (in the article set in general) is higher in Dutch than in German, and higher in German than in Italian. The difference in probability distribution of the articles in the case of an inherently unique noun will lead to a decrease in the value of relative entropy of the article set associated with the noun, but, as I argued in chapter 5, the *order* of the entropy values will not

However, in two experimental conditions we found results that, at first sight, cannot be accounted for straightforwardly by differences in the relative entropy value:¹¹¹

- (A) In the condition in which the finiteness effect was tested (section 5.3.1) German was not different from Dutch
- (B) In the conditions where the finite auxiliary verb was omitted the difference between the languages does not reach significance. We find a high preference for omission in all three languages.

Let me start with (A). In the condition where a finite auxiliary verb was present, German was not different from Dutch. However, both German and Dutch were different from Italian. We would expect that the omissions in German were in-between the omissions in Italian and Dutch. In this condition the prediction was not borne out. How can we explain this finding?

We may speculate that since in this condition passive constructions were used, the fact that the relative entropy value alone does not make the right prediction may be related to the differences in processing resources that are necessary to form a passive construction. And, in fact,

change. Thus, selecting an article still takes longer in Dutch than in German, and longer in German than in Italian. Hence, we still find: retrieval time (t) article in Italian < retrieval time (t) article in German < retrieval time (t) article in Dutch. A headline writer strives to maximize the use of the information channel. In a formula: the goal is to make R (the rate of encoding, or the actual channel use, $R = I/t$) come as close as possible to the maximal channel capacity, $C_{max} = I_{max}/t$. Use of a definite article leads to an increase of I , and to an increase of t . The small increase of I in Italian may be compatible with the resulting increase in t , and therefore we find a very low preference for omission. In Dutch and German the increase in t may be disproportionately higher than the increase of I , hence lead to a less optimal channel use, and therefore we find a preference for omission. This preference is stronger in Dutch than in German, because the retrieval time t article German < retrieval time article t Dutch.

¹¹¹ It is, however, important to note that in none of these apparently problematic conditions do we find an opposite effect, that is in none of these conditions do we find more omissions in Italian than in German or more omissions in German than in Dutch. In fact we even find a small tendency for more omissions in the expected direction, hence more omissions in German than in Italian and more omissions in Dutch than in German. However the difference does not reach significance.

the experiment provides us with preliminary evidence for this tentative hypothesis.

In the Dutch experiment we used (perfective) passive constructions that were all formed by a form of 'zijn' (be) + a past participle, as illustrated in example 11:

- (11) HET VREDESPLAN VOOR HAITI IS VERWORPEN
The peace plan for Haiti is rejected

In the German experiment we used two different types of passive constructions:

- passive constructions in the perfective, which were formed by 'sein'+ past participle. These were similar to the passive constructions we used in Dutch (see example 12)
- passive constructions which were formed by a simple past/present tense form of 'werden' + past participle, which had no equivalent in the Dutch version of the experiment (see example 13)

- (12) DER MORD AN ANGELA B. IST AUFGEKLÄRT
The murder of Angela B. is solved.

- (13) DIE VERURTEILUNG VON SCHLEUSERN WIRD
VERTAGT
The conviction of Schleusern is postponed.

Interestingly, there is a difference in the preference responses we found for the two construction types. The German constructions where the passive was formed in the same way as in the Dutch constructions, ('sein'+ past participle) led to preferences for omission in 55,5% of the cases. The constructions where the passive was formed by a simple past/present form of 'werden' led to preferences for omission of 77,5%. This means that the German preferences for omission are in conformity with what would be expected by the differences in relative entropy value. They are in between the preferences for omission in Italian and Dutch, but *only* if the passive construction is in the perfective tense, formed by a form of 'sein'. If the construction is formed by a simple form of 'werden'

plus participle, German shows a higher preference for omission than would be expected.¹¹²

Let me now discuss (B). In the conditions where the finite auxiliary verb was omitted we found no differences between the three languages. We find a high preference for omission in all three languages. Let me explain how we can account for this in the model I propose. Recall that, as we saw in chapter 2, headline readers expect and want to get fast ‘information retrieval’. Headlines have to provide a high amount of information in a short time, and should do so right from the beginning of the headline.

Let me explain what happens in the headlines in which the finite auxiliary is omitted, using the terminology of the model of article selection which I proposed and outlined in chapter 4. I will start with a headline without an article, like in example 14.

(14) HOOLIGAN ARRESTED IN ROME

The processor starts with the noun in initial position, preceding the non-finite verb. Since the article is omitted, the channel can (and will) be used very effectively, with an actual channel use $(R) = I/t$ as close as possible to the maximum capacity level. And this is what we expect, if we process a headline: a high amount of information, in a short time. Then the processor encounters a non-finite verb, containing a relatively high amount of information (as it is non-finite). Since we were already using the channel very effectively, with high processing speed per unit of information, right from the beginning of the sentence on, we can process the high amount of information the non-finite verb contains *fast*. No problems arise. Hence, the headline is fine and the reader gets what he wants: a high amount of information, in a short time.

¹¹² This of course leaves us with the question what could be the reason for the difference in preference in omission of articles in ‘sein’-passives and ‘werden’-passives. We may speculate for example that there are differences in the processing resources required for the formation of a *perfective passive construction*, like ‘IST aufgeklärt’ and a construction with a *simple form* of ‘werden’, like ‘WURDE aufgeklärt’, which influence the use of the article in the two construction types. However, further research, for example based on entropy calculations of sets of auxiliaries, is needed to confirm this speculative account.

Let us now look at the same headline, but now with the article, as in example 15:

(15) A HOOLIGAN ARRESTED IN ROME

The processor starts with the noun, in initial position, preceding the non-finite verb. Since the article is used, the channel cannot be used very effectively, the actual channel use $(R) = I/t$ is further away from the maximum capacity level than in the version without the article. (Note that the processing time per unit of information is higher in A HOOLIGAN than in HOOLIGAN). Per unit of time less information is sent through the channel than in the version without the article. It is not what we expect, if we process a headline, but, still, it can be handled, and the reader may (reluctantly) accept it (with Dutch readers being the most reluctant, German readers a little less and Italian readers least reluctant). But then, the processor encounters a non-finite verb, with a high amount of information. It is important to observe that in this situation, differently from the version without the article, the reader encounters this verb that contains a high amount of information when the processor is working at a slower rate than in the version without the article. The goal of a headline is to *enable* the reader to process the information *as fast as possible*. This means that the reader, at this point, has to speed up the processing rate (increase the amount of information that is sent through the channel per unit of time). If he does not, he has to accept a relatively 'long' time to process the verb (but this long processing time is incompatible with the expectation of fast information retrieval of the headline reader). Thus, the reader has to change to a higher gear, he has to accelerate the processing speed per unit of information. Even if this acceleration is possible (assuming that changing to a higher gear is possible as long as the actual amount that is sent through the channel does not exceed the maximal channel capacity, $R < C$), if asked which headline we prefer, we prefer the first version, where we already arrive in higher gear at the point of the non-finite verb.

Let me illustrate this with fictitious numbers: in example 15 we have to accelerate from for example a processing speed of 4 bits per unit of time to 5 bits per unit of time, while in example 14 we arrive with a higher processing speed, perhaps something like 4.2. The reluctance to change to a higher gear is the same, in all three languages. It does not (only) depend on the value of entropy of the articles, it is mainly caused by the informative load of the non-finite verb, The processing resources

that are required to process this (high) informative load as fast as possible exceed the actual use that is made of the channel at that specific point in the processing of the headlines, in all three languages.¹¹³

Summarizing we can say that the headline with a non-finite verb with an article (A HOOLIGAN ARRESTED IN ROME) is unwanted for two reasons:

- First, the article at the beginning of the headline is unwanted in light of the reader's expectation of fast information processing: the time necessary to acquire one unit of information is (too) high.
- Second, after the reader has (reluctantly) accepted the lower processing speed caused by the time necessary to process the article, the reader encounters a non-finite verb in the headline. This will disturb the reader even more. First he has to accept a lower than preferred processing rate, and now, all of a sudden, he is forced to speed up (or to accept an even longer processing time for the headline). This reluctance to change gear (or accept a longer processing time) will be the same in all three languages, as it does not depend on the value of entropy of the articles. It is caused by the informative load of the non-finite verb. The processing resources that are required to process this (high) informative load exceed the actual use that is made of the channel at that specific point in the processing of the headlines, in all three languages. And that is why we find no differences between the languages: in all three languages we find a strong preference for omission.

5.6 Conclusion and further predictions

In chapters 3 and 4 I discussed the results of my study on article omission in Dutch and Italian child speech and headlines, and proposed an account based on differences in processing resources necessary to

¹¹³ It may even be the case that it is not at all possible to 'change to a higher gear' when we are in the course of processing a sentence. In that case the reader will have to accept a longer processing time, which is incompatible with the goal of fast information retrieval of a headline.

retrieve the articles. In this chapter I repeated the study with German. The results showed that in the majority of the examined cases German was in-between Dutch and Italian. I observed more omissions, or a higher preference for omission, in Dutch than in German, and likewise more in German than in Italian.

These findings cannot easily be explained by an account that relates differences in omission of articles to specific characteristics of Germanic and Romance languages (as for example differences in parameter setting). These accounts would predict that German would behave similar to Dutch with respect to article omission. Moreover, these accounts would predict more categorical judgements, rather than the observed optionality in preferences.

Further, the results for German provide additional evidence for the fact that structural accounts do not always make the right predictions with respect to article omission. In preposition contexts, for example, the three languages behave differently: in all three languages we find less omissions if the preposition is present. However, only the findings on Italian confirm the prediction of structural accounts. We hardly find any omission of the article after a preposition in Italian. In German and Dutch we find optionality: sometimes the article is used, sometimes it is omitted. And here too, we find more omissions in Dutch than in German.

I calculated the relative entropy value of the German article set. This value is in-between the value found for Dutch and Italian. In this chapter I showed that it is exactly this order of entropy values that accounts for all the observed crosslinguistic differences in omission pattern of articles. At this point the question should be asked what predictions we could make if the Relative Entropy value of the article set in language A would be 0.85 and in language B 0.87. Do we still expect to find differences in omission pattern then? The argumentation in chapter 4, section 4.5 on the comparison of the predicted effect of large and small differences in informative load on the processing of articles, should also hold for differences in relative entropy value. Relative entropy influences the processing cost that are required for the selection of an article from a set. The processing cost are higher if the relative entropy value is higher. The higher the differences between the entropy values, the higher the probability that speakers' sensitivity to the entropy value will be reflected in the omission rate. In other words, the higher the differences between the entropy values, the higher the probability that, in a situation of limited processing time or limited processing resources, speakers in the

language with the higher entropy value will omit significantly more articles than speakers in the language with the lower entropy value. It is obvious that if we compare sets with relative entropy values of 0.73 and 0.99 we will find larger differences between omission rates than if we compare sets with relative entropy values of 0.85 and 0.87. However, also in the comparison of the latter set we should find at least a tendency for less omissions in the set with the lower relative entropy value.

CONCLUSION

The results of the crosslinguistic database analysis and experiments presented in this study have shown that there are interesting similarities between the pattern of article omission by adults in newspaper headlines and the pattern of article omission observed in language acquiring children. Looking at Dutch and Italian I found that in both headlines and child speech we find more omissions in Dutch. Further, we observed that in both languages and in both groups of speakers article omission is optional. Moreover, article omission in both languages and both categories of speakers appears to be influenced by the presence of a finite verb or a preposition and by the sentence position of the article-requiring noun. Although this effect works in the same direction in the two languages (less omissions if the verb is finite, less omissions after a preposition, less omissions in sentence-internal position), the concrete consequences are different in both languages. The aim of my study was to come to a unified account for these findings. More in particular, the aim was to come to a model of child language development that accounts for omission of articles and for crosslinguistic differences in this omission pattern, and to show that this model can at the same time explain why we find similar article omission patterns and crosslinguistic differences in adults' special registers.

The basis of the model of article selection I propose to explain these findings is introduced in chapter 2. It is a combination of Levelt's speech production model, some of the ideas expressed in Avrutin and the model for article selection proposed by Caramazza and colleagues and describes the process of article selection under normal and specific circumstances, with normal speakers and speakers with limited processing resources. According to this base model, for an article to be produced, the targeted form needs to receive the highest activation among the members of the set it is a member of. The studies of Caramazza and his colleagues have shown that other members of the article set too receive a certain amount of activation. Thus the final output of the article selection process is the result of a competition between all members of the set. In a sense, the article selection process therefore can be characterized as resolving *uncertainty* with regard to which element of the set will win the competition. Selecting an article from a set with a high level of

uncertainty may be a more 'costly' operation, from a processing point of view, than selecting an article from a set with a lower level of uncertainty. If article selection is a relatively 'easy' process in a language, the activation level necessary for determiner selection will be low. If the activation level is low, an article can be selected, in spite of the fact that the available processing resources of the speaker or the available processing time are low. If, however, because of a higher uncertainty level of the article set, the necessary activation level is high, the outcome of the process of article selection may be influenced by the amount of processing resources or processing time the speaker has available.

In chapter 4 I have shown that the uncertainty level of a set is not just an abstract, theoretical notion, but that it can be measured in an objective way. In fact it is the basic notion of the Information Theoretical approach. In this study I presented an information-theoretical method, adopted from Kostić, to calculate the uncertainty level in the article set, the relative entropy value. The relative entropy of the article system reflects the differences in probability distribution between the elements: the higher the differences in probability distribution (and, thus the higher the value of relative entropy), the more processing resources are required to select an article from the article set. I showed that the relative entropy value in Dutch is higher than the relative entropy value in Italian. I argued that both children and headline writers are sensitive to differences in the processing resources necessary for article selection, children because of their limited processing resources, headline writers (and readers) because of the restricted time readers have available for reading the headline. I showed that with the measure of relative entropy we can account not only for the crosslinguistic differences in omission pattern, but also for the relation between article omission and finiteness, for the influence of the use of a preposition before the article-requiring noun on the omission of articles, for the relation between article omission and sentence position and for differences in omission pattern of different articles (particularly the differences in omissions of 'de' and 'het' in Dutch and the differences between omission of definite and indefinite articles in Italian).

As a further test of the strength of relative entropy as a measure of the complexity level of the article set I used data from a third language, German. I calculated the relative entropy value of the German article system. This value was 'in-between' the Dutch and Italian values. If article omission in child speech and headlines is related to the amount of processing resources necessary to select an article, and if the differences

in the required processing resources are reflected by differences in relative entropy, we should expect the omission rate in German to be 'in-between' the omission rates found for Dutch and Italian. And this was what I found in a repetition of the Dutch-Italian study on article omission in child speech and headlines with German.

The current study shows that both children and adults omit articles because of the processing cost necessary to retrieve the articles from the article set. Children have a limited amount of processing resources, headline writers have to cope with time restrictions. Relative entropy reflects the differences in probability distribution between the elements in a set. The crucial factor in the omission of articles is the optimal use of the (available) information channel capacity, which effectiveness is measured by the amount of information transmitted per unit of time. There is no difference between the brain maturation of Dutch, German and Italian children, nor is there a difference in the intentions of Dutch, German and Italian headline writers, or the amount of space available in their respective publications. What differs is the level of complexity of the article selection process in the three languages. This differs in such a way that producing an article in Italian becomes possible at an earlier age, and that omitting an article in Italian becomes less advantageous in the headline writers' search for the best possible headline.

Bibliography

- Alario, F.X. and Caramazza, A. (2002). The Production of Determiners: Evidence from French, *Cognition*, 82, 179-223.
- Anderson, J.R. and Reder, L.M.(1999). The Fan Effect: New Results and New Theories. *Journal of Experimental Psychology*, Volume 128, 2, 186-197.
- Arnold, E.C. (1969). *Modern Newspaper Design*. New York: Harper & Row.
- Avrutin, S. (1999). *Development of the Syntax-Discourse Interface*, Dordrecht: Kluwer Academic Publishers.
- Avrutin, S. (2004a). Beyond Narrow Syntax. In L. Jenkins (Ed.), *Genetics of Language*, London: Routledge.
- Avrutin, S. (2004b). Optionality in Child and Aphasic Speech. *Lingue e Linguaggio*, 1(1), 67-93.
- Baauw, S., de Roo, E. and 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.
- Baauw, S., Avrutin S., De Lange, J. (2005). The omission of D and T in Dutch child language, (submitted).
- Baddeley, R., Hancock, P., Földiák, P. (2000). *Information Theory and the Brain*. Cambridge: Cambridge University Press.
- Baayen, R.H., Dijkstra, T. and Schreuder, R. (1997). Singulars and Plurals in Dutch: Evidence for a Parallel Dual-route Model. *Journal of Memory and Language*, 37, 94-117
- Bates, E., Dale, P.S. and Thal, D. (1995). Individual Differences and Implication for Theories of Language Development. In P. Fletcher and B. MacWhinney (Eds.), *Handbook of Child Language*, Oxford: Blackwell.
- Bell, A. (1991). *The Language of the News Media*. Oxford: Blackwell Publishers.
- Blake, B.J. (1994). *Case*. Cambridge: Cambridge University Press.
- Blom, E. (2006). The Acquisition of Dutch Gender by early and late Learners, *Paper presented at ELiTU*, Utrecht 19 december.
- Bock, K. (1989). Closed-class Immanence in Sentence Production, *Cognition*, 31, 163-186.

- Bol, G. W. (1996). Optional Subjects in Dutch Child Language. In C. Koster and F. Wijnen (Eds.), *Proceedings of the Groningen Assembly on Language Acquisition*, 125–135.
- Borsley R.D. and Kathol, A.. (2000). Breton as a V2 language. *Linguistics*, 38, 665-710.
- Bradley, D.C.(1978). *Computational Distinctions of Vocabulary Type*. Doctoral dissertation, MIT
- Burzio, L.(1981). *Intransitive verbs and Italian auxiliaries*. Dordrecht: Reidel.
- Caramazza, A., Miozzo, M., Costa, A., Schiller, N., and Alario, F.X. (2001). A crosslinguistic investigation of Determiner Production. In Emmanuel Dupoux, (Ed.), *Language, Brain and Cognitive Development, essays in Honor of Jacques Mehler*. Cambridge: MIT Press.
- Chierchia, G. (1998). Reference to Kinds across Languages. *Natural Language Semantics* 6, 339–405.
- Chierchia, G., Guasti, M.T. and Gualmini, A. (1999). Nouns and Articles in Child Grammar and the Syntax/Semantics Map. *Paper presented at GALA, Potsdam*
- Chomsky, N. (1956). Three models for the Description of Language. I.R.E. *Transactions of Information Theory*, IT-2, 113-124.
- Chomsky, N. (1970). Remarks on Nominalization. In Jacobs, R.A. & P.S. Rosenbaum (Eds.), *Readings in English Transformational Grammar*. Waltham, Massachusetts: Ginn and Company.
- Chomsky, N. (1977). On Wh-Movement. In Culicover, P., Wasow T. and Akmajian, A., (Eds.), *Formal Syntax*. New York: Academic Press.
- Chomsky, N. (1986). *Knowledge of Language, its Nature, Origin and Use*. New York: Praeger
- Chomsky, N.(1993). A Minimalist Program for Linguistic Theory. In Hale, K. and Keyse, S.J. (Eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger* (pp. 1-52). Cambridge, Massachusetts.: MIT Press.
- Chomsky, N. (1995). *The Minimalist Program*, Cambridge, Massachusetts.: MIT Press
- Cinque, G. (1983). Topic Constructions in some European Languages and ‘Connectedness’. In Ehlich K. and van Riemsdijk H.(Eds.), *Connectedness in Sentence, Discourse and Text* Proceedings of the Tilburg Conference held on 25 and 26 January 1982, (pp. 7-43). Tilburg Studies in Language and Literature.
- Cipriani, P., Pfanner, P., Chilosi, A., Cittadoni, L., Ciuti, A., Maccari, A., Pantano, N., Pfanner, L., Poli, P., Sarno, S., Bottari, P., Cappelli, G., Colombo, C. and Veneziano, E. (1989). *Protocolli Diagnostici e Terapeutici*

- nello Sviluppo e nella Patologia del Linguaggio* (1/84 Italian Ministry of Health). Stella Maris Foundation.
- Clahsen, H., Eissenbeiss, S. and Vainikka, A. (1994). The Seeds of Structure: A Syntactic Analysis of the Acquisition of Case Marking. In T. Hoekstra and B. Schwartz (Eds.), *Language Acquisition Studies in Generative Grammar*, Amsterdam: John Benjamins.
- Clahsen, H., Eisenbeiss, S. and Penke, M. (1996). Lexical Learning in Early Syntactic Development. In H. Clahsen (Ed.), *Generative Perspectives on Language Acquisition, Empirical Findings, Theoretical Considerations and Crosslinguistic Comparisons* (pp.129-160). Amsterdam: John Benjamins.
- Clark, R. (2001). Information Theory, Complexity and Linguistic Descriptions. In S. Bertolo (Ed.), *Language Acquisition and Learnability* (pp. 126-171). Cambridge: University Press.
- Cole, P., Segui, J. and Taft, M. (1997). Words and Morphemes as Units for Lexical Access. *Journal of Memory and Language*, 37, 312-330.
- Costa, A., Sebastián-Galles, N., Miozzo, M. and Caramazza, A. (1999). The Gender Congruity Effect: Evidence from Spanish and Catalan. *Language and Cognitive Processes*, 14 (4), 381-391.
- Cover, T.M. and Thomas, J.A. (1991). *Elements of Information Theory*. New York: John Wiley & Sons Inc.
- Dardano, M. (1981). *Il Linguaggio dei Giornali Italiani*. Roma: Laterza.
- De Bot, K. (2000). A Bilingual Production Model: Levelt's 'Speaking' Model Adapted. In Li Wei (Ed.), *The Bilingualism Reader*, London, Routledge.
- Dell, G.S. (1985). Effect of Frequency and Vocabulary Type on Phonological Speech Errors. Paper presented at the meeting of the Psychonomic Society, Boston, Massachusetts.
- Dell, G.S. (1986). A Spreading Activation Theory of Retrieval in Sentence Production. *Psychological Review*, 93-3, 283-321.
- Dell, G.S., Burger, L.K. and Svec, W.R. (1997). Language Production and Serial Order: a Functional Analysis and a Model. *Psychological Review*, 104, 123-144
- De Lange, J., Avrutin, S. and Guasti, M.T. (2006). Crosslinguistic Differences in Child and Adult Speech Optional Omissions: A Comparison of Dutch and Italian. *Proceedings of GALA 2005*.
- De Mauro, T., Mancini, F., Vedovelli, M. and Voghera, M. (1993). *Lessico di Frequenza dell'Italiano Parlato*. Milano: Etaslibri.

- Demuth, K. (1994). On the Underspecification of Functional Categories in early Grammars. *Syntactic Theory and First Language Acquisition*, Vol 1. Hillsdale New Jersey: Lawrence Erlbaum Associates, Inc.
- De Roo, E. (1999). *Agrammatic Grammar. Functional Categories in Agrammatic Speech*. Den Haag: Holland Academic Graphics.
- De Roo, E., Kolk, H. and Hofstede, B. (2003). Structural Properties of Syntactically Reduced Speech: A Comparison of Normal Speakers and Broca's Aphasics. *Brain and Language*, 86, 99-115.
- Dor, D.(2003). On Newspaper Headlines as Relevance Optimizers. *Journal of Pragmatics*, 35, 695-721.
- Enç, M. (1987). Anchoring Conditions for Tense, *Linguistic Inquiry*, 18, 633-657
- Escobar, L. (1995). *Leftband Satellites in Spanish*. Doctoral Dissertation, Utrecht University, OTS Dissertation Series.
- Frazier, L.(1987). Sentence Processing: A Tutorial Review. In M.Coltheart (Ed.), *Attention and performance*, vol.XII. London:Erlbaum, 559-585.
- Friederici, A.D.(2002). Towards a Neural Basis of Auditory Sentence Processing. *Trends in Cognitive Science*, 6, 78-84.
- Gaillard, W.D., Hertz-Pannier, L., Mott, S.HI, Barnett, A.S., LeBihan, D., Theodore, W.H.(2000). Functional Anatomy of Cognitive Development:fMRI of Verbal Fluency in Children and Adults. *Neurology*, 54(1), 180 -185.
- Garrett, M.F. (1982). Production of Speech: Observations from Normal and Pathological Language Use. In A. Ellis (Ed.) *Normality and pathology in cognitive functions*. London: Academic Press, 19-76.
- Gerken, L.A.(1991). The Metrical Basis for Children's Subjectless Sentences. *Journal of Memory and Language*, 30, 1-21.
- Goldin-Meadow, S. (2006). Talking and Thinking with our Hands. *Current Directions in Psychological Science*, 15, 34-39.
- Goodglass, H. and Kaplan, E. (1972). *The assessment of Aphasia and Related Disorders*. Philadelphia: Lea & Febiger.
- Guasti, M.T.(1993/1994). Verb Syntax in Italian Child Grammar: Finite and Non-finite Verbs. *Language Acquisition* 3, 1-40.
- Guasti, M.T., Gavarró, A., de Lange, J. and Caprin, C. (2008). Article omission across languages (to appear in *Language Acquisition*).
- Guasti, M. T., de Lange, J., Gavarró, A.and Caprin, C. (2004). Article Omission: across Child Languages and across Special Registers. In J. van Kampen and S. Baauw (Eds), *Proceedings of GALA 2003*. Utrecht: LOT Occasional Series, vol. 1, 199–210.

- Guéron, J. and Hoekstra, T. (1995). The Temporal Interpretation of Predication. In A.Cardinaletti and M.T. Guasti (Eds.), *Syntax and Semantics, Volume 28, Small Clauses*, San Diego: Academic Press.
- Haegeman, L. (1990). Non-overt Subjects in Diary Contexts. In J. Mascaro and M. Nespore (Eds.), *Grammar in Progress* (pp.167 – 174). Dordrecht: Foris
- Hagoort, P. (2003). How the Brain solves the Binding Problem for Language: a Neurocomputational Model of Syntactic Processing. *NeuroImage* 20, S18-S29.
- Haider, H. (1984). Was zu haben ist und was zu sein hat. *Papiere zur Linguistik*, 30, 23-36
- Hankamer, J. (1971). *Constraints on Deletion in Syntax*, PhD dissertation, Yale University.
- Heim, I., (1982), *The Semantics of Definite and Indefinite Noun Phrases*, Diss. U. Mass., Amherst
- Heim, I. (1991). Articles and Definiteness. Published in German as “Artikel und Definitheit”. In A.v.Stechow & D. Wunderlich (Eds), *Semantics: An international handbook of contemporary research*. Berlin: de Gruyter.
- Hoekstra, T. (1986). Passives and Participles. *LIN-Bundel 1986*, 95-104.
- Hoekstra, T., Hyams, N., and Becker, M. (1996). The Underspecification of Number and the Licensing of Root Infinitives. *Proceedings of the annual Boston University Conference on Language Development 21* (pp.293-306). Somerville, Massachusetts: Cascadilla Press.
- Hoekstra, T. and Hyams, N. (1998). Agreement and Finiteness of V2: Evidence from Child Language, In A.Greenhill, M. Hughes, H. Littlefield and H. Walsh (Eds.), *Proceedings of the annual Boston University Conference on Language Development 22* (pp360-373). Somerville, Massachusetts: Cascadilla Press.
- Holland, S.K., Plante, E., Weber Byars, A., Strawsberg, R.H., Schmithorst, V.J., Ball, W.S.(2001). Normal fMRI Brain Activation Patterns in Children Performing a Verb Generation Task. *NeuroImage*, 14, 837-843.
- Horton, W.S.and Keysar, B. (1996). When do Speakers take into Account Common Ground? *Cognition* 59, 91-117.
- Hyams, N. (1996).The underspecification of functional categories in early grammar. In H. Clahsen (Ed.) *Generative perspectives on language acquisition*. Amsterdam:John Benjamins.

- Indefrey, P., Brown, C.M, Hellwig, F., Amunts, K., Herzog, H., Seitz, R.J. and Hagoort, P. (2001). A Neural Correlate of Syntactic Encoding during Speech Production. *PNAS* 98, 5933-5936.
- Janssen, N. and Caramazza, A. (2003). The Selection of Closed-class Words in Noun Phrase Production: The Case of Dutch Determiners. *Journal of Memory and Language* 48, 635-652.
- Kamide, Y., Altmann, G.T.M. and Haywood, S.L. (2003). The Time Course of Prediction in Incremental Sentence Processing: Evidence from Anticipatory Eye Movement. *Journal of Memory and Language*, 49, 133 -156.
- Keysar, B., Barr, D.J., Balin, J.A. (1998). Definite Reference and Mutual Knowledge: Process Models of Common Ground in Interpretation. *Journal of Memory and Language* 39, 1-20.
- Klee, T. and Fitzgerald, M.D. (1985). The Relation between Grammatical Development and Mean Length of Utterance in Morphemes. *Journal of Child Language*, 12, 251-269.
- Kolk, H.H.J. (2000). Does Agrammatic Speech constitute a Regression to Child Language? A three-way Comparison between Agrammatic, Child and Normal Ellipsis. *Brain and Language*, 74, 549-551.
- Kostić, A. and Katz, L. (1987). Processing Differences between Nouns, Adjectives and Verbs. *Psychological Research*, 49, 1, 229-336.
- Kostić, A. (2004). The Effects of the Amount of Information on Processing of Inflected Morphology, ms. University of Belgrade.
- Kostić, Đ. (1965). Sintaktičke funkcije padež "ih oblika u srpskohrvatskom jeziku Beograd. Institut za eksperimentalnu fonetiku i patologiju govora.
- Kronrod, A. and Engel, O. (2001). Accessibility Theory and Referring Expressions in Newspaper Headlines. *Journal of Pragmatics*, 33, 683-699.
- Kupisch, T. (2006). The Emergence of Article Forms and Functions in a German-Italian Bilingual Child. In C.Lléo (Ed.), *Interfaces in Multilingualism: Acquisition, Representation and Processing* (pp.45-109). Amsterdam: John Benjamins.
- Kupisch, T. (2007). Testing the Effects of Frequency on the Rate of Learning: Determiner Use in Early French, German and Italian. In I.Gülzow, N.Gagarina (Eds.), *Frequency effects in language acquisition*. Berlin: Mouton de Gruyter.
- La Heij, W., Mark, P., Sander, J. and Willeboordse, E. (1998). The Gender Congruency Effect in Picture Word Tasks. *Psychological Research*, 61, 209 – 219.

- Landau, I. (2004). The Scale of Finiteness and the Calculus of Control. *Natural Language & Linguistic Theory*, 22, 811-877.
- Levelt, W.J.M.(1989). *Speaking. From intention to articulation*. Cambridge M.A.: MIT Press.
- Levelt, W.J.M. (2001). Spoken Word Production: A Theory of Lexical Access. *PNAS* 98, 13464-13471.
- Levelt, W.J.M., Roelofs, A. and Meyer, A.S. (1999). A Theory of Lexical Access in Speech Production. *Behavioral and Brain Sciences*, 22, 1-75.
- Lightfoot, D. (1999). *The development of language*. Massachusetts: Blackwell Publishers.
- Linebarger, M., Schwartz, M. and Saffran, E. (1983). Sensitivity to Grammatical Structure in so-called Agrammatic Aphasia. *Cognition*, 13, 361-393.
- Lléo, C. and Demuth, K. (1999). Prosodic Constraints on the Emergence of Grammatical Morphemes: Crosslinguistic Evidence from Germanic and Romance languages, *Proceedings of the 23rd BUCLD*. Somerville: Cascadilla Press, 407-418.
- Longobardi, G. (1994). Reference and Proper Names: a Theory of N-Movement in Syntax and Logical Form. *Linguistic Inquiry*, 25, 609-665.
- Longobardi, G. (2001). How Comparative is Semantics? A Unified Parametric Theory of Bare Nouns and Proper Names. *Natural Language Semantics* 9, 335-369.
- Lukatela, G., Carello, C. and Turvey, M.T. (1987). Lexical Representation of Regular and Irregular Inflected Nouns. *Language and Cognitive Processes*, 2\1, 1-17.
- MacWhinney, B. and Snow, C. (1985). The Child Language Data Exchange System. *Journal of Child Language* 12, 271-296.
- Magni, M. (1992). *Lingua Italiana e Giornali d'Oggi*. Milano: Miano.
- Maiden, M. and Robustelli, C. (2000). *A Reference Grammar of Modern Italian*. London: Arnold.
- Manelis, L. and Tharp, D. (1977). The Processing of Inflected Words. *Memory and Cognition*, 5 (6), 690-695.
- Marantz, A. (1995). The Minimalist Program. In G. Webelhuth (Ed.), *Government and Binding Theory and the Minimalist Program*. Oxford: Blackwell.
- Mardh, I. (1980). *Headlines: on the grammar of English front page headlines*. Gleeurp: C.W.K.
- Mehler, J., Dupoux E., Nazzi, T. and Dehaene-Lambertz, G. (1996). Coping with Linguistic Diversity: The Infant's Viewpoint. In James I.

- Morgan and K. Demuth (Eds.), *Signal to syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp.101-116). Mahwah, NJ, Erlbaum.
- Mehler, J., Sebastian, N., Altmann, G., Dupoux, E., Christophe, A. and Pallier, C. (1993). Understanding Compressed Sentences: The Role of Rhythm and Meaning. In P. Tallal, A.M. Galaburda, R.R. Llinás and C. Von Euler (Eds), *Temporal Information Processing in the Nervous System: Special Reference to Dyslexia and Dysphasia. Annals of the New York Academy of Sciences*, 682, 272-282
- Menn, L. and Obler, L. (1990). *Agrammatic Aphasia: a Cross-Language Narrative Sourcebook*. Amsterdam: John Benjamins.
- Miozzo M. and Caramazza, A. (1999). The Selection of Determiner in Noun Phrase Production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 25, 907-922.
- Moscoso del Prado, M., Kostić, A. and Baayen, H. (2004). Putting the Bits together: An Informational Perspective on Morphological Processing. *Cognition* 94(1), 1- 18.
- Müller O. and Hagoort, P. (2006). Access to Lexical Information in Language Comprehension: Semantics before Syntax. *Journal of Cognitive Neuroscience*, 18:1, 84-96.
- Neville, H., Mills, D.L and Lawson, D.L. (1992). Fractionating Language: Different Neural Subsystems with different sensitive Periods. *Cerebral Cortex*, 2, 244-258.
- Nir, R. (1993). A Discourse Analysis of News Headlines. *Hebrew Linguistics*, 37, 23-31.
- Nortier, J. (1995). Code Switching in Moroccan Arabic / Dutch vs. Moroccan Arabic/French Language Contact. *International Journal of the Sociology of Language*, 112, 81-95.
- Nortier, J.(2001). *Murks en Straattaal*. Amsterdam: Prometheus.
- Ouhalla, J. (1994). *Introducing Transformational Grammar. From Rules to Principles and Parameters*. London: Edward Arnold.
- Ouhalla, J.(1999). *Introducing Transformational Grammar. From Principles and Parameters to Minimalism*. London: Edward Arnold.
- Pascal, B. (1656). *Lettres provinciales*.
- Pérez-Leroux, A.T. and Roeper, T. (1999). Scope and the Structure of Bare Nominals: Evidence from Child Language. *Linguistics* 37, 927 – 960.
- Perlmutter, D. (1978). Impersonal Passives and the Unaccusative Hypothesis. *Proceedings of the 4th annual meeting of the Berkeley Linguistic Society*.

- Peters, A.M. (2001). Filler Syllables: What is their Status in Emerging Grammar? *Journal of Child Language*, 28, 229-242.
- Phaff, A.(2006). *Lidwoordgebruik in Kindertaal*. Bachelor-eindwerkstuk, Universiteit Utrecht.
- Philips, C. (2003). Linear Order and Constituency. *Linguistic Inquiry*, 34, 37-90.
- Piñango, M. and Burkhardt, P. (2001). Pronominals in Broca's Aphasia Comprehension: the Consequences of Syntactic Delay. *Brain and Language*, 79 (1), 167- 168.
- Prince, E.F. (1992). Subjects, Definiteness and Information Status. In W.C. Mann and S.Thompson (Ed.), *Discourse Description* (pp 295-325). Philadelphia: John Benjamins Publishing Company.
- Prince, E.F.(1985). Fancy Syntax and Shared Knowledge. *Journal of Pragmatics*, 9, 65-81.
- Prince, E.F.(1981). Towards a Taxonomy of Given-New Information. In Cole, P. (Ed.), *Radical Pragmatics* (pp 223-256). New York:Academic Press.
- Prins, R. (1987). *Afasie: Classificatie, Behandeling en Herstelverloop*. Doctoral dissertation, Amsterdam. Dordrecht: ICG Printing.
- Radford, A. (1990). *Syntactic Theory and the Acquisition of English Syntax: the Nature of Early Child Grammars of English*. Oxford: Blackwell Publishers.
- Renzi, L. (1988). *Grande Grammatica Italiana di Consultazione*. Bologna: Il Mulino.
- Rizzi, L. (1993/1994). Some Notes on Linguistic Theory and Language Development: The Case of Root Infinitives. *Language Acquisition* 3, 371 – 393.
- Rizzi, L. (2000). Remarks on Early Null Subjects. In M.A. Friedemann and L.Rizzi (Eds.), *The Acquisition of Syntax*. Harlow, England: Longman.
- Rizzi, L. (2004). From Inventories to Computations: Open/Closed Class Items and Substantive/ Functional Heads. *Dialectica*, 3, 437-451.
- Robertson, D.A., Gernsbacher, M.A., Guidotti, S.J., Robertson, R.R.W., Irvin, W., Mock, B.J. and Campana, M.E. (2000). Functional Neuroanatomy of the Cognitive Process of Mapping during Discourse Comprehension. *Psychological Science*, 11, 255 – 260
- Roelofs, A. (1992). A Spreading-Activation Theory of Lemma Retrieval in Speaking. *Cognition*, 42, 107-142.
- Roelofs, A.(1997). The WEAVER-Model of Word-Form Encoding in Spoken Production. *Cognition*, 64, 249-284.

- Roeper, T., Ramos, E., Seymour H. and Abdul-Karim, L. (2001). Language Disorders as Window on Universal Grammar: An Abstract Theory of Agreement for IP, DP and V-PP. *Brain and Language*, 77, 378-397.
- Rozendaal, M. (2006). Linguistic and Non-linguistic Influences on the Acquisition of Referent Introductions, a Cross-Linguistic Perspective. Paper presented at ELiTU meeting, Utrecht, June 30.
- Ruigendijk, E. (2002). *Case-assignment in Agrammatism: a crosslinguistic study*, Doctoral Dissertation Groningen University.
- Sapir, E. (1921). *Language, an Introduction to the Study of Speech*. New York: Harcourt, Brace and Company.
- Schaeffer, J. and Matthewson, L. (2005). Grammar and Pragmatics in the Acquisition of Article Systems. *Natural Language and Linguistic Theory*, 23, 53-101.
- Schaerlakens, A.M. (1980). *De Taalontwikkeling van het Kind*. Groningen: Wolters Noordhoff.
- Schiller, N.O. and Caramazza, A. (2002). The Selection of Grammatical Features in Word Production: The Case of Plural Nouns in German. *Brain and Language*, 81, 342-357.
- Schiller, N.O. and Caramazza, A. (2003). Grammatical Feature Selection in Noun Phrase Production: Evidence from German and Dutch. *Journal of Memory and Language*, 48, 169-194.
- Schmitt, B.M., Schiltz K., Zaake, W., Kutas M. and Münte T.F. (2001). An Electrophysical Analysis of the Time Course of Conceptual and Syntactical Encoding during Tacit Picture Naming. *Journal of Cognitive Neuroscience*, 13:4, 510-522.
- Schneider, T. (1984). *The Information Content of Binding Sites on Nucleotide Sequences*. PhD thesis University Colorado.
- Schoenenberg, M., Penner, Z. and Weissenborn, J. (1997). Object Placement in Early German Grammar. In: A. Hughes, M. Hughes and A. Greenhill (Eds.), *Proceedings of the annual Boston University Conference on Language Development*, 21. Somerville, MA.: Cascadilla Press.
- Schriefers, H. (1993). Syntactic Processes in the Production of Noun Phrases. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19, 841-850.
- Schriefers, H., Jescheniak, J. D and Hantsch, A. (2002). Determiner Selection in Noun Phrase Production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 941-950.

- Shannon, C. and Weaver, W. (1949). *The Mathematical Theory of Communication*, London: University of Illinois Press.
- Skipper, J.I., Goldin-Meadow, S., Nusbaum, H.C. and Small, S.L. (2007). Speech Associated Gestures, Broca's Aerea and the Human Mirror System. *Brain and Language* 101, 260-277.
- Simon-Vandenberghe, A.M. (1981). *The Grammar of Headlines in The Times 1870 – 1970*. Brussel: Paleis der Academieën.
- Sperber, D. and Wilson, D. (1986). *Relevance: Communication and Cognition*. Cambridge, Massachusetts: Harvard U.P.
- Stowell, T. (1981). *Origins of phrase structure*. Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Stowell, T. (1999). Words Lost and Syntax Found in Headlines: The Hidden Structure of Abbreviated English in Headlines, Instructions and Diaries. *Paper presented at York University, Toronto, November 24*.
- Straumann, H. (1935). *Newspaper Headlines, a Study in Linguistic Method*. London:Unwin Brothers Ltd.
- Stuart, A., Kalinowski, J., Rastatter, M.P. and Lynch, K. (2002). Effect of Delayed Auditory Feedback on Normal Speakers at Two Speech Rates. *The Journal of the Acoustical Society of America* , 111-5, 2237-2241.
- Swinney, D. (2003). Psycholinguistic Approaches. *Paper presented at the 4th Science of Aphasia Conference, Trieste*.
- Taft, M. (1979). Recognition of Affixed Words and the Word Frequency Effect. *Memory and Cognition*, 7(4), 263-272.
- Taft, M. (1981). Prefix Stripping Revisited. *Journal of Verbal Learning and Verbal Behavior*, 20, 289-297.
- Taft, M. (1994). Interactive-Activation as a Framework for Understanding Morphological Processing. *Language and Cognitive Processes*, 9, 271-294.
- Thornton, R. (1996). Elicited Production. In D. MacDaniel, C.MacKee and H. Smith Cairns (Eds.), *Methods for Assessing Children's Syntax*. Cambridge, Massachusetts: MIT Press.
- Todorović, D. (1988). Hemispheric Differences in Case Processing. *Brain and Language*, 33, 365-389.
- M.J. Traxler and Pickering, M.J. (1996). Plausibility and the Processing of Unbounded Dependencies:An Eye-Tracking Study. *Journal of Memory and Language* 35, 454-475.
- Traxler, M.J., Bybee, M. D. and Pickering, M. J. (1997). Influence of Connectives on Language Comprehension: Eye tracking Evidence for Incremental Interpretation. *The Quarterly Journal of Experimental Psychology Section A*, 50- 3, 481 – 497.

- Tribus, M. (1961). *Thermostatistics and Thermodynamics*. Princeton, N.J.: D. van Nostrand Company Inc.
- Valian, V. (1991). *Syntactic Subjects in the Early Speech of American and Italian children*. *Cognition* 40, 21–81.
- Van Der Velde, M. (2004). L'Acquisition des Articles Définis en L1, *Acquisition et Interaction en Langue Étrangère*, 21.
- Van Dijk, T. A. (1988). *News as discourse*. New Jersey: Lawrence Erlbaum associates Publishers.
- Van Kampen, N. J. (1994). The Learnability of the Left Branch Condition. In R. Bok-Bennema and C. Cremers (Eds.), *Linguistics in the Netherlands 1994*. Amsterdam: John Benjamins.
- Van Turenhout, M., Hagoort, P. and Brown, C.M. (1998). Brain Activity during Speaking: from Syntax to Phonology in 40 milliseconds. *Science*, 280, 572-574.
- Vasić, N. (2006). *Pronoun Comprehension in Agrammatic Aphasia, the Structure and Use of Linguistic Knowledge*. Doctoral dissertation. Utrecht Institute for Linguistics OTS. LOT Series.
- Vergnaud, J.R. and Zubizarreta, M.L. (1992). The Definite Determiner and the Inalienable Constructions in French and in English. *Linguistic Inquiry* 23: 595-652.
- Vincenzi, M. and Di Domenico, E. (1999). A Distinction among Φ -Features: The Role of Gender and Number in the Retrieval of Pronoun Antecedents. *Revista di Linguistica*, 11, 1, 41-74.
- Voinescu, I., Fradis, A. and Mihailescu, L. (1967). The First Degree Entropy of Phonemes in Aphasics. *Revue Roumaine de Neurologie*, 4, 1, 67 – 79.
- Ward, G.L. and Prince, E.F. (1991). On the Topicalization of Indefinite NPs. *Journal of Pragmatics*, 16-2, 167-177.
- Zonneveld, W. (1992). Het jonge hoofd, De Righthand Head Rule bij kinderen van 4 tot 7 jaar. *De Nieuwe Taalgids* 85-1, 37-49.
- Zurif, E.B., Swinney, D., Prather, P., Solomon, J. and Bushell, C. (1993). An On-line Analysis of Syntactic Processing in Broca's and Wernicke's Aphasia. *Brain and Language*, 45, 448-464.

Appendix-A: Test-Headlines**Test Headlines Dutch¹¹⁴**

I + II YES/NO AUX.

1. A. Vermiste krantenbezorger (is) teruggevonden.
B. De vermiste krantenbezorger (is) teruggevonden.
2. A. De criminaliteit (is) vorig jaar licht gedaald.
B. Criminaliteit(is) vorig jaar licht gedaald.
3. A. De NOS-internetsite (is) totaal vernieuwd.
B. NOS-internetsite (is) totaal vernieuwd.
4. A. Italiaanse voetbalclub Parma (is) failliet verklaard.
B. De Italiaanse voetbalclub Parma (is) failliet verklaard.
5. A. Complete Bandidos-bende (is) gepakt.
B. De complete Bandidos-bende (is) gepakt.
6. A. Het 'Wandelvrouwtje' (is) beestachtig vermoord.
B. 'Wandelvrouwtje' (is) beestachtig vermoord.
7. A. Het dossier van Mabel (is) gestolen uit archief.
B. Dossier van Mabel (is) gestolen uit archief.
8. A. Onderzoeksinstituut Nyfer (is) mogelijk afgeluisterd.
B. Het onderzoeksinstituut Nyfer (is) mogelijk afgeluisterd.
9. A. Mobiel telefoonverkeer in Madrid (was) overbelast.
B. Het mobiele telefoonverkeer in Madrid (was) overbelast
10. A. Het vredesplan voor Haiti (is) verworpen.
B. Vredesplan voor Haiti (is)verworpen.

III IniSubDef.

1. A. Regering wil vliegend materieel in Irak.
B. De regering wil vliegend materieel in Irak.
2. A. De Britse overheid gaat bezuinigen.
B. Britse overheid gaat bezuinigen
3. A. De politie pakt straatracers aan.
B. Politie pakt straatracers aan.
4. A. Busbranche eist meer concurrentie.
B. De busbranche eist meer concurrentie.
5. A. CDA-top bekritiseert Balkenende.
B. De CDA-top bekritiseert Balkenende.
6. A. Het enthousiasme voor Europa neemt af.
B. Enthousiasme voor Europa neemt af.
7. A. Het parlement pioniert in veenkolonien.
B. Parlement pioniert in veenkolonien.

¹¹⁴ With this overview of headlines native speakers can check their own intuitions on the use or omission of articles in the different conditions discussed in this study.

8. A. Filmmuseum wil naar Amsterdam-Noord.
B. Het filmmuseum wil naar Amsterdam-Noord.
9. A. Nederlands drugsbeleid ontzet Parijs.
B. Het Nederlands drugsbeleid ontzet Parijs.
10. A. Het Terra College moet leraren ontslaan door bezuinigingen.
B. Terra College moet leraren ontslaan door bezuinigingen.

IV IntObjDef.

1. A. Giscard d' Estaing verlaat de Franse politiek.
B. Giscard d' Estaing verlaat Franse politiek.
2. A. Woerden zal cent niet missen.
B. Woerden zal de cent niet missen.
3. A. Beatrix moet troonrede niet meer voorlezen.
B. Beatrix moet de troonrede niet meer voorlezen.
4. A. Duizenden herdenken Februaristaking.
B. Duizenden herdenken de Februaristaking.
5. A. Moodys houdt de Russische kredietstatus op niveau.
B. Moodys houdt Russische kredietstatus op niveau.
6. A. Albers leidt het klassement na vijf overwinningen.
B. Albers leidt klassement na vijf overwinningen.
7. A. Hervormde tegenstanders verliezen het kort geding.
B. Hervormde tegenstanders verliezen kort geding.
8. A. Nesselande krijgt verlossend punt cadeau.
B. Nesselande krijgt het verlossende punt cadeau.
9. A. Halsema blaast systeem op.
B. Halsema blaast het systeem op.
10. A. Balkenende wil het conflict oplossen.
B. Balkenende wil conflict oplossen.

V. IniSubIndef.

1. A. Een politieman doet aangifte van diefstal.
B. Politieaanpak doet aangifte van diefstal.
2. A. Een Indonesier krijgt 10 jaar voor 'Tanjung Priok'.
B. Indonesier krijgt 10 jaar voor 'Tanjung Priok'.
3. A. Amsterdammer bedreigt Balkenende.
B. Een Amsterdammer bedreigt Balkenende.
4. A. Nieuwe fusiegolf komt eraan in medialand.
B. Een nieuwe fusiegolf komt eraan in medialand.
5. A. Een Hagenaar sjoemelde met duizenden energiemeters.
B. Hagenaar sjoemelde met duizenden energiemeters.
6. A. Een familiebedrijf koopt 25 Megapool-winkels.
B. Familiebedrijf koopt 25 Megapool-winkels.
7. A. Lovend advies levert niet altijd geld op.
B. Een lovend advies levert niet altijd geld op.

8. A. Een treinramp in Noord-Korea eist mogelijk 3000 levens.
B. Treinramp in Noord-Korea eist mogelijk 3000 levens.

9. A. Een grotere EU maakt Oostenrijk vuiler.
B. Grotere EU maakt Oostenrijk vuiler.

10. A. Een koffertje zaait paniek in Haarlem.
B. Koffertje zaait paniek in Haarlem.

VI. IntObjIndef.

1. A. Joling krijgt reallife soap.
B. Joling krijgt een reallife soap.
2. A. Aboriginals leveren veldslag in Sydney.
B. Aboriginals leveren een veldslag in Sydney.
3. A. Astronomen ontdekken een diamantster.
B. Astronomen ontdekken diamantster.
4. A. Sharon lijdt een gevoelige nederlaag.
B. Sharon lijdt gevoelige nederlaag.
5. A. F-16's onderscheppen onbekende Boeing.
B. F-16's onderscheppen een onbekende Boeing.
6. A. MVRDV gaat paviljoen bouwen in Londen.
B. MVRDV gaat een paviljoen bouwen in Londen.
7. A. Zes lidstaten willen een lagere bijdrage.
B. Zes lidstaten willen lagere bijdrage.
8. A. Eindhoven kiest een gecontroleerde tippelzone.
B. Eindhoven kiest gecontroleerde tippelzone.
9. A. Ambtenaren dwingen dienstauto af.
B. Ambtenaren dwingen een dienstauto af.
10. A. Opa en oma krijgen eigen feestdag.
B. Opa en oma krijgen een eigen feestdag.

VII + VIII Order of omission, definite articles:

1. A. Kamer verlengt de fiscale regeling voor tweeverdieners.
B. De kamer verlengt fiscale regeling voor tweeverdieners.
2. A. De Duitse regering wil groei aanwakkeren.
B. Duitse regering wil de groei aanwakkeren.
3. A. De beurs begroet verkoop van GNC met gejuich.
B. Beurs begroet de verkoop van GNC met gejuich.
4. A. Krab verovert de Noordzee.
B. De krab verovert Noordzee.
5. A. Betuwelijn blijft de burger geld kosten.
B. De Betuwelijn blijft burger geld kosten.
6. A. De minister past Walviswet aan voor artiesten.
B. Minister past de Walviswet aan voor artiesten.
7. A. De Duitse import overstijgt export.
B. Duitse import overstijgt de export.

8. A. De toetreding vervult droom van vrijheid.
B. Toetreding vervult de droom van vrijheid.
9. A. Nederlandse teamsport is de aansluiting kwijtgeraakt.
B. De Nederlandse teamsport is aansluiting kwijtgeraakt.
10. A. Turkse regering stelt de onderwijswet uit.
B. De Turkse regering stelt onderwijswet uit.

VII + VIII Order of omission, indefinite articles:

1. A. Een trein rijdt winkel binnen in Den Haag.
B. Trein rijdt een winkel binnen in Den Haag.
2. A. Een mollige tiener wint missverkiezing.
B. Mollige tiener wint een missverkiezing.
3. A. Meisje redt een kleuter van verdrinking.
B. Een meisje redt kleuter van verdrinking.
4. A. Ziekenhuis laat een gewonde vrouw op de stoep liggen.
B. Een ziekenhuis laat gewonde vrouw op de stoep liggen.
5. A. Een Steenwijker springt auto na in het water.
B. Steenwijker springt een auto na in het water.
6. A. Een vuurwerkbom vernielt ruit.
B. Vuurwerkbom vernielt een ruit.
7. A. Vrachtwagen ramt een ventilator.
B. Een vrachtwagen ramt ventilator.
8. A. Spaanse regisseur krijgt een oeuvreprijs.
B. Een Spaanse regisseur krijgt oeuvreprijs.
9. A. Een Chinese tank vuurt granaat af.
B. Chinese tank vuurt een granaat af.
10. A. Een Nederlandse kunstenaar wint grote Britse prijs.
B. Nederlandse kunstenaar wint een grote Britse prijs.

IX Nouns in Isolation

1. A. Ondergang van een cowboy.
B. De ondergang van een cowboy.
2. A. De macabere tocht naar Dutrouxs gruwelkooi.
B. Macabere tocht naar Dutrouxs gruwelkooi.
3. A. De jarige Donald Duck in het Cobra museum.
B. Jarige Donald Duck in het Cobra museum.
4. A. Eerste nederlaag voor Oranje.
B. De eerste nederlaag voor Oranje.
5. A. Clash tussen hoofddoek en decollete.
B. De clash tussen hoofddoek en decollete.
6. A. Het lastige kwalificatiepad op de marathon.
B. Lastig kwalificatiepad op de marathon.
7. A. Het drukste station van Spanje.
B. Drukste station van Spanje.

8. A. Recht van het kapmes in Lira.
B. Het recht van het kapmes in Lira.
9. A. Icoon van de cosmetica.
B. Het icoon van de cosmetica.
10. A. Het centrum van de cybermisdaad.
B. Centrum van de cybermisdaad.

X Hanging Topic Constructions.

1. A. Begroting, universiteiten moeten bezuinigen.
B. De begroting, universtiteiten moeten bezuinigen.
2. A. De school, tekort aan jonge docenten.
B. School, tekort aan jonge docenten.
3. A. Het kerstdiner, een uitdaging voor de portemonnee.
B. Kerstdiner, een uitdaging voor de portemonnee.
4. A. Mediawet, regering in grote problemen.
B. De mediawet, regering in grote problemen.
5. A. Geschil tussen Nijs-van der Hoeven, forse kritiek uit Kamer.
B. Het geschil tussen Nijs-van der Hoeven, forse kritiek uit Kamer.
6. A. Het onderzoek naar de cocainehandel, een politieke affaire.
B. Onderzoek naar de cocainehandel, een politieke affaire.
7. A. De hypotheekrenteaftrek, Groen Links hoopt op discussie.
B. Hypotheekrenteaftrek, Groen Links hoopt op discussie.
8. A. Selectie studenten, goed plan volgens universiteiten.
B. De selectie van studenten, goed plan volgens universiteiten.
9. A. Openbaar vervoer, dag vol chaos.
B. Het openbaar vervoer, dag vol chaos.
10. A. De ramp gereconstrueerd, horrorfilm over vliegtuigcrash.
B. Ramp gereconstrueerd, horrorfilm over vliegtuigcrash.

Test Headlines Italian**I + II YES/NO AUX**

1. A. Un comandante (è stato) arrestato per spionaggio.
B. Comandante (è stato) arrestato per spionaggio
2. A. Rapinatore di 17 anni (è stato) ferito al mercato.
B. Un rapinatore di 17 anni (è stato) ferito al mercato
3. A. Un commercialista (è stato) ucciso per ripicca.
B. Commercialista (è stato) ucciso per ripicca
4. A. Un commercialista (è stato) ucciso per ripicca.
B. Commercialista (è stato) ucciso per ripicca
5. A. Camion dei rifiuti partenopeo (è stato) respinto a Milano.
B. Un camion dei rifiuti partenopeo (è stato) respinto a Milano.
6. A. Giornalista (è stata) rapinata in Egitto.
B. Una giornalista (è stata) rapinata in Egitto.
7. A. Una sala bingo (è stata) condannata per oltraggio.
B. Sala bingo (è stata) condannata per oltraggio.
8. A. Scuola (è stata) chiusa per meningite.
B. Una scuola (è stata) chiusa per meningite.
9. A. Una turista (è stata) dispersa in Svizzera.
B. Turista (è stata) dispersa in Svizzera
10. A. Rapina (è stata) sventata da turisti a Venezia.
B. Una rapina (è stata) sventata da turisti a Venezia.

III IniSubDef.

1. A. La fiducia in Berlusconi diminuisce.
B. Fiducia in Berlusconi diminuisce.
2. A. La Ferrari ringrazia Schumacher.
B. Ferrari ringrazia Schumacher.
3. A. Lista Prodi si riunisce a Chianciano.
B. La lista Prodi si riunisce a Chianciano.
4. A. Sentenza Sme porta dritto a Berlusconi.
B. La sentenza Sme porta dritto a Berlusconi.
5. A. La cicogna arriva a casa.
B. Cicogna arriva a casa.
6. A. Il PM chiede condanne fino a 8 anni.
B. PM chiede condanne fino a 8 anni.
7. A. Governo accusa Prodi.
B. Il governo accusa Prodi.
8. A. Governo italiano vuole risparmiare.
B. Il governo italiano vuole risparmiare.
9. A. Il sindacato chiede maggiore compartecipazione.
B. Sindacato chiede maggiore compartecipazione.

10. A. Futuro richiede investimenti.
B. Il futuro richiede investimenti.
- IV IntObjDef.
1. A. Moratti rassicura calcio italiano.
B. Moratti rassicura il calcio italiano.
 2. A. Ciampi esorta il governo.
B. Ciampi esorta governo.
 3. A. Violante spiazza il parlamento.
B. Violante spiazza parlamento.
 4. A. Berlusconi vuole ridurre debito pubblico.
B. Berlusconi vuole ridurre il debito pubblico.
 5. A. Alitalia riapre consiglio direttivo.
B. Alitalia riapre il consiglio direttivo.
 6. A. Del Piero cerca la partita perfetta.
B. Del Piero cerca partita perfetta.
 7. A. Bush teme la sindrome Aznar.
B. Bush teme sindrome Aznar.
 8. A. Previti corrompeva finanza.
B. Previti corrompeva la finanza.
 9. A. Israele accusa Francia di antisemitismo.
B. Israele accusa la Francia di antisemitismo.
 10. A. 'Presto Space' salva la video-storia italiana.
B. 'Presto Space' salva video-storia italiana.
- V. IniSubIndef.
1. A. Un giornalista attacca Biondi.
B. Giornalista attacca Biondi.
 2. A. Un buttafuori uccide due giovani.
B. Buttafuori uccide due giovani.
 3. A. Pitbull ferisce due bambini.
B. Un pitbull ferisce due bambini.
 4. A. Camorrista dirige due ristoranti famosi.
B. Un camorrista dirige due ristoranti famosi.
 5. A. Un carpentiere falsifica banconote.
B. Carpentiere falsifica banconote.
 6. A. Una signora anziana denuncia falsari.
B. Signora anziana denuncia falsari.
 7. A. Ditta italiana compra 25 supermercati in Cina.
B. Una ditta italiana compra 25 supermercati in Cina.
 8. A. Valigia terrorizza Madrid.
B. Una valigia terrorizza Madrid.
 9. A. Una missione italiana arriva a Bagdad.
B. Missione italiana arriva a Bagdad.
 10. A. Una sportiva italiana vince in Grecia.

B. Sportiva italiana vince in Grecia.

V. IntObjIndef.

1. A. Sei stati europei richiedono riduzione fiscale.
B. Sei stati europei richiedono una riduzione fiscale.
2. A. Sharon subisce sconfitta memorabile.
B. Sharon subisce una sconfitta memorabile.
3. A. Schumacher investe una persona anziana.
B. Schumacher investe persona anziana.
4. A. Si scola una bottiglia di grappa e fugge.
B. Si scola bottiglia di grappa e fugge.
5. A. Clooney compra villa a Como.
B. Clooney compra una villa a Como.
6. A. Spunta inventore da 50 mila euro.
B. Spunta un inventore da 50 mila euro.
7. A. Benigni vince un premio in America.
B. Benigni vince premio in America.
8. A. Cassano cerca un nuovo attaccante.
B. Cassano cerca nuovo attaccante.
9. A. Boselli strappa nuovo vertice.
B. Boselli strappa un nuovo vertice.
10. A. Fiat teme futuro nero.
B. Fiat teme un futuro nero.

VII + VIII Order of omission, definite articles:

1. A. La Bresaola conquista cucina olandese.
B. Bresaola conquista la cucina olandese.
2. A. La madre fa catturare figlia drogata.
B. Madre fa catturare la figlia drogata.
3. A. Camera firma la nuova Carta Europea.
B. La camera firma nuova Carta Europea.
4. A. Camera vara la riforma Moratti.
B. La camera vara riforma Moratti.
5. A. La par condicio cambia legge TV.
B. Par condicio cambia la legge TV.
6. A. Il professore smentisce magistrato.
B. Professore smentisce il magistrato.
7. A. Mago Zurli presenta il nuovo show per bambini.
B. Il mago Zurli presenta nuovo show per bambini.
8. A. Governo ratifica il condono edilizio.
B. Il governo ratifica condono edilizio.
9. A. Il consiglio riapre dibattito.
B. Consiglio riapre il dibattito.
10. A. Il parlamento boccia dialogo.
B. Parlamento boccia il dialogo.

VII + VIII Order of omission, indefinite articles:

1. A. Leone ferisce un bambino.
B. Un leone ferisce bambino.
2. A. Poliziotto uccide un killer.
B. Un poliziotto uccide killer.
3. A. Un ladro ferisce carabiniere.
B. Ladro ferisce un carabiniere.
4. A. Un treno investe ferroviere.
B. Treno investe un ferroviere.
5. A. Regista spagnolo riceve un premio.
B. Un regista spagnolo riceve premio.
6. A. Bomba distrugge una vetrina.
B. Una bomba distrugge vetrina.
7. A. Una bambina senza genitori cerca famiglia.
B. Bambina senza genitori cerca una famiglia.
8. A. Una dose eccessiva di droga uccide turista spagnola.
B. Dose eccessiva di droga uccide una turista spagnola.
9. A. Poliziotta tende una trappola.
B. Una poliziotta tende trappola.
10. A. Presentatrice TV offre una spilla in beneficenza.
B. Una presentatrice TV offre spilla in beneficenza.

IX Nouns in Isolation.

1. A. La prima sconfitta per Trapattoni.
B. Prima sconfitta per Trapattoni
2. A. La musica a volto coperto.
B. Musica a volto coperto.
3. A. Menengite a scuola.
B. La menengite a scuola.
4. A. Protesta radicale di fronte a Montecitorio.
B. La protesta radicale di fronte a Montecitorio.
5. A. La festa di primavera a Milano.
B. Festa di primavera a Milano.
6. A. Il maltempo in Friuli.
B. Maltempo in Friuli.
7. A. Pane a prezzo scontato.
B. Il pane a prezzo scontato
8. A. Terremoto in Marocco.
B. Il terremoto in Marocco
9. A. Il tempo di saldi.
B. Tempo di saldi.
10. A. Il compleanno di Ciampi.
B. Compleanno di Ciampi.

X Hanging Topic Constructions.

1. A. La polmonite ignota, primi casi in Italia.
B. Polmonite ignota, primi casi in Italia.
2. A. La Grecia, nuovo governo a tempo di record.
B. Grecia, nuovo governo a tempo di record.
3. A. Benzina record, vertice anti-rincari.
B. La benzina record, vertice anti-rincari.
4. A. Confindustria, scatta domani l'ora di Montezemolo.
B. La Confindustria, scatta domani l'ora di Montezemolo.
5. A. La lista Prodi, nomine già congelate.
B. Lista Prodi, nomine già congelate.
6. A. Il delitto d'Antona, denunciata brigatista.
B. Delitto d'Antona, denunciata brigatista
7. A. Governo battuto, vietato avere tre reti TV.
B. Il governo battuto, vietato avere tre reti TV.
8. A. Commercio mondiale, la carica di Cina e Brasile.
B. Il commercio mondiale, la carica di Cina e Brasile.
9. A. Il colesterolo, ora gli Usa processano quello 'buono'.
B. Colesterolo, ora gli Usa processano quello 'buono'.
10. A. Il casinò di Lugano, i giochi sporchi dei cambisti.
B. Casinò di Lugano, i giochi sporchi dei cambisti.

Test Headlines German.

I + II YES/NO AUX

1. A. Indische Großbank (ist) stark gefragt
B. Die indische Großbank (ist) stark gefragt
2. A. Die Martinstraße (wird) ab Freitag gesperrt
B. Martinstraße (wird) ab Freitag gesperrt
3. A. Verurteilung von Schleusern (wird) erschwert
B. Die Verurteilung von Schleusern (wird) erschwert
4. A. Der Mord an Angela B. (ist) aufgeklärt.
B. Mord an Angela B. (ist) aufgeklärt.
5. A. Neue Ledermesse in Düsseldorf (wird) geöffnet von DJ Bono
B. Die neue Ledermesse in Düsseldorf (wird) geöffnet von DJ Bono
6. A. "Hollywood-Schriftzug" (wird) versteigert
B. Der "Hollywood-Schriftzug" (wird) versteigert
7. A. Der Urteilspruch gegen Chodorkowskij (ist) verschoben
B. Urteilspruch gegen Chodorkowskij (ist) verschoben
8. A. Verurteilung von Schleusern (wird) vertagt
B. Die Verurteilung von Schleusern (wird) vertagt
9. A. Der Großraum-Airbus (ist) erfolgreich getestet
B. Großraum-Airbus (ist) erfolgreich getestet
10. A. Mord an Rudolph M. (ist) aufgeklärt
B. Der Mord an Rudolph M. (ist) aufgeklärt

III IniSubDef.

1. A. Das Kartellamt ermittelt gegen Fleurop
B. Kartellamt ermittelt gegen Fleurop
2. A. Weltwirtschaft wächst langsamer
B. Die Weltwirtschaft wächst langsamer
3. A. Der Handy-Markt wächst langsamer
B. Handy-Markt wächst langsamer
4. A. Post gibt weitere Filialen auf
B. Die Post gibt weitere Filialen auf
5. A. Der Finanzstreit eskaliert weiter
B. Finanzstreit eskaliert weiter
6. A. Die Koalition streitet über Kombilöhne
B. Koalition streitet über Kombilöhne
7. A. Konsumklima trübt sich
B. Das Konsumklima trübt sich
8. A. Der Sportbund schreibt wieder schwarze Zahlen
B. Sportbund schreibt wieder schwarze Zahlen
9. A. Ölpreis steigt kräftig
B. Der Ölpreis steigt kräftig
10. A. Die deutsche Handelsflotte wächst weiter
B. Deutsche Handelsflotte wächst weiter

IV IntObjDef.

1. A. EV Duisburg gewinnt Play-off-Finalspiel
B. EV Duisburg gewinnt das Play-off-Finalspiel
2. A. Hommelhoff erhält den Leo-Baek-Preis
B. Hommelhoff erhält Leo-Baek-Preis
3. A. Polen senkt Leitzins
B. Polen senkt den Leitzins
4. A. Trichet setzt die Zinswende durch
B. Trichet setzt Zinswende durch
5. A. Delphi verringert Streikgefahr
B. Delphi verringert die Streikgefahr
6. A. Bayern München gewinnt Fußball-Finalspiel
B. Bayern München gewinnt das Fußball-Finalspiel
7. A. Konjunkturdaten belasten den Dollar
B. Konjunkturdaten belasten Dollar.
8. A. Washington will bosnische Regierung stärken
B. Washington will die bosnische Regierung stärken
9. A. Pechstein verpasst den Weltcup-Sieg
B. Pechstein verpasst Weltcup-Sieg
10. A. Osthoff meidet Öffentlichkeit
B. Osthoff meidet die Öffentlichkeit

V IniSubIndef.

1. A. Ein Kartellskandal erfasst Breton
B. Kartellskandal erfasst Breton
2. A. Postmitarbeiter versteckt 8000 Briefe
B. Ein Postmitarbeiter versteckt 8000 Briefe
3. A. Ein Fan bedroht Dick Advocaat
B. Fan bedroht Dick Advocaat
4. A. Zweiter Vulkan rumort in Indonesien
B. Ein zweiter Vulkan rumort in Indonesien
5. A. Ein Taxi überfährt Touristinnen
B. Taxi überfährt Touristinnen
6. A. Ein Feuerball erschreckt dutzende Spanier
B. Feuerball erschreckt dutzende Spanier
7. A. Bullterrier verletzt Britney Spears
B. Ein Bullterrier verletzt Britney Spears
8. A. Ein Postbote versteckt 8000 Briefsendungen
B. Postbote versteckt 8000 Briefsendungen
9. A. Zeuge belastet Michael Jackson
B. Ein Zeuge belastet Michael Jackson
10. A. Ein Bahnstreik legt weite Teile Frankreichs lahm
B. Bahnstreik legt weite Teile Frankreichs lahm

VI IntObjIndef

1. A. Putin schlägt Nahost-Konferenz in Moskau vor
B. Putin schlägt eine Nahost-Konferenz in Moskau vor
2. A. Rupprath sucht einen neuen Trainer
B. Rupprath sucht neuen Trainer
3. A. Vucicevic droht lange Sperre
B. Vucicevic droht eine lange Sperre
4. A. Merkel beruft einen Energiegipfel ein
B. Merkel beruft Energiegipfel ein
5. A. Heidelberg Cement kauft dänischen Betonhersteller
B. Heidelberg Cement kauft einen dänischen Betonhersteller
6. A. Japan plant Verfassungsänderung
B. Japan plant eine Verfassungsänderung
7. A. Arcor kündigt einen Angriff auf Mobilfunk an
B. Arcor kündigt Angriff auf Mobilfunk an
8. A. Fujitsu erleidet herben Rückschlag
B. Fujitsu erleidet einen herben Rückschlag
9. A. Stoiber beruft einen neuen Regierungssprecher
B. Stoiber beruft neuen Regierungssprecher
10. A. Berlin sucht neue Ostpolitik
B. Berlin sucht eine neue Ostpolitik

VII + VIII Order of omission, definite articles:

1. A. Die Stahlindustrie drosselt Produktion
B. Stahlindustrie drosselt die Produktion
2. A. Fifa regiert die Fussball-Republik
B. Die Fifa regiert Fussball-Republik
3. A. Die deutsche Regierung lädt Papst ein
B. Deutsche Regierung lädt den Papst ein
4. A. Parlament in Kiew stützt die Regierung
B. Das Parlament in Kiew stützt Regierung
5. A. Die Telekom-Offerte lässt Fernsehbranche kalt
B. Telekom-Offerte lässt die Fernsehbranche kalt
6. A. Die Deutsche Bank verliert Heimspiel
B. Deutsche Bank verliert das Heimspiel
7. A. Handel verliert den Glauben ans Weihnachtsgeschäft
B. Der Handel verliert Glauben ans Weihnachtsgeschäft
8. A. Regierung lädt den Papst ein
B. Die Regierung lädt Papst ein
9. A. Der Finanzaufsichtsrat untersucht EDF-Börsengang
B. Finanzaufsichtsrat untersucht den EDF-Börsengang
10. A. Fiskus nimmt das Filmfonds ins Visier
B. Der Fiskus nimmt Filmfonds ins Visier

VII + VIII Order of omission, indefinite articles:

1. A. Brutaler Räuber verletzt eine Seniorin
B. Ein brutaler Räuber verletzt Seniorin
2. A. Ein Drogenabhängiger erdrosselte Taxifahrer
B. Drogenabhängiger erdrosselte einen Taxifahrer
3. A. Deutscher Millionär kauft eine britische Internetfirma
B. Ein deutscher Millionär kauft britische Internetfirma
4. A. Eine britische Zeitarbeitsfirma erprobt neues Marketingkonzept
B. Britische Zeitarbeitsfirma erprobt ein neues Marketingkonzept
5. A. Rentnerin fährt ein Auto zu Schrott
B. Eine Rentnerin fährt Auto zu Schrott
6. A. Minister-Sohn überfährt eine Rentnerin
B. Ein Minister-Sohn überfährt Rentnerin
7. A. Unbekannte Firma erhält einen Milliardenauftrag
B. Eine unbekannte Firma erhält Milliardenauftrag
8. A. Britische Elektronikfirma erwirbt ein Rekordergebnis
B. Eine britische Elektronikfirma erwirbt Rekordergebnis
9. A. Ein Polizist erschießt Bankräuber
B. Polizist erschießt einen Bankräuber
10. A. Fischvergiftung löst einen Großeinsatz aus
B. Eine Fischvergiftung löst Großeinsatz aus

IX Nouns in Isolation:

1. A. Jagd auf streunende Katzen
B. Die Jagd auf streunende Katzen
2. A. Der Sondereinsatz in Küche und Kinderzimmer
B. Sondereinsatz in Küche und Kinderzimmer
3. A. Wahl im Bundestag
B. Die Wahl im Bundestag
4. A. Der Rückkauf von Berlin-Fonds
B. Rückkauf von Berlin-Fonds
5. A. Prozess um den Mord am 'Banker Gottes'
B. Der Prozess um den Mord am 'Banker Gottes'
6. A. Verwirrende Reise eines Mädchens
B. Die verwirrende Reise eines Mädchens
7. A. Die erste Generalaudienz des neuen Papstes
B. Erste Generalaudienz des neuen Papstes
8. A. Anschlag auf Papst Johannes II
B. Der Anschlag auf Papst Johannes II
9. A. Die neue Koalition der kleinen Minderheit
B. Neue Koalition der kleinen Minderheit
10. A. Skandal um Thomas Gottschalk
B. Der Skandal um Thomas Gottschalk

Appendix B

Calculation Informative Load Italian Articles with ‘expletive’ as a separate meaning.

Table 1 *Calculation informative load individual articles Italian, with expletives as separate meaning.*

	Frequency (=F) (corpus De Mauro)	number of functions/ meanings (=R)	av.freq. per funct/m.= F/R	(F/R det)/ (sum F/R paradigm) (p)	I = -log ₂ p
il	7111	2	3555,5	0,145416249	-2,7817396
la	7254	2	3627	0,148340525	-2,7530153
le	2536	2	1268	0,05185988	-4,2692373
lo	425	2	212,5	0,008691029	-6,8462573
i	2637	2	1318,5	0,053925278	-4,2128945
l'	3120	2	1560	0,063802376	-3,970246
gli	820	2	410	0,016768573	-5,8980962
un	6624	1	6624	0,270914705	-1,8840894
un'	800	1	800	0,032719167	-4,9337202
una	3838	1	3838	0,156970205	-2,6714374
uno	227	1	227	0,009284064	-6,7510279
degli	111	1	111	0,004539784	-7,7831605
dei	394	1	394	0,01611419	-5,9555245
del	18	1	18	0,000736181	-10,407651
della	15	1	15	0,000613484	-10,670686
delle	470	1	470	0,019222511	-5,7010594
dello	1	1	1	4,0899E-05	-14,577576
dell'	1	1	1	4,0899E-05	-14,577576

Samenvatting in het Nederlands

In dit proefschrift onderzoek ik omissie van lidwoorden in het Nederlands en het Italiaans. Hierbij richt ik me niet alleen op de al in veel eerdere studies onderzochte omissie van lidwoorden door taalverwervende kinderen. Ik ga ook uitgebreid in op omissie van lidwoorden door een categorie van sprekers die tot nu toe in eerdere studies weinig aandacht heeft gekregen, namelijk volwassen sprekers in zogenaamde ‘speciale registers’, zoals krantenkoppen. Ik toon aan dat er interessante overeenkomsten bestaan tussen het patroon van omissie van lidwoorden door taalverwervende kinderen en volwassenen. We vinden dezelfde crosslinguïstische verschillen: meer omissies in het Nederlands dan in het Italiaans. We vinden dezelfde relatie tussen omissie van het lidwoord en de positie van het zelfstandig naamwoord in de zin: meer omissies als het zelfstandig naamwoord aan het begin van een zin staat. We vinden dezelfde relatie tussen omissie van het lidwoord en de finietheid van het werkwoord: meer omissies als een finiet werkwoord in de zin ontbreekt. Deze overeenkomsten zouden erop kunnen wijzen dat omissie van lidwoorden in deze twee zo verschillende categorieën van sprekers een gemeenschappelijke oorzaak hebben, en roepen dan ook de intrigerende vraag op wat deze gemeenschappelijke oorzaak is. Het beantwoorden van deze vraag is het doel van dit proefschrift.

In het tweede hoofdstuk belicht ik de psycholinguïstische achtergrond van dit proefschrift. Ik doe hierin een voorstel voor een model voor de productie van lidwoorden dat gebaseerd is op een combinatie van Levelt’s (1989) model voor taalproductie, een aantal voorstellen die zijn gedaan door Avrutin (1999, 2004a, 2004b) en het model voor lidwoordselectie zoals dat is voorgesteld door Caramazza en zijn collega’s op basis van hun bevindingen in verschillende crosslinguïstische studies (Alario and Caramazza, 2002; Caramazza et al., 2001; Schiller en Caramazza, 2002, 2003). Het door mij in dit hoofdstuk voorgestelde basismodel beschrijft het proces van lidwoordselectie onder normale en speciale omstandigheden, met normale sprekers en sprekers met een verminderde verwerkingscapaciteit van de hersenen (zoals bijvoorbeeld kinderen).

In het derde hoofdstuk presenteer ik de resultaten van mijn onderzoek naar omissie van lidwoorden door Nederlandse en Italiaanse

volwassenen in krantenkoppen en door taalverwervende kinderen. Mijn onderzoek naar de omissie van lidwoorden door volwassenen in krantenkoppen bestaat uit twee delen. In de eerste plaats heb ik een analyse gemaakt van de omissies van lidwoorden in een database die ik heb samengesteld aan de hand van krantenkoppen uit Nederlandse en Italiaanse kranten. Uit deze analyse blijkt dat lidwoorden vaker worden weggelaten in Nederlandse krantenkoppen dan in Italiaanse krantenkoppen. Bovendien toont de analyse aan dat omissie van het lidwoord beïnvloed wordt door de zinspositie van het zelfstandig naamwoord (aan het begin van de zin vinden we vaker omissies), door het al dan niet aanwezig zijn van een finiet werkwoord in de zin (we vinden vaker omissies als het finiet werkwoord afwezig is) en door het gebruik van een voorzetsel (we vinden minder omissies als het zelfstandig naamwoord direct volgt op een voorzetsel). Vervolgens heb ik een experiment uitgevoerd met Nederlandse en Italiaanse krantenlezers, waarin ik hun oordeel heb gevraagd over krantenkoppen waarin het lidwoord werd gebruikt of weggelaten, in de verschillende soorten linguïstische contexten die in de analyse van de database het gebruik van het lidwoord bleken te beïnvloeden. De resultaten bevestigden de conclusies van de analyse van de database, zowel voor wat betreft de crosslinguïstische verschillen (een hogere voorkeur voor omissie van het lidwoord in het Nederlands) als voor wat betreft de invloed van zinspositie en finietheid op de voorkeur voor omissie. Tenslotte bespreek ik in dit hoofdstuk de resultaten van de analyse van omissie van lidwoorden in spontane spraakdata van vier Nederlandse en vier Italiaanse taalverwervende kinderen, aan de hand van files afkomstig uit de Childes database. De resultaten van deze analyse tonen aan dat er interessante overeenkomsten zijn tussen omissie van lidwoorden door kinderen en volwassenen. We vinden dezelfde crosslinguïstische verschillen (meer omissies in het Nederlands), en we vinden dezelfde invloed van zinspositie en finietheid op het patroon van omissie.

Levelt beargumenteert dat het gebruik van het lidwoord niet berust op een bewuste beslissing van de spreker om een lidwoord te gebruiken; ook omissie van het lidwoord berust niet op een bewuste beslissing van een spreker om in een speciaal register het lidwoord weg te laten. Problemen in de productie van taal (bijvoorbeeld omissies van functionele categorieën) kunnen veroorzaakt worden door gebrekkige kennis van de taal of door gebrek aan de benodigde verwerkingscapaciteit om taal voort te brengen. Het is niet aannemelijk dat normale volwassenen, die onder normale omstandigheden het

lidwoord op juiste wijze kunnen gebruiken, in speciale registers, als krantenkoppen, opeens de kennis over het gebruik van het lidwoord verloren hebben. We kunnen dus aannemen dat de kennis van normale volwassenen over het gebruik van het lidwoord intact is. Dat ze toch lidwoorden weglaten in speciale registers zou dus veroorzaakt moeten worden door beperkingen in de verwerkingscapaciteit die benodigd is om lidwoorden te produceren. Als dit de juiste benadering is moeten twee vragen beantwoord worden:

- We vinden alleen omissies in speciale registers, dus kennelijk is daar de verwerkingscapaciteit beperkter dan in normaal taalgebruik: waarom is dat zo?
- We vinden meer omissies in speciale registers in het Nederlands dan in speciale registers in het Italiaans. We gaan ervan uit dat de verwerkingscapaciteit van de hersenen van Italiaanse sprekers niet verschilt van die van Nederlandse sprekers, dus dit zou impliceren dat de productie van een lidwoord in het Nederlands meer verwerkingscapaciteit vergt dan productie van een lidwoord in het Italiaans. Bestaat er een manier om verschillen in de verwerkingscapaciteit die nodig is om een lidwoord te produceren in verschillende talen op objectieve, kwantificeerbare wijze te meten?

En, als we het antwoord op deze vragen hebben gevonden, kunnen we dan hiermee ook omissie van lidwoorden in kindertaal verklaren?

In het model voor lidwoordselectie zoals ik dat in het tweede hoofdstuk heb voorgesteld, wordt, in navolging van de bevindingen van de experimentele crosslinguïstische studies van Caramazza en zijn collega's, beschreven dat tijdens het selectieproces van een lidwoord niet alleen het beoogde lidwoord, maar ook andere, concurrerende, lidwoorden geactiveerd worden. Het lidwoord dat uiteindelijk wordt geselecteerd is het lidwoord dat het sterkst is geactiveerd. In feite is de selectie van een bepaald lidwoord dus het resultaat van een competitieproces tussen alle lidwoorden binnen de groep van lidwoorden in een taal. Er is tijdens het selectieproces onzekerheid over de vraag welk lidwoord de competitie zal winnen, hoe sterker de competitie is tussen de verschillende lidwoorden in een groep, hoe groter de onzekerheid is over de uitkomst van het selectieproces. Hoe zwaarder de competitieproces tussen de lidwoorden, en dus: hoe groter de onzekerheid, des te meer verwerkingscapaciteit nodig zal zijn om het beoogde lidwoord te selecteren.

In hoofdstuk 4 toon ik aan dat het niveau van onzekerheid binnen een bepaalde groep van elementen niet een abstracte notie is, maar dat we de onzekerheidsgraad op een objectieve manier exact kunnen berekenen door gebruik te maken van een methode die is voorgesteld in Shannon en Weaver's (1949) Informatie Theorie. Deze theorie richtte zich in eerste instantie op technische datatransmissie over mechanische communicatiekanalen. Maar de resultaten van eerdere studies van Kostić die ik in dit hoofdstuk bespreek tonen aan dat ook de menselijke taalproductie, dus in feite de datatransmissie die plaatsvindt tijdens menselijke communicatie, gevoelig is voor de in Informatie Theorie gedefinieerde noties. Ik presenteer in hoofdstuk 4 een methode die ons in staat stelt de onzekerheidsgraad binnen de groep van lidwoorden van een taal op exacte wijze te bepalen. Deze onzekerheidsgraad wordt, binnen Informatie Theorie, relatieve entropie genoemd, en toont aan hoe groot de verschillen in waarschijnlijkheid van selectie tussen de elementen zijn. Als er grote verschillen zijn in de waarschijnlijkheidsverdeling van de elementen is de onzekerheid, relatieve entropie, laag, als de verschillen in waarschijnlijkheid van de elementen klein zijn is de relatieve entropie hoog. Hoe hoger de relatieve entropie, des te moeilijker zal het selectie proces zijn, en dus, des te meer verwerkingscapaciteit benodigd zal zijn voor selectie van het beoogde lidwoord. In hoofdstuk 4 toon ik aan de hand van de berekening van de relatieve entropie van de lidwoordenset in het Nederlands en het Italiaans aan dat deze hoger is in het Nederlands. Daarom is selectie van een lidwoord in het Nederlands een inspannender proces dan in het Italiaans. Zowel kinderen als schrijvers van krantenkoppen zijn gevoelig voor dit verschil in inspanning die benodigd is voor lidwoordselectie. Kinderen, omdat de verwerkingscapaciteit die ze ter beschikking hebben voor taalproductie nog niet het volwassen niveau heeft bereikt. Schrijvers van krantenkoppen omdat ze rekening houden met de beperkte tijd die lezers willen besteden aan het lezen van een krantenkop. Ook toon ik in hoofdstuk 4 aan hoe we met behulp van relatieve entropie de invloed kunnen verklaren die uitgaat van de positie van het zelfstandig naamwoord in de zin, van finietheid en van de aanwezigheid van een voorzetsel op het gebruik van een lidwoord.

In het laatste hoofdstuk bespreek ik de resultaten van het onderzoek dat ik heb gedaan naar lidwoordomissie door volwassen en kinderen in een derde taal: Duits. De resultaten van het onderzoek in deze derde taal tonen aan dat de crosslinguïstische verschillen die we vinden tussen omissie van lidwoorden in het Italiaans en het Nederlands niet

veroorzaakt worden door specifieke eigenschappen van Germaanse of Romaanse talen. De resultaten van mijn onderzoek in het Duits zijn namelijk niet alleen verschillend van de resultaten in het Italiaans, maar ook van het Nederlands. We vinden meer omissies in het Duits dan in het Italiaans, en minder omissies in het Duits dan in het Nederlands. Omdat, zoals ik in dit hoofdstuk aantoon, de relatieve entropie in het Duits hoger is dan in het Italiaans en lager dan in het Nederlands zijn dit ook precies de verschillen in omissie die we zouden verwachten. Dit betekent dat relatieve entropie een betrouwbare maatstaf voor de onzekerheidsgraad, en dus voor de benodigde verwerkingscapaciteit, is.

Deze studie toont aan dat zowel kinderen als volwassenen lidwoorden weglaten vanwege de hoge verwerkingskosten die selectie van lidwoorden uit de lidwoordenset met zich meebrengt. Kinderen hebben een beperkte verwerkingscapaciteit, krantenkoppenschrijvers moeten rekening houden met de beperkte tijd die lezers beschikbaar hebben voor het lezen van de koppen. Er bestaan geen verschillen tussen de ontwikkeling van de beschikbare verwerkingscapaciteit van de hersenen van Nederlandse, Duitse en Italiaanse kinderen, en er bestaan ook geen verschillen tussen de intenties van krantenkoppenschrijvers in de drie talen. Wat wel verschilt zijn de eigenschappen van de lidwoordensets in de drie talen. Deze verschillen dusdanig dat het produceren van een lidwoord in het Italiaans op vroegere leeftijd mogelijk wordt, en dat het weglaten van een lidwoord in het Italiaans minder voordelen biedt in de zoektocht van journalisten naar de best-mogelijke krantenkop.

Curriculum Vitae

Joke de Lange was born on the 13th of August, 1961, in Gouda (The Netherlands). She attended the Rijksscholengemeenschap in Gouda. From 1999 to 2003 she studied Italian Language and Culture at the University of Utrecht, in 2003 she obtained her MA degree with distinction (Cum Laude). In 2002 she spent a year at the Università degli Studi di Milano-Bicocca working on her MA-thesis and attending psycholinguistic courses.

Joke joined the Comparative Psycholinguistics Project in 2003 as an AIO (Assistent in Opleiding) and started her PhD research. This dissertation is the result of the research carried out during this time.