The Role of the Phonological Loop in Sentence Comprehension

Withaar, R.G.

Publication date:
2002

The Role of the Phonological Loop in Sentence Comprehension

Rienk Withaar

The research reported in this thesis was carried out under auspices of the School of Behavioral and Cognitive Neurosciences (BCN) and the University of Groningen

Printed by Print Partners Ipskamp, Enschede

Groningen Dissertations in Linguistics (GRODIL), # 40
ISSN 0928-0030

Rijks*universiteit* Groningen

The Role of the Phonological Loop in Sentence Comprehension

Proefschrift

ter verkrijging van het doctoraat in de
Letteren
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. D.F.J. Bosscher,
in het openbaar te verdedigen op
maandag 24 juni 2002
om 14:15 uur

door

Rienk Gerwin Withaar

geboren op 14 mei 1971
te Joure

Promotor:	Prof. dr. J. Koster
Co-promotor:	Dr. L.A. Stowe

Beoordelingscommissie:	Prof. dr. Y.R.M. Bastiaanse,
	Prof. dr. J. Nerbonne en
	Prof. dr. D. Swinney

Contents

Acknowledgments

Now that my PhD research project has come to a conclusion and I am about to publish my thesis, I would like to express my gratitude to a number of people without whom the project would not be what it is today.

First, I would like to thank Jan Koster and Laurie Stowe for supervising this research project and for keeping up the pressure at the end of the process. They encouraged me to go on and finish the job.

Second, I would like to thank the members of my reading committee. I would like to say thanks to David Swinney for commenting on my thesis and coming all the way to Groningen for my defense. And to John Nerbonne for the helpful suggestions for improvement of Chapter 6. Finally, I would like to thank Roelien Bastiaanse for her suggestions on the aphasia data and, of course, for showing us Mr. Tan's brains.

Especially, I would like to thank Laurie for the innumerable discussions we have had about the project and for her constructive comments and encouragement, not to mention for the force-feeding sessions at her place with people of the project, or any other lot for that matter.

Additionally, my defence would not have been possible without the support of André and Laura, who will be standing by my side on the 24th of June and who helped me out in the frantic days before the defence.

I would also like to thank my colleagues in the Neurological Basis of Language project, Christer, Hein, John, Laura, Laurie, Marco, Monika, and Monique for the fruitful discussions and suggestions about experiments, conference abstracts, and rudimentary versions of this thesis. Especially Hein's suggestion about changing the order of presentation of the articulatory suppression and no-load blocks and John's idea about the response bias turned out to be major break throughs in this research. And I would also like to thank my project members for spicing up conference visits, for showing the courage to celebrate Sinterklaas with teasing poems and all.

Further, I would like to thank my colleagues at the departments of Dutch and Linguistics, Bart, Claartje, Danielle, Dicky, Dieuwke, Dirk-Bart, Esther, Femke, Gerard, Jacques, Jan-Wouter, Joanneke, Judith, Maaike, Maartje, Roel, Roelien, Shalom, Sible, Stasinos, Tjeerd, Tony, and Victor for providing the much-needed distraction from the job. I really appreciated the coffee breaks, corridor chats, lunch conversations, and football matches and tournaments.

And I would also like to mention the staff at the administration, Alice, Anna, Belinda, Jolanda, Nathalie, Tineke, and Wyke in these acknowledgments. They have always been more than willing to help me out with the reproduction of yet another pretest.

Finally, I would like to thank my colleagues at the Faculty of Arts Library, Aly, Annet, Geert, Geke, Ger, Greetje, Ineke, Jean-Marc, Jelga, Lideke, Mary, Michiel, Mimi, and Tom, for having put up with my final dissertation stress.

Chapter 1
Introductory Remarks

Introduction

Every day, people spend a considerable amount of time understanding language. They are good at it: they seem to do it exceptionally fast and seemingly effortlessly. Because language comprehension plays such an important role in everyday life, it is important to know how people understand language. In order to say something about this issue, one must know which information language communicates and how people deal with each information type.

The first issue is investigated within the field of theoretical linguistics. Theoretical linguists look at language at several levels. At the sound level, phoneticians attempt to systematically classify the sounds made in the spoken utterances of a particular language, whereas phonologists study how sound is structured in a particular language. Of course, sounds can be combined to form words. How linguistic elements can be combined at the word level is looked at by morphologists, who study how words are formed through inflection, derivation, and compounding. Then at yet a higher level, syntacticians study how words are combined into phrases and phrases into clauses. Obviously, there is more to language than structure alone. Several branches of theoretical linguistics are concerned with the meaning of language utterances. Semantics is concerned with theories about meaning, including theories of denotation, extension, naming, and truth. Additionally, pragmatics is concerned with the relationship of sentences to the environment in which they occur.

As can be seen from this overview, language is quite complex, consisting of structure at multiple levels, and yet we seem to process it easily. How we manage to accomplish this is studied within psycholinguistics. It may not be very surprising that psycholinguists make much use of theoretical linguistics, since theoretical linguistics provides a detailed and well-researched framework of concepts which are clearly relevant to language processing. In addition, psycholinguists often refer to models of information processing that have been developed within cognitive psychology.

The research presented in this book falls within the domain of psycholinguistics. At the broadest level, it is concerned with the interaction between the different kinds of linguistic information during processing. Specifically, it was directed at the development of a model about the use of working memory (WM) during sentence comprehension. The concept of WM has been extensively studied within cognitive psychology. One of the most prominent models of WM was developed by Baddeley and coworkers (e.g., Baddeley, 1986; Baddeley & Hitch, 1974; Gathercole & Baddeley, 1993).

The experiments reported in Chapters 3 to 5 were conducted in order to specify the role of the *phonological loop* during sentence comprehension. The phonological loop is a module of Baddeley's WM model that deals with the temporary storage of phonological information. The model presented in this book concerns how and when readers make use of phonological memory representations during syntactic and semantic integration processes.

Assuming a Role for (Phonological) WM in Sentence Comprehension

Many researchers have explained sentence complexity effects in terms of WM limitations (e.g., Berwick & Weinberg, 1984; Caplan & Waters, 1999; Daneman & Carpenter, 1980; Gibson, 1998; Just & Carpenter, 1992; King & Just, 1991; Marcus, 1980; Martin & Romani, 1994; Stowe, 1986, 1991; Waters, Caplan, & Hildebrandt, 1987). The assumption of a role for WM in sentence comprehension is intuitively plausible because language comprehension occurs over time. For auditory comprehension, this is because a speaker cannot utter two words simultaneously. Therefore, the listener has to store the words just spoken until they can be chunked (e.g., by combining them to a phrase structure).

Although readers are able to view more than one word at a time, visual language processing happens serially too. Only a limited number of words can be inspected and processed at the same time. This serial aspect of both auditory and visual language comprehension calls for the involvement of WM. If words that have not yet been integrated into an already existing (phrase) structure are not maintained in one way or another, sentence comprehension simply cannot succeed.

Defining the Role of WM in Sentence Comprehension

It is certainly true that language comprehension occurs over time, but it happens in a very quick and automatized way. Therefore, one may object that the use of WM during sentence comprehension is minimal or even nonexistent. One may argue that the major challenge of a reader/listener is to immediately combine words into phrases, phrases into clauses etc. In these aspects of sentence comprehension, processing is very prominent, but memory aspects hardly come into play.

However, within the WM literature, it is generally assumed that WM entails both the manipulation of information and the temporary storage of intermediate products that result from these manipulations (Baddeley, 1986; Daneman & Carpenter, 1980; Just & Carpenter, 1992; Salthouse, 1990, 1994). Under this view, the processing aspects of sentence comprehension just mentioned are also part of WM. In this book, WM processes are taken to entail both the storage and processing aspects of language comprehension.

Position of the Current Research

Many psycholinguistic and computational models that assume a role for WM in sentence processing specify what must be stored at a particular moment (e.g., Berwick & Weinberg, 1984; Gibson, 1998; Marcus, 1980). However, these models are not always clear about in what format or manner the elements to be stored are maintained.

The latter aspect received more attention within the WM literature (e.g., Babcock & Salthouse, 1990; Baddeley, 1986; Gathercole & Baddeley, 1993; Salthouse, 1990, 1994). One model that is particularly well researched and generally accepted is Baddeley's model of WM. This model is based on a large body of experimental evidence. However, the language experiments usually involve storage of lists of words and syllables rather than (parts of) sentences. Therefore, Baddeley's model has clear limitations when applied to sentence processing.

Several researchers have developed linguistic processing theories that are based on Baddeley's (1986) WM model (e.g., Baddeley, Vallar, & Wilson, 1987; Caplan & Waters, 1999; Just & Carpenter, 1992; Martin & Romani, 1994; Vos, 1999; Waters et al., 1987). These models specify both the comprehension processes and the way information is processed and/or stored.

The current research fits in with the latter approaches. The model presented in this book is a specification of Baddeley's WM model along the lines of Baddeley et al. (1987) and of Caplan & Waters (1999).

A Limited Role for Phonological WM in Sentence Comprehension?

Several researchers have discussed the role of phonological WM representations in sentence comprehension. This role differs from model to model. Four important hypotheses about the role of phonological memory in language processing will be discussed below.

One of the first studies that proposes a role for the phonological loop in sentence comprehension was conducted by Caramazza, Basili, Koller, & Berndt (1981). They proposed that the phonological loop is used when sentence interpretation requires full syntactic analysis of a sentence. This is the case with reversible sentences, such as (1a); nonreversible sentences, such as (1b), do not require full syntactic analysis, because the lexical content constrains the relations among the words in the sentence:

(1) a. The guy is courting the girl
b. The guy is buying a beer

Caramazza et al. (1981) made this proposal on the basis of the performance of a conduction aphasic, M.C., who had a severe limitation in auditory-verbal WM. He was poor at repeating word lists and failed completely on lists of function words. Additionally, his comprehension of reversible sentences was much worse than that of nonreversible sentences. Caramazza et al. concluded that function words were stored in a phonological format and that for interpretation of these function words, a whole representation of the sentence was needed. Given the severe phonological WM limitations observed in M.C., M.C. could not retain such a representation.

Baddeley et al. (1987) made a similar proposal about the role for phonological WM representations in sentence comprehension. They reported a patient with a pure auditory-verbal WM deficit who had problems judging the truth value of sentences as in (2):

(2) The earth divides the equator into two hemispheres, the northern and the southern

Based on this pattern of performance, they claimed that phonological WM representations are used only in those cases where word order is crucial for sentence comprehension. This proposal is similar to the one made by Caramazza et al. (1981), because in the reversible sentences tested by Caramazza et al., the use of word order as a cue for sentence interpretation is crucial as well.

Waters et al. (1987) made a different proposal about the role of phonological WM representations. They argued that these representations are addressed during the checking of propositional content. They tested this checking procedure by comparing sentences with one action with two-action sentences as shown in (3a and b):

(3) a. It was the thief that broke into the warehouse
b. The man hit the landlord that requested the money

They found that blocking access to the phonological representation by use of articulatory suppression made comprehension more difficult, particularly for two-proposition sentences.

Given the fact that the meaning of sentences is usually constrained by context, Waters et al. (1987) assumed that this checking procedure is hardly ever used. They argued that phonological-loop involvement in sentence comprehension can be observed in sentences presented in isolation (i.e., in cases where there is little or no constraining influence from context).

In contrast to the three studies just reported, Martin & Romani (1994) assumed that semantic WM representations rather than phonological representations are of critical importance for sentence comprehension, but phonological WM representations are not. They found that healthy volunteers had more difficulty judging sentences with stacked adjectives in cases where the adjectives preceded the noun they modified than in cases where the adjectives followed the noun they modified as in (4a and b):[1]

(4) a. # A fluffy, small, surprised shriek came out of the room
b. # A shriek that was fluffy, small, and surprised came out of the room

Martin & Romani (1994) explained this difference between the processing of pre- and post-modifying adjectives in terms of semantic WM load. In the premodification conditions, adjectives must be stored until the head noun is encountered but in the postmodification conditions, the adjectives can be integrated immediately.

On the other hand, Martin & Romani (1994) considered the contribution of phonological WM systems less crucial. This was because E.A., a patient with a severe impairment of phonological WM who also performed the experiment just discussed, showed the same pattern of performance as the healthy volunteers.

---

[1] The pound sign (#) is used in this book for sentences that are syntactically correct, but semantically unacceptable.

Reconsidering the Role of Phonological WM in Sentence Comprehension

The proposals of the role of the phonological loop in sentence comprehension presented above differ with respect to whether they assume a role for phonological memory in sentence comprehension and with respect to how large that role is. Two factors may have influenced these proposals and be responsible for the limited occurrence of phonological WM effects in the studies just reported.

First, the studies reporting little involvement of phonological memory in sentence comprehension mainly looked at English data. In English, word order is critical for syntactic analysis (e.g., King & Just, 1991; MacDonald & Christiansen, in press). This type of information may be maintained in other forms of WM than the phonological loop.

Second, the specific model proposed by Waters et al. (1987) is based on potentially confounded data. We will return to this issue on page 28.

It may well be that the use of phonological WM representations becomes more apparent when syntactic analysis depends on phonologically subtle cues. This is the case in Dutch relative clauses, where subject-verb agreement is an important syntactic cue (Lamers, 2001; Mak, 2001). In Dutch, subject-verb agreement is expressed by phonologically nonsalient morphosyntactic elements. This book presents five whole-sentence anomaly-judgment experiments that investigate the role of phonological WM representations in sentence comprehension using Dutch materials.

A Quick Walk-Through

The model presented in this book was developed on the basis of five whole-sentence anomaly-judgment experiments. In the current section, I will briefly sketch the contributions of each experiment to certain features of the model.

The Basis for the Current Research

A whole-sentence anomaly-judgment task carried out by Waters et al. (1987) served as the basis for the current research. Waters et al. assumed a partition of WM resources that are used during sentence comprehension. They assumed separate stages for syntactic and postsyntactic (or propositional) processing. In the syntactic stage, phrase structure is assembled, and in the postsyntactic stage, sentence level integrations are carried out that are semantic in nature. During the latter stage, a phonological back-up representation of the sentence that is being analyzed is consulted.

This latter assumption seemed theoretically underspecified and based on potentially confounded data: Waters et al. (1987) used center-embedded structures in the postsyntactically complex, but not in the simple sentences.

Replication and Revision

Therefore, a replication of their experiment was carried out using Dutch materials (Experiment 1 on p. 42). Experiment 1 differed from the Waters et al. (1987) experiment in that it used right-embedded structures only and it contained reversible instead of nonreversible sentences. Experiment 1 yielded results that are consistent with a partition of processing into separate syntactic and postsyntactic stages along the line of Waters et al. However, there was no evidence for a role for a phonological back-up in postsyntactic processing since the effects of articulatory suppression were equally detrimental for both one- and two-proposition languages.

Since Experiment 1 did not use center-embedded structures, it appeared that the phonological back-up is instead involved in the processing of center-embedded structures. This suggestion was corroborated by the results of Experiment 2 (cf. p. 56), which found interactions of center embedding with phonological memory load.

Furthermore, the manipulation of Clause Type (subject- vs. object-relative clause), which is taken to affect the syntactic processing stage, interacted with phonological memory load in Experiment 2 as well. Dutch subject-relative clauses, in which the subject and verb are further apart than in object-relative clauses, were read less accurately during irrelevant articulation.

Additionally, it turned out that participants were faster in phonological memory load conditions (irrelevant articulation) than in no-load condition. However, these faster RTs were associated with a detrimental effect on accuracy on implausible sentences.

On the basis of these findings, it was concluded that Waters et al.'s (1987) proposal needs revision. It was proposed that a phonological back-up is addressed during the checking of subject-verb agreement features across a distance as in subject relatives and center-embedded structures and that readers of Dutch verb final constructions go faster under irrelevant articulation in order to prevent the back-up from decaying too badly to carry out this check.[2]

---

[2] This proposal is made under the assumption that articulatory suppression exerts a detrimental influence on phonological memory representations (Baddeley, 1986; Gathercole & Baddeley, 1993). As a consequence, these representations are taken to decay more rapidly.

The Influence of Reversibility

A potential objection against the approach taken in Experiments 1 and 2 is that reversible sentences were used instead of nonreversible sentences. The difference between reversible and nonreversible sentences is shown in (1a and b) repeated here as (5a and b):

(5) a. The guy is courting the girl
b. The guy is buying a beer

Because of this difference in materials, two similar experiments were conducted that used nonreversible sentences (Experiments 3 and 4; cf. Chapter 4). These experiments corroborated the conclusions from Experiments 1 and 2 with respect to the use of the phonological loop by the syntactic processor, the postsyntactic analyzer, and the phonological back-up. However, the differences in materials produced some differences in strategy.

Investigating the Checking Hypothesis

On the basis of Experiments 1 to 4, one may ask whether the phonological back-up is also consulted during integrations other than subject-verb agreement. Therefore, Experiments 5a and b looked at the long-distance integration of adjectives and nouns in Dutch.[3] It turned out that no interactions between phonological memory load and long-distance adjective-noun integration could be found. Therefore, it was concluded that the use of the phonological back-up is restricted to subject-verb integration only.

A Reader's Guide to This Book

This book consists of seven chapters. Chapter 1 serves as a general introduction and as a reader's guide to the book. Chapter 2 gives an overview of the background literature about the role of WM and of phonological WM representations in sentence comprehension and introduces the main research issues addressed in this book. Chapters 3 to 5 report five whole-sentence anomaly-judgment experiments that served as a basis for a model of WM and sentence comprehension. Chapter 6 discusses the memory span data from the five experiments reported in the Chapters 3 to 5. Finally, the model itself is presented in Chapter 7.

---

[3] In Dutch, there is an agreement relation between nouns and adjectives.

Chapter 2 offers an introduction to the field for those who are not familiar with the sentence processing and/or WM literature. Readers already familiar with this literature (e.g., Just & Carpenter, 1992; Caplan & Waters, 1999) and with Baddeley's (1986) WM model, may want to proceed directly to Chapters 3 to 5 (cf. p. 32).

Chapters 3 to 5 report the experiments on which the WM model presented in this book was based. Chapter 3 reports two follow-ups to an experiment conducted by Waters et al. (1987, exp. 3) that used reversible sentences, whereas the Waters et al. experiment used nonreversible sentences. Chapter 4 reports two follow-ups to Waters et al. that used nonreversible sentences like the Waters et al. study. Chapter 5 looked at the WM resources used during the integration of stacked adjectives. The Chapters 3 to 5 can be read without having read the other chapters of the book.

Chapter 6 is about the predictive power of two different WM span tasks, a nonword span test and a Dutch version of the Salthouse Listening Span Test. Because correlations with complexity effects reported in Chapters 3 to 5 are discussed in Chapter 6, knowledge of the designs of the experiments reported in Chapters 3 to 5 is necessary in order to understand Chapter 6.

Chapter 7 compares the experiments presented in Chapters 3 to 5 with each other and presents a new model of WM and sentence processing. Because it refers back to the data in Chapters 3 to 5, it may be impossible to understand Chapter 7 without knowing the designs and results presented in the latter chapters.

Chapter 2
Background and Goals

Introduction

This book presents a model of the use of working memory (WM) resources during sentence comprehension. Specific attention will be given to the role of phonological WM representations during the process of comprehension. These phonological representations are assumed to be kept in the phonological-loop component of WM (Baddeley, 1986; Baddeley, Vallar, & Wilson, 1987; Gathercole & Baddeley, 1993; Waters, Caplan, & Hildebrandt, 1987).

The current chapter will begin with a sketch of sentence comprehension models that explain linguistic complexity effects in terms of WM demands. This description is followed by a side path into Sternberg's (1969) logic of interactions. Sternberg's interaction logic plays an important role in Baddeley's (1986) WM model, which will be discussed next. After the sketch of Baddeley's (1986) WM model, several models will be described that applied Baddeley's model to sentence comprehension. Finally, the goals of the research reported in this book will be outlined and motivated.

With respect to the overview of literature, it is important to keep in mind that in order to specify the role of phonological WM representations in sentence processing, one must characterize other WM resources as well. Therefore, the overview of the literature will focus on the use of WM resources in general.

The Sentence Comprehension Literature

In the current section, I will give an overview of models that have explained linguistic complexity effects in terms of WM demands. The first subsection covers one of the first parsing models with a WM component, which was developed by

Wanner & Maratsos (1978).[1] The next subsection deals with the syntax-first approach to parsing. This subsection is followed by a subsection on the Berwick & Weinberg (1984) parser, and finally the syntactic prediction locality theory (Gibson, 1998) will be addressed.

An important body of research directed at the use of WM resources during sentence comprehension uses Baddeley's (1986) WM model as a starting point (e.g., Caplan & Waters, 1992; Just & Carpenter, 1992; Martin & Romani, 1994). These studies will not be discussed in the current section. In order to understand these models, one must be familiar with Baddeley's model of WM; therefore, Baddeley's WM model will be discussed before the coverage of the models based on his WM model. Baddeley's model will be discussed on page 20, and the approaches based on his model will be covered on page 24.

An Early WM Approach to Sentence Comprehension

An early model of sentence comprehension with a WM component was developed by Wanner & Maratsos (1978). Wanner & Maratsos's proposal modeled the processing of subject- and object-relative clauses in English, as in (1a and b) respectively:

(1) a. The witch who despised sorcerers frightened little children
  b. The witch who sorcerers despised frightened little children

Wanner & Maratsos (1978) used an augmented transition network (ATN) to model the syntactic aspects of relative-clause comprehension. The network goes through a series of states. Each state has particular actions associated with it. Each action will result in a next state. The actions of the network are integration actions. Part of this network is the so-called HOLD list, a memory store.

The network combines the words of a sentence into a syntactic representation. When the network encounters a relative pronoun, such as *who,* it stores the pronoun's antecedent in the HOLD list. The sentence analyzer of the network processes a clause in a number of steps: first, it looks for a subject noun phrase (NP), then for a verb to go with it, and then for objects.

In the subject-relative clause in (1a), it will find that the relative pronoun is not followed by an NP, but by a verb. Given the fact that there is a verb, but no subject,

---

[1] Sentence comprehension, more specifically syntactic processing, is often referred to as *parsing*. In this book, *syntactic processing* and *parsing* will be used interchangeably i.e., as synonyms.

the ATN checks the HOLD list for a potential subject and inserts the words in the HOLD list, *the* and *witch*, in the subject position.

In the object-relative clause, the ATN does find a subject and must continue processing the clause before it encounters the position where the content of the HOLD list can be inserted (the direct-object position). Thus, the model accounts for the fact that object-relative clauses are more difficult to read than subject-relative clauses, because the content of the HOLD list must be stored longer in object-relative clauses.

Wanner & Maratsos (1978) collected experimental evidence for their model in a design where sentences were presented one word at a time. Presentation of each sentence was interrupted with the presentation of five names; after presentation of the names, the sentence continued. At the end of each sentence, participants were asked either to answer a comprehension question on the sentence or to recall the five names presented during the interruption. There were four interruption points in the experimental sentences; these are indicated in (2) below:

```
           1                2              3          4
           ↓                ↓              ↓          ↓
(2) a.  The witch who despised sorcerers frightened little children
    b.  The witch who sorcerers despised frightened little children
```

The second interruption point was in the critical region; the other points served as controls.

Wanner & Maratsos (1978) found that participants had higher error proportions for object- than for subject- relative clauses in both the recall and the comprehension task when the interruption occurred at position 2, but not at the other positions. In terms of their model, participants performed worst when storage of the names and addressing the HOLD list occurred at the same time.

Since Wanner & Maratsos (1978), the subject- vs. object-relative comparison is typically used in investigations of WM load in syntactic processing, a tradition that will be continued in this thesis.

About the Syntax-First Approach and WM Limitations

A second line of research that makes appeals to WM limitations of the parser is the syntax-first approach (e.g., Frazier, 1979, 1987). This approach states that syntactic information is used before other information, such as semantic, and pragmatic information, comes into play. The syntactic processor is designed so that it consumes as few WM resources as possible. It does so by following the minimal attachment strategy and the active filler strategy.

The Minimal Attachment Strategy. The first strategy of the syntactic processor is adhering to the principle of minimal attachment. This principle states that when multiple syntactic analyses are permitted at a particular point, the processor initially opts for the simplest possibility.

Consider sentence (3) for example:

(3) The officers taught at the military academy were very demanding

The words *taught at the military academy* are temporarily ambiguous. This clause can serve as a main clause or as a reduced relative clause. The relative-clause reading is structurally more complex because it requires the construction of both a main clause and an embedded clause, whereas for the main-clause reading, the construction of a main clause suffices.

The minimal attachment strategy states that readers will initially interpret this sentence part as a main clause, because that is least demanding to WM. When it turns out that this choice is incorrect, the parser has to go back and re-analyze the sentence. This is the case when readers encounter the word *were*, which is the main-clause verb.

The Active Filler Strategy. A second parsing strategy assumed by the syntax-first approach is the active filler strategy. This strategy also aims to minimize WM demands. It states that when readers encounter a filler, such as the interrogative pronoun who in (4), they will try to identify its grammatical function as soon as possible:

(4) Who does the teacher talk to the students about?

The first possible location for an empty position to occur in (4) is at subject position, but this position is occupied by *the teacher.* The second possible location is after the preposition *to.* However, this position is already occupied by the NP *students.* Therefore, reading the words *teacher* and *students* should yield structural re-analyses and lead to longer maintenance of the filler. Like Wanner & Maratsos's (1978) proposal, this strategy minimizes the time during which the filler is maintained.

Challenges to the Syntax-First Approach. The assumption that syntactic processing occurs before other information, such as lexical content, can be evaluated has not gone unchallenged. Kaan (1997) found that participants have no difficulty processing wh-clauses when the subjects of these clauses are personal pronouns, as shown in (5):

(5) Wij wisten welke vrienden hem in het dorp aan de kust hadden opgezocht
We knew which friends him in the village at the shore had visited
'We knew which friends had visited him in the village at the shore'

This indicates that readers make early use of the case information expressed by that pronoun and do not attempt to get rid of the filler as soon as possible in these constructions.

Furthermore, Mak (2001) showed that object-relative clauses are just as complex as subject-relative clauses when the object is inanimate, suggesting that animacy information influences syntactic parsing decisions at a very early stage.

Third, Stowe (1986) proposed a slight modification of the active filler strategy. She conducted an experiment that looked at the following sentence types:

(6) a. My brother wanted to know if Ruth will bring us home to Mom at Christmas
b. My brother wanted to know who_i t_i will bring us home to Mom at Christmas
c. My brother wanted to know who_i Ruth will bring t_i home to Mom at Christmas
d. My brother wanted to know who_i Ruth will bring us home to t_i at Christmas

Sentence (6a) served as the base line, since no empty position had to be identified in that sentence. In (6b), the empty position is in the subject position, in (6c) in object position, and in (6d) it is in the complement position of a prepositional phrase.[2]

Under the active filler strategy, one would expect the sentences (6a and b) to be easiest. No empty position must be detected in (6a), and the empty position is found as soon as possible in (6b).

Furthermore, the syntax-first approach predicts a structural re-analysis at the subject position in (6c and d), because an empty position was expected by the parser. The presence of an overt NP forces the parser to re-analyze the sentence.

However, such an increase was not found, suggesting that participants expected the empty category to be in direct-object position instead of in the first possible position as predicted by the active filler strategy. Alternatively, it may indicate that readers have no trouble recovering from the incorrect assumption of an empty category in subject position.

The minimal attachment strategy has been challenged as well. Stowe (1991) found that structure is assigned immediately if the preference for the simplest syntactic construction matches with the semantic content of the sentence to be understood. If this is not the case, the parsing decision is delayed. This decision is not delayed indefinitely, but restricted by limitations of processing resources.

---

[2] The empty positions in the sentences are indicated with a *t*. The fact that the empty position is the underlying position of the filler, *who*, is indicated with a subscript *i*.

Additionally, Stowe (1991) proposed that the parser does not decide for a structure on the basis of its simplicity, but on the basis of the amount of structure that is needed in order to complete the sentence. This proposal is similar to Gibson's (1998) assumption that the fewest number of predictions needed to complete a sentence grammatically determines the processing load of that sentence.

The syntax-first approach and Stowe's (1986, 1991) proposals differ with respect to the specification of the strategies readers use during sentence comprehension. However, the two approaches are similar in that they both assume that parsing strategies are determined by limitations in WM capacity. This does not go for constraint-based approaches that radically oppose to the syntax first-approaches.

Constraint-Based Approaches. The constraint based approaches, assume that a number of factors guide the parsing process (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; Tanenhaus & Trueswell, 1995). These are factors such as the frequency of occurrence of a particular construction, thematic fit etc. Since these approaches assume that many different information types are used during sentence comprehension and that alternatives are maintained until full resolution occurs, they do not assume a limited capacity WM in the way the syntax-first approach does.

The Berwick & Weinberg Parsing Model

Berwick & Weinberg (1984) proposed a model of sentence processing, which is inspired by Chomsky's (1965) derivational theory of complexity (DTC). The DTC assumes a direct relationship between the grammar and the way that grammar is used for comprehension. According to this theory, every grammatical operation has a unit time cost; therefore, the DTC predicts that sentences with many grammatical operations are processed more slowly than sentences with few operations. The Berwick & Weinberg model is an extension of the DTC which is more specific about how constituents are processed than the DTC.

There are two assumptions that underlie the Berwick & Weinberg (1984) parsing model. The first assumption about the model is that it takes already segmented words as its input. Second, the model is taken to access the lexical representations of the words it processes. This information entails grammatical category and subcategorization information.

Subcategorization properties of a word determine which constituents go with a particular word and which do not. For example, the subcategorization information of the word *frighten* states that the verb requires a direct object. This is illustrated in (7a and b), where one can see that a sentence with the word *frighten,* but without a direct object is ungrammatical:[3]

(7) a. The villain frightens the old lady
b. * The villain frightens

Berwick & Weinberg (1984) made two proposals about the parser's architecture. First, they proposed that the parsing mechanism is deterministic without a backtracking mechanism. The parser is deterministic in the sense that it analyzes sentences from left to right. The fact that the parser lacks a backtracking mechanism implies that if the parser fails to produce the right structure, it must re-analyze the sentence and start all over again.

Berwick & Weinberg's (1984) second proposal about the parsing mechanism is that it has an input buffer with three cells. It takes the information in the first cell as input, but it is allowed to look ahead at the information of the next two cells. Berwick & Weinberg assumed a deterministic parser with limited look-ahead because:

1. They want to mimic real-time parsing;
2. with limited look-ahead, the parser can resolve some local ambiguities; and
3. with the limited look-ahead and deterministic nature of the parser, Berwick & Weinberg want to mimic the limited amount of WM resources of humans

The Berwick & Weinberg's (1984) parser consists of a syntactic processor and a propositional list. In the syntactic processor, words come in an input buffer. The syntactic processor combines these words into a phrase, which is kept active in an *active node stack.* Within this stack, the phrase is extended as long as the incoming words can still be attached. When a new word cannot be integrated, the phrase is shunted to a pushdown stack; the new word is now stored in the active node stack. Phrases stored in the different stacks are combined when they can form a clause.

Then the clause is transferred to a propositional list, which carries out the second stage of the process. Berwick & Weinberg (1984) assumed two processing stages in order to account for the distributional properties of reflexives and pronouns. In their model, antecedents of reflexives are detected in the syntactic stage, but the antecedents of pronouns are identified in the propositional list.

The distribution of pronouns and reflexives is illustrated below: in sentence (8a), the reflexive *himself* refers to an element within the same clause, *who.* In (8b), on the

---

[3] In this book, ungrammatical example sentences are preceded by an asterisk (*).

other hand, *him* cannot have an antecedent within the same clause: its antecedent is *John*, the subject of the main clause:[4]

(8) a. John talks to the man_i who_i is shaving $himself_i$
b. $John_j$ talks to the man_i who_i is shaving him_j

The Syntactic Prediction Locality Theory

Yet another model of WM and sentence processing was developed by Gibson (1998). In his syntactic prediction locality theory, he proposed that humans have a processing system of limited capacity. In this system, storage of linguistic information and integration of words compete for resources.

Gibson's (1998) idea was that memory load during processing is determined by the number of active predictions that are necessary to complete a sentence in its simplest form. If a reader encounters the first word, *the,* (s)he must make at least two predictions: one for a subject noun and one for a verb because grammatical English sentences need both. Consider sentence (9) for example:

(9) The man that walks the dog greets the woman

As soon as a prediction is fulfilled, it is not maintained any longer. This means that the prediction of a subject noun is dropped once *man* is read, and that the prediction of the main verb is dropped later, at *greets*.

Gibson (1998) assumes that the demands of a particular integration depend on the number of new discourse referents over which the integration takes place. Verbs and nouns, and in some cases adjectives as well, are considered new discourse referents. Thus, the integration of the words *the* and *man* in (9) costs one integration unit because *man* is the only active discourse referent. However, the integration of *man* and *greets* costs four integration units because *man*, *walks*, *dog*, and *greet* must be active during this integration.

The amount of resources available for integration processes equals the total amount of WM resources available minus the amount of resources needed for maintenance of open predictions (i.e., the memory load).

[4] In the example sentences in this book, words with the same referent are coindexed with the same subscript letter, in these examples, the letters *i* and *j*.

Methodological Interlude

After the overview of sentence processing approaches that assume involvement of WM during comprehension, we shift our perspective to the WM model developed by Baddeley (1986). In order to understand this model correctly, one must first be familiar with two techniques that are used to test whether certain cognitive tasks are processed within the same module or not. Therefore, it is important to dwell on the logic of interactions and on the logic of double dissociations before discussing Baddeley's WM model in detail.

Sternberg's Interaction Logic

Sternberg's Interaction Logic in a Nutshell. Sternberg (1969) argued that if two factors show additive effects, they draw on different cognitive resources; if on the other hand, they produce a superadditive effect (i.e., they interact), they share a common cognitive resource pool.

Whether effects interact or not can best be tested using an orthogonal design. Suppose one would like to conduct an experiment with two factors, A and B that both have two levels of complexity, simple and complex. An orthogonal design would yield the following experimental conditions: A simple/B simple, A simple/B complex, A complex/B simple, and A complex/B complex.

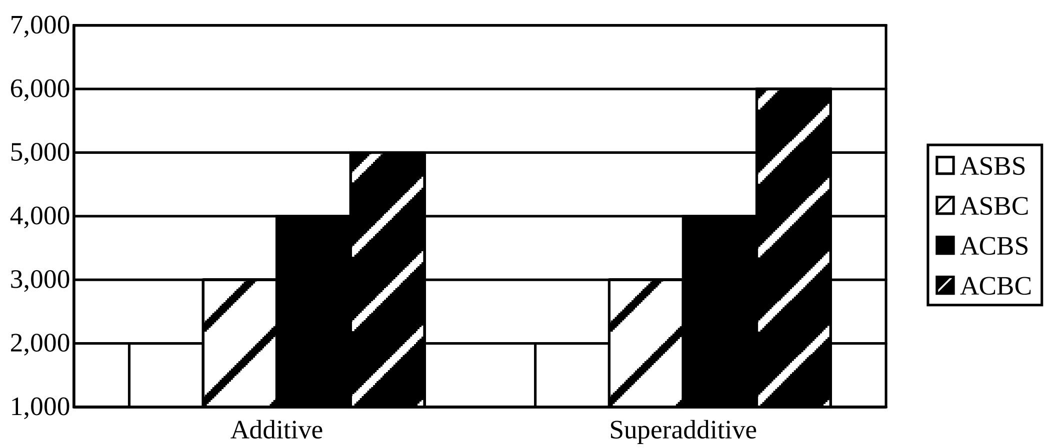


Figure 1. Two additive (left) and two superadditive effects (right) in ms for the hypothetical conditions A simple/B simple (ASBS), A simple/B complex (ASBC), A complex/B simple (ACBS), and A complex/B complex (ACBC).

If two effects are additive, the response to the most complex condition equals the sum of the effect sizes of factor A and factor B. This situation is depicted on the left-hand side of Figure 1. If two effects are superadditive, the response to the most complex condition exceeds that sum; this situation is shown in Figure 1 on the right-hand side.

Limitations of Sternberg's Interaction Logic. Sternberg's (1969) interaction logic has two limitations. First, it is important to notice that effects arising in a single cognitive system tend to add up; they only produce an interaction once the system has been taxed to its limits. This implies that independent effects may reflect either the fact that the factors manipulated are processed by separate systems, or the fact that the system dealing with the two factors has not been sufficiently loaded.

Second, an interaction tells you that two factors share a common resource pool, but this need not always lead to the conclusion that the two factors are processed by one and the same system. One could also imagine a situation where the two factors are processed by two independent cognitive systems that are both dependent on a third module. In that case, an interaction reflects a competition for resources within the third module. This situation is depicted in Figure 2 below.

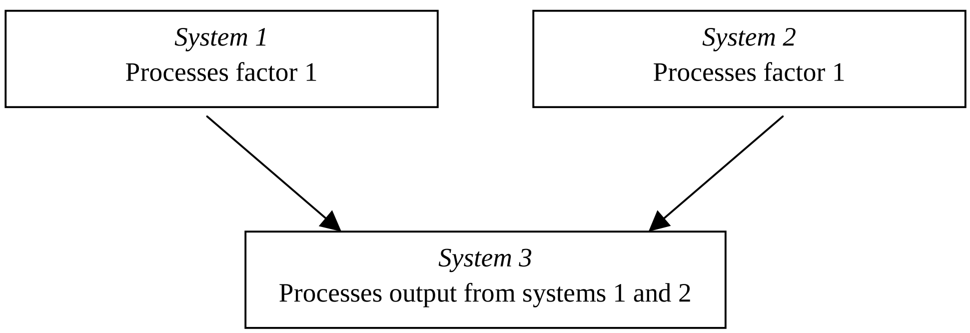


Figure 2. Model in which two independent systems cause a two-way interaction.

The Logic of Double Dissociations

Another way to test the independence of two cognitive processes is looking for double dissociations. Two systems can be considered separate if there are two patients suffering from damage to the central nervous system who show opposite patterns of impairment. This is the case when patient A has damage to one system, but intact functioning of a second system and patient B has intact functioning of the first, but not of the second system.

An example of a double dissociation is the well-documented dissociation between long and short-term memory (Baddeley, 1986, p. 114). There are patients who have extreme limitations of short-term storage capacity, but whose memory for long-term events is unaffected. On the other hand, patients suffering from Alzheimer's disease often show good short-term storage capacity (as demonstrated by the considerable number of digits they can rehearse) in the context of impaired long-term storage (as shown by the forgetting of events that happened in the near past.)

Baddeley's WM Model

The most influential WM model was developed by Baddeley and coworkers. Their model is primarily concerned with storage of verbal and visual information rather than with sentence processing (Baddeley, 1986; Baddeley & Hitch, 1974; Gathercole & Baddeley, 1993). Baddeley proposed a WM model consisting of several components. He developed this model based on experimental data from healthy volunteers (applying the logic of interactions) and from patients suffering from damage to the central nervous system (applying the logic of double dissociations). Within his model, a number of dissociations can be made. These dissociations will be discussed next.

Visual vs. Verbal WM

Baddeley and colleagues have proposed a number of subcomponents within WM. The first dissociation in their model is between visual and verbal WM. This distinction was proposed because concurrent visual WM tasks impede visual, but not verbal processing, and vice versa. For example, repeatedly touching the four corners of a monitor deteriorates performance on visuo-spatial tasks, but has no detrimental effect on verbal tasks, such as remembering a list of words. Conversely, concurrent articulation of verbal material impairs memory for word lists, but not for spatial information. In terms of Sternberg's (1969) interaction logic, visual and verbal WM are independent because visual and verbal tasks produce additive effects.

Additional evidence for the independence of visual and verbal WM comes from neuropsychological data. Patients have been reported with selective impairments of verbal, but not of visual WM; on the other hand, there are patients with the opposite pattern of impairment.

The distinction between visual and verbal WM is incorporated in Baddeley's (1986) model by the assumption of two modules that Baddeley refers to as *slave systems.* One is for visuo-spatial processing (the *visuo-spatial sketchpad*), and the other is for verbal processing, (the *phonological loop*).[5] In addition to these two slave systems, Baddeley assumes a central control mechanism, the *central executive*. He makes this assumption for the following reasons:

> "(…) Graham Hitch and I abandoned the view of working memory as a single unitary store. We substituted the idea of a number of subsystems controlled by a limited capacity executive system. The danger here, of course is to produce a system that is so complex as to be untestable and unproductive. The problem with a multiple system is to ensure that one does not simply invent another store or control process whenever an embarrassing result occurs. (…) We therefore chose to operate initially with a tripartite system, comprising a supervisory controlling system, the *Central Executive* aided by two slave systems, one which was specialized for processing language material, the *Articulatory Loop* and the other concerned with visuo-spatial memory, the *Visuo-Spatial Scratch Pad* or *Sketch Pad.*" (Baddeley, 1986, pp. 70-71)

The visuo-spatial sketchpad is not very relevant from a language processing point of view. Therefore, it will be ignored from now on. In the next subsection, we will look at subdivisions that can be made within the phonological-loop component of WM.

Partitioning the Phonological Loop: Phonological Storage vs. Articulatory Rehearsal

Once Baddeley had concluded that there were separate memory mechanisms for visual and verbal information, he proposed a partition of these mechanisms into even smaller units (Baddeley, 1986; Gathercole & Baddeley, 1993). Baddeley and coworkers proposed a division of the phonological loop into a *phonological store* and an *articulatory rehearsal mechanism* because of interactions between a number of effects. These effects will be discussed next, followed by an overview of the interactions that occur between the different effects.

The Phonological-Similarity Effect. The first finding to be mentioned is that subjects have more difficulty remembering phonologically similar items than phonologically dissimilar items. For example, it is harder to remember lists of letters that rhyme to each other (e.g., *B, G, C, D*) than it is to remember lists of letters that do not (e.g., *B, K, L, Y*). Retrieving the right items is not problematic, but subjects tend to confuse the order in which the words were presented.

---

[5] The phonological loop is also referred to by Baddeley (1986) as *articulatory loop*. The term *articulatory loop* is also used in cases of perceptual processing. The two expressions will be used interchangeably i.e., as synonyms in this book.

Baddeley and colleagues accounted for this effect by positing a *phonological store,* which is responsible for processing auditory input. The phonological-similarity effect is explained in terms of confusability of traces. Dissimilar words are stored as phonological traces that are clearly distinct from each other, whereas the similar words share so many features their traces are confused in the phonological store.

The Auditory-Suppression Effect. When subjects have to remember word lists or understand sentences, their performance deteriorates when irrelevant speech sounds are presented relative to performance in no-load conditions. This is the case for both auditorily and visually presented materials. The effect is called the *auditory-suppression effect* because it suppresses the auditory memory capacity.

It is worth noting that only speech sounds have this effect on verbal material. Martin, Wogalter, & Forlano (1988) found that sentence comprehension was affected by the concurrent presentation of speech sounds, but not by that of music, whereas identification of music deteriorated under concurrent presentation of irrelevant melodies, but not of irrelevant speech.

Baddeley (1986) explained this effect by assuming that the phonological store has limited capacity and that exposing participants to irrelevant speech reduces the available storage capacity of the phonological store.

Interactions of Phonological-Loop Effects. The phonological-similarity effect is exacerbated under auditory suppression. According to Sternberg's (1969) interaction logic, the two factors should therefore draw on the same resource pool.

The Word Length Effect. Baddeley (1986) assumed that the word length effect arised in a different part of the phonological loop, the articulatory rehearsal mechanism. Subjects have lower memory spans for lists containing long words than for lists with short words. Word length has to do with the time it takes to pronounce or perceive a word. Not only the numbers of phonemes in the words matter, but also the duration of these phonemes. For example, although the words *bead* and *bid* have the same number of phonemes, the duration of the /i:/ in *bead* is longer than that of the /ɪ/ in *bid.*

Baddeley (1986) assumed that this effect reflects the fact that people use a kind of repetitive tape loop for the verbal rehearsal of words. Because it takes longer to pronounce long words, less of these words fit in the tape loop.

The Articulatory-Suppression Effect. Articulatory suppression (or irrelevant speech) means that subjects have to pronounce sounds or words out loud while remembering word lists. In the articulatory-suppression conditions, subjects have worse memory for word lists than in a no-load condition. Articulatory suppression reduces the availability of the articulatory rehearsal mechanism because part of its capacity is used for the pronunciation of the words.

Interactions of Effects Arising in the Articulatory Rehearsal Mechanism. When participants do not have to perform a secondary task, they show a word length effect: longer words are remembered less well than short words. However, this effect disappears under articulatory suppression.

The interaction of word length and articulatory suppression suggests that the advantage of shorter words is only present as long as participants can use articulatory rehearsal. In other words, when their tape loop is loaded with irrelevant material, the word length effect disappears because words cannot be remembered via articulatory rehearsal anymore.

The Connection Between the Phonological Store and the Rehearsal Mechanism. The interaction of articulatory suppression by phonological similarity shows how the subcomponents of the phonological loop work together. When words are presented visually, the phonological-similarity effect disappears under articulatory suppression. With auditory presentation, the phonological-similarity effect is present, even under articulatory suppression.

This interaction suggests a model in which auditory input enters the phonological loop via the phonological store. There, it is stored temporarily as an acoustic trace. In order to prevent the trace from decaying, articulatory rehearsal can be used: when a person articulates the acoustic trace, that trace is refreshed and fed into the phonological store. Moreover, with auditory presentation, a phonological representation exists and it is susceptible to phonological similarity effects, even in articulatory-suppression conditions.

On the other hand, a phonological store representation apparently can exist with visual presentation also. Baddeley (1986) argues that in addition to its rehearsal function, the rehearsal mechanism also seems to mediate grapheme-to-phoneme conversion. When written words are pronounced, they become available in an acoustic format that can be stored in the phonological store. The fact that articulatory suppression abandons the phonological-similarity effect in the written modality suggests that it prevents the phonological loop from doing the grapheme-to-phoneme conversion.

Central-Executive Processing

In contrast to the slave systems, which have very specific functions, the central executive is a general-purpose cognitive system. It is taken to play a major role in the interpretation of sentence meaning (Gathercole & Baddeley, 1993; Just & Carpenter, 1992). The exact nature and functions of the central executive remain issues that are far from resolved. It is taken to be active in dual-task processing (Babcock & Salthouse, 1990; Gathercole & Baddeley, 1993; Salthouse, 1994). Carpenter and coworkers have made important contributions to research on central-executive functioning and language (Gathercole & Baddeley, p. 222; Just & Carpenter, p. 123). They assume that the central executive is the module that is involved in sentence processing. Figure 3 shows the different components of Baddeley's WM model. Just & Carpenter's theory of language processing and WM is outlined in the following section.

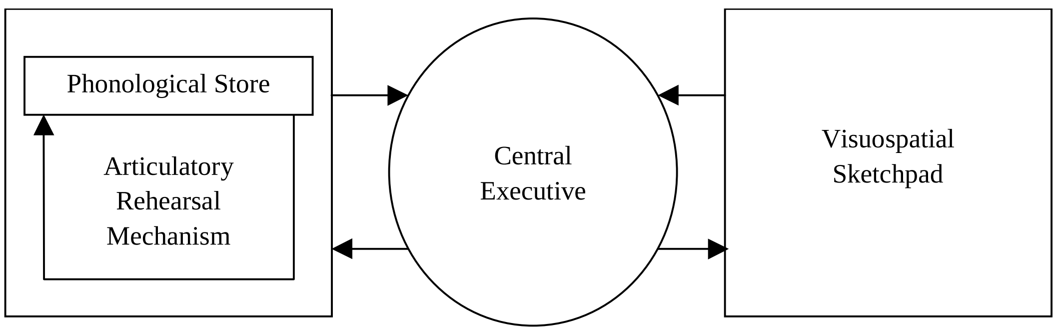


Figure 3. Baddeley's WM model

Attempts to Apply Baddeley's WM Model to Sentence Processing

Over the last decades, several attempts have been made that applied Baddeley's (1986) WM model to sentence comprehension. The most important ones were proposed by Carpenter and colleagues (e.g., Daneman & Carpenter, 1980; Just & Carpenter, 1992; Just, Carpenter, Keller, Eddy, & Thulborn, 1996), Caplan & colleagues (e.g., Caplan & Waters, 1999; Waters et al., 1987), and Martin & Romani (1994).

The Single-Resource Theory

Just & Carpenter (1992) claim that a single WM system is addressed during sentence processing. It performs both storage and processing of information. Just &

Carpenter posit that this single WM system corresponds to Baddeley's (1986) central executive.

In the single-resource theory, information elements have levels of activation associated with them. For an element to be active, its activation must exceed threshold level. Threshold is reached through cycles of activation. The activations of individual elements add up until the maximum capacity of the WM system is reached. Then all elements will become less active, which leads to de-allocation of the least active elements and/or a slow down of processing.

If a substantial part of WM is used, it will take more cycles of activation for each element to reach threshold, and in some cases elements will not reach threshold at all, which results in forgetting of those elements. Thus, under high WM demands, processing will be slower and less accurate.

According to Just & Carpenter (1992), there are individual differences in WM capacity. People with more resources should be better at processing complex sentences than people with fewer resources.

Daneman & Carpenter (1980) developed the Reading Span Test to measure a participant's resources. In this task, participants read sets of sentences aloud while they have to remember the last word of each sentence. Sets vary in length from two to seven sentences. The largest number of sentences a participant is able to remember determines his/her span size. Daneman & Carpenter claim that the Reading Span Test provides a reliable measure of a participant's linguistic processing and storage capacity.

Whether high-spanners are better comprehenders than low spanners can be demonstrated in self-paced reading experiments: high-capacity participants should have faster reaction times than low-capacity participants. King & Just (1991) found that low-span participants had longer reading times for syntactically difficult sentences than high-span participants. In addition, low-span participants had more difficulty understanding complex sentences than high-spanners in conditions with extrinsic memory loads. These data were also the basis for the assumption of a limited-capacity WM system within Gibson's (1998) SPLT model.

Since the single-resource theory does not deal with peripheral storage, such as phonological storage or articulatory rehearsal, it does not make clear predictions about phonological-loop manipulations.

The Separate Language Interpretation Resource Hypothesis

The separate language interpretation resource or SLIR hypothesis (Caplan & Waters, 1999) states that (at least) two WM systems are involved in language comprehension. The first one is used in *interpretive* processing. The interpretive stage entails processes like acoustic-phonetic conversion, lexical access, syntactic processing, recognition of intonational contours, and determining the discourse level semantic value of a sentence.

Since these processes are dealt with by the same system, they should produce superadditive interactions when sufficiently loaded. Caplan & Waters (1999) found no super-additive interactions of interpretive processes with additional memory load or correlations with memory spans.

The second WM system is concerned with *postinterpretive* processing. Postinterpretive processing entails the integration of propositions and interpretation of referential elements such as pronouns. Caplan and coworkers usually manipulate post-interpretive complexity by contrasting sentences with one and two propositions (cf. 10a and b) because in two-proposition sentences, the propositions must be linked and additional thematic-role relations must be set up:

(10) a. It was the jeweller who adjusted the broken clock
b. The man hit the landlord that requested the money

Thematic role assignment is a process in which it is determined who does what in a sentence. Waters et al. (1987) claim that thematic role assignment occurs in the postinterpretive stage as part of the linking of propositions.

Unlike the interpretive system, processes attributed to the second system show interactions with articulatory suppression and correlations with standard measures of WM capacity, including the reading span task. Caplan & Waters (1999) claim that in postinterpretive, but not in interpretive processing, a phonological backup representation in the phonological loop is addressed. They assume that addressing the phonological back-up representation is necessary to re-initiate a parse for thematic role assignment.

The dichotomy of resource pools proposed by Caplan & Waters (1999) is similar to the one made by Berwick & Weinberg (1984). This is because earlier versions of Caplan & Waters's theory were heavily influenced by Berwick & Weinberg's two-stage parser (e.g., Waters et al., 1987).

A Multiple Component View

Martin & Romani (1994) propose an extension of the Baddeley model along different lines than Waters et al. (1987): in addition to the phonological loop and the visuo-spatial sketchpad, they assume separate slave systems for syntactic and semantic storage. They argue for separate phonological and semantic slave systems on the basis of an apparent dissociation between two patients: E.A. and A.B.. E.A. is a patient with a phonological deficit along with unimpaired processing and maintenance of semantic information. She performed better on a word span than on a nonword span task. This suggests that she has intact semantic storage but impaired phonological storage. For patient A.B., the opposite held: he was better at a rhyme probe task than at a category probe task. This seems to indicate that A.B. has intact phonological storage and impaired semantic storage.

Additionally, Martin & Romani (1994) claim that semantic and syntactic storage are separate on the basis of a double dissociation between patients A.B. and M.W.. A.B. was not able to retain unintegrated semantic information, whereas M.W. had difficulty keeping incomplete syntactic structures in memory.

Martin & Romani (1994) argue that both the lexicosemantic and the syntactic storage component is important for language processing, but that phonological WM representations play little or no role during the comprehension of sentences since E.A. shows no clear comprehension deficit.

Predictions of the Three Theories

Table 1

Superadditive Interactions Predicted by Just & Carpenter, Caplan & Waters, and Martin & Romani

Manipulations	Phonological Loop	Lexicosemantic	Syntactic	Propositional
		Interactions predicted by Just & Carpenter (1992)		
Phonological Loop	yes	yes	yes	yes
Lexicosemantic		yes	yes	yes
Syntactic			yes	yes
Propositional				yes
		Interactions predicted by Caplan & Waters (1999)		
Phonological Loop	yes	no	no	yes
Lexicosemantic		yes	yes	no
Syntactic			yes	no
Propositional				yes
		Interactions predicted by Martin & Romani (1994)		
Phonological Loop	yes	no	no	?
Lexicosemantic		yes	no	?
Syntactic			yes	?
Propositional				?

Note. This table gives an overview of superadditive interactions predicted by the different theories. *Yes* indicates that an interaction is predicted by a theory; *no* indicates that it is not; *?* indicates that a theory does not make predictions about a particular complexity effect.

Table 1 gives an overview of differences in predictions made by the single-resource theory proposed by Just & Carpenter (1992), Caplan & Waters's (1999), and Martin & Romani's (1994) theories.

Goals of the Current Research

As already stated, the current research aims to develop a model of WM and sentence comprehension that specifies the WM resources used during sentence comprehension. The research discussed in this book has a strong emphasis on the use of phonological WM representations during sentence comprehension. This book will present a number of Dutch experiments that were based on an experiment conducted by Waters et al. (1987, exp. 3).

The goals of these experiments are the following:

1. To re-evaluate the model proposed by Waters et al. (1987) and Caplan & Waters (1999);
2. to replicate the experiment conducted by Waters et al. (1987, exp. 3) with Dutch materials; and
3. to develop a new model of WM and sentence processing with a stronger emphasis on the theoretical bases of a role of the phonological loop in sentence processing.

These three issues will be discussed in the following parts.

An Evaluation of the SLIR Hypothesis

Caplan & Waters (1999) and Waters et al. (1987) proposed a model with two WM resource pools for sentence comprehension. One resource pool is used during syntactic processing and the other for propositional processing. Waters et al. argue that during the latter stage, a phonological back-up representation of the sentence is addressed.

Caplan and colleagues based their model on a whole-sentence anomaly-judgment experiment (Waters et al., 1987, exp. 3) that varied three factors. The first stage of their model was tested by varying subject- and object-relative clauses (SRs and ORs). Waters et al. tested the second stage of their model by contrasting sentences with one and two propositions (1P and 2P). Since their design was orthogonal, the experiment had four sentence conditions: 1PSR, 1POR, 2PSR, and 2POR. Example sentences of each condition are given below in (11a to d):[6]

(11) a. It was the thief that broke into the warehouse
b. It was the broken clock that the jeweller adjusted
c. The man hit the landlord that requested the money
d. The meat that the butcher cut delighted the customer

The involvement of a phonological back-up was tested by contrasting a no-load condition with an articulatory-suppression condition.

Waters et al. (1987) found independent effects of Clause Type (SR vs. OR) and Propositional Complexity (1P vs. 2P), which supports the idea that the two factors are processed by separate systems. Furthermore, they found a two-way interaction of Propositional Complexity x Articulatory Suppression (no load vs. counting out loud), but not of Clause Type x Articulatory Suppression. This finding led them to believe that a phonological back-up was addressed during propositional, but not during syntactic processing.

The assumption of a role for a phonological back-up during propositional processing must be questioned on both theoretical and empirical grounds. During sentence processing, readers see letters, identify words in the letter strings, maybe retrieve sound representations of the words, combine words into phrases, and combine phrases into propositions to arrive at an interpretation.

During this process, readers go from lower- to higher-level representations. Then why would they need a very low-level representation such as a phonological one for very high-level processes that occur during propositional analysis? This aspect of their theory is unlikely and no clear argument is given for how it could do so in practice.

Second, Waters et al.'s (1987) model is based on potentially confounded data. Consider examples (11a to d) above. The one-action sentences (11a and b) are right embedded, but (11d), a two-proposition sentence, is center embedded. Center-embedded structures have long been known to be harder than right-embedded structures (Gunter, 1995; Miller & Isard, 1964). Thus, it may be that the interaction found by Waters et al. is not an interaction of Propositional Complexity x Articulatory Suppression, but of Embedding Type x Articulatory Suppression.

[6] Technically speaking, sentences (11a and b) are not relative clauses, but cleft sentences; for convenience sake, these sentences will be referred to as one-propostion subject- and object-relative sentences.

Because of this methodological flaw in the Waters et al. (1987) design, their data deserve replication using an experiment that uses right-embedded structures only.

The Importance of Cross-Linguistic Replication

The second goal of the research presented in this book is cross-linguistic replication of a number of WM experiments using Dutch rather than English. Most of the models discussed above are based on English data, although the results are assumed to extend to all languages.[7] However, Dutch and English differ in structure. An important difference between the two languages is that Dutch embedded clauses are verb final necessitating frequent integration over longer time frames. This may have consequences for the information that readers use and for the strategies they apply to maintain it.

The fact that embedded clauses are verb final has consequences for the processing of the subject- vs. object-relative ambiguity. Subject- and object-relative clauses have identical surface orders, as can be seen in (12a and b):

(12) a. Dit is de jongen die de meisjes kust
This is the boy that the girls kiss$_{sg.}$
'This is the boy who kisses the girls'

b. Dit is de jongen die de meisjes kussen
This is the boy that the girls kiss$_{pl.}$
'This is the boy who the girls kiss'

Readers of English rely mainly on word order when processing these sentences, but readers of Dutch will not benefit from using word order as a cue. Instead, they must rely on subject-verb agreement information: if the embedded verb agrees with the first NP of the embedded clause, the clause is subject relative; if it agrees with the second NP, the clause is object relative.

Subject-verb agreement is a much subtler phonological cue than word order only. The difference between the subject- and object-relative reading of the clause hinges on the difference between the distinction between the third person singular inflection, /t/, and the third person plural inflection, /ə/.

Because of the reliance on subtle morphophonological information, readers of Dutch may depend on phonological memory more than English readers do,

[7] Note that this does not apply to Gibson's (1998) theory and to a number of case studies conducted by Baddeley and colleagues (e.g., Baddeley et al., 1987).

particularly given the fact that agreement occurs over a long distance given the verb-final order. Therefore, it is important to do cross-linguistic replications using Dutch rather than English data.

A New Role for Phonological Memory

Because of the differences in structure discussed above, it is quite possible that with Dutch data, different results will be obtained with respect to the role of phonological memory in sentence comprehension. These different results may have consequences for the specification of WM resources that are active during sentence comprehension. If different results are obtained using Dutch data, the existing models of (phonological) WM and sentence comprehension need to be reconsidered.

Chapter 3
About the Role of the Phonological Loop in Syntactic Analysis

Introduction

Over the last decades, there has been a lively debate about the involvement of some form of working memory (WM) in sentence comprehension. Several researchers have argued that some sentences are harder to understand than others because it takes more WM resources to process them (Caplan & Waters, 1999; Frazier & Rainer, 1988; Gibson, 1998; Just & Carpenter, 1992; King & Just, 1991; Waters, Caplan, & Hildebrandt, 1987).

Some people may challenge the assumption that WM plays an important role in sentence processing on the grounds that sentence comprehension occurs so fast and so automatized that WM is not needed in this process at all. However, it is commonly accepted that WM involves both processing and storage of information (e.g., Babcock & Salthouse, 1990; Baddeley, 1986; Gathercole & Baddeley, 1993; Gibson, 1998; Just & Carpenter, 1992).

A topic that re-emerges repeatedly in the discussion about WM and sentence processing is how many specific WM components are involved in sentence comprehension. On the one hand there are researchers who argue for a single WM component dealing with all sorts of language tasks (e.g., Just & Carpenter, 1992; King & Just, 1991). On the other hand, there are researchers who argue for a partition of the WM system involved in language comprehension (e.g., Caplan & Waters, 1999; Waters et al., 1987).

The goal of the current chapter is to re-examine the evidence put forward in favor of a partition of WM resources in Waters et al. (1987, exp. 3). Waters et al. proposed a particular division based on potentially confounded data. This chapter reports a replication of their experiment without this potential confound and a follow-up to that replication.

This chapter is written on a stand-alone basis. This means that readers should be able to read it without having read Chapter 2. For those who have skipped that chapter, the current chapter contains a small introduction. Readers who have read Chapter 2 may want to proceed to Experiment 1 on page 42 although some elements are more thoroughly discussed here.

Sternberg's Interaction Logic

In the discussion about how many WM resource pools are used during sentence comprehension, Sternberg's (1969) logic of interactions plays an important role. If one wants to study potential interactions, one needs a design that varies at least two factors. It is crucial whether two factors in a particular design interact or not. Suppose, for the sake of the argument, that one varies two factors, A and B, in a reaction time (RT) experiment. Let us assume that each factor has two levels, simple and complex. This will yield the following experimental conditions: A simple/B simple, A simple/B complex, A complex/B simple, and A complex/B complex. Now let us also assume that the simple condition of each factor yields faster RTs than the complex conditions.

It is important to find out what happens in the most complex condition of the experiment. One possibility is that the complexity effects of the two factors add up in that condition. However, if the most complex condition yields a response latency that exceeds the sum of the effect sizes of factors A and B, these factors produce a so-called superadditive interaction. Sternberg proposed that if factors add up, they are processed by different cognitive systems, but if they produce a superadditive interaction, they are processed a single cognitive resource pool.

Although this logic seems straightforward, there are a number of issues to be kept in mind. First, interactions only become apparent when a cognitive system is taxed to its limits. Even if two factors are processed by a single system, they may not interact, because that system has much more capacity than the amount needed for the processing of the two factors. Additionally, interactions may also occur due to ceiling or floor effects. In this case, the interactions are not superadditive (like in the example above), but non-additive: the response to the most complex condition is less than the sum of effects of A and B. In the case of nonadditive interactions, the Sternberg logic can not be applied.

Baddeley's WM Model

I will now discuss the most important proposals about the involvement of WM systems during sentence processing. Baddeley developed a model in which WM is used for both storage and processing of information (e.g., Baddeley, 1986; Gathercole

& Baddeley, 1993). The model consists of several components. Its principal component is the *central executive,* which is involved in higher level attentional processes, dual-task processing, storage of information, and manipulation of partial results.

If the central executive needs additional storage capacity, it addresses so-called *slave systems.* Baddeley (1986) assumes at least two of these slave systems, the *visuo-spatial sketchpad* (for visual processing) and the *phonological loop* (for auditory processing).[1] The visuo-spatial sketchpad is not important for the purposes of the current study and will be ignored from now on.

The phonological-loop component of Baddeley's (1986) model is taken to consist of two parts. The first part is a module called the phonological store. This is a passive storage facility for acoustic traces, which are subject to decay. If the traces are in danger to decay too quickly, a second module, the *articulatory rehearsal mechanism,* can help refreshing them via inner speech. In addition to refreshing the phonological traces, the rehearsal mechanism also mediates conversion of written text to a sound representation, which is then kept in the phonological store.

Gathercole & Baddeley (1993) assume that processes such as syntactic analysis occur in the central executive. Baddeley, Vallar, & Wilson (1987) proposed that under some circumstances, the central executive makes use of the phonological loop during syntactic analysis.

> "We would like to propose a (…) view whereby the phonological storage component of the articulatory loop acts as a 'mnemonic window,' holding sequences of incoming discourse and allowing the components of such sequences to be processed and interrelated. We assume that comprehension occurs on-line, with words, and possibly fragments of words, accessing their phonological and semantic representation as they are heard (…). The comprehension of connected discourse, however, involves more than the understanding of individual discrete items, with order information typically being very important in performing the necessary syntactic analysis." (Baddeley et al., 1987, pp. 526-27)

This proposal predicts that manipulations of syntactic complexity should produce interactions with manipulations that tax the phonological loop.

A Single-Resource Theory of Sentence Processing

Just & Carpenter (1992) developed another model that combined Baddeley's (1986) model with a computational model for sentence processing. They proposed

---

[1] The phonological loop is also referred to by Baddeley (1986) as *articulatory loop.* The two expressions will be used interchangeably in this book.

that a single WM system is addressed during sentence processing. It performs both storage and processing of information. In their model, information elements have levels of activation associated with them. For an element to be active, its activation must exceed threshold level. Threshold is reached through cycles of activation. The activations of individual elements add up until the maximum capacity of the WM system is reached. Then all elements will become less active, finally leading to de-allocation of the least active elements and/or the slow-down of processing.

If a substantial part of WM is used, it will take more cycles of activation for each element to reach threshold, and in some cases elements will not reach threshold at all, resulting in forgetting of those elements. Thus, under high WM demands, processing will be slower and less accurate.

Just & Carpenter (1992) claim that there are individual differences in WM capacity. People with more resources should be better at processing complex sentences than people with fewer resources. Daneman & Carpenter (1980) developed the Reading Span Task to measure a participant's resources. Whether high spanners are better comprehenders than low spanners can be demonstrated in self-paced reading experiments: high-capacity participants should have faster reaction times than low-capacity participants.

King & Just (1991) found that low-span participants had longer reading times for syntactically difficult sentences than high-span participants. In addition, low-span participants had more difficulty understanding complex sentences than high spanners in conditions with extrinsic memory loads. In extrinsic memory load conditions, participants have to remember a number of words while reading a sentence. The amount of words to be remembered usually varies from one to three. Unlike the Waters et al. (1987) model, Just & Carpenter's (1992) model predicts interactions of manipulations of all sorts of complexities related to language processing.

Just & Carpenter (1992) proposed that this parsing system is similar to Baddeley's (1986) central executive. They do not assume any other WM systems, such as a phonological loop. Therefore, clear-cut predictions with respect to interactions with phonological-loop loads cannot be derived from their theory.

The Syntactic Prediction Locality Theory

A fourth model that relates sentence complexity effects to WM demands was developed by Gibson (1998). Like Baddeley (1986), Gibson proposed that WM resources can be used for both storage and processing of information. He assumed a parsing mechanism with limited capacity, in which storage and processing compete for resources. The more capacity is used up for storage, the less there is left for processing. Thus, high storage demands imply slower processing times.

In Gibson's (1998) model, memory demands are defined in terms of the minimum number of predictions needed to finish a sentence grammatically. For example, at the verb *puts* in (1), there are minimally two predictions, since the verb *put* requires at least a direct object and a locative preposition:

(1) John puts the cake on the table

The number of open predictions is not the only demanding factor here: when predictions must be kept active for a longer time, they become more demanding.

The second component in Gibson's (1998) model is integration difficulty. Gibson assumes that integration becomes harder when it has to be established over a longer distance. In his complexity metric, distance is determined as the number of new discourse referents over which the integration takes place. Thus in (1), integration of the word *cake* with the verb *put* requires only two active new discourse referents, *cake* and *put*. Integration of the word *table* requires three active referents: *put*, *cake*, and *table.* Therefore, *table* has a higher integration cost than *cake.*

Given the fact that Gibson does not mention additional WM systems such as the articulatory loop, it is uncertain what his theory predicts with respect to articulatory-loop loads.

The Separate Language Interpretation Resource Theory

A second proposal about the role of WM in sentence comprehension was made by Waters et al. (1987). They conducted a series of experiments that were aimed at establishing a connection between Baddeley's (1986) WM model and the parsing model developed by Berwick & Weinberg (1984).[2] Waters et al. used Berwick & Weinberg's parsing model as a specification of the actions carried out by the central executive.

The reason for choosing Baddeley's (1986) model is that it is both well researched and widely accepted. Waters et al. (1987) adapted the Berwick & Weinberg (1984) model because it has three important features: it is specific, it is computationally plausible, and it makes clear-cut predictions. The model is specific in the sense that it states every step the parser takes in the process of sentence comprehension. Berwick

[2] Sentence comprehension, more specifically syntactic processing, is often referred to as *parsing*. From now on, *syntactic processing* and *parsing* will be used interchangeably and as synonyms.

& Weinberg tried to make their model computationally plausible by requiring that it mimic human sentence comprehension: it has to be able to understand sentences in real time, and it has a limited processing capacity. In addition, because the model is so detailed, it yields testable predictions.

The Berwick & Weinberg Parser. Berwick & Weinberg's (1984) sentence comprehension model uses government and binding theory (Chomsky, 1981) as its linguistic framework. For a concise description of the Berwick & Weinberg parser, the reader is referred to Caplan & Hildebrandt (1988, pp. 32-42), who have provided an outstanding and informative summary of the model. For now, a very brief outline of the parser suffices.

For reasons of simplification, Berwick & Weinberg (1984) made two assumptions. First, the input to the parser has already been segmented into words, and second, the parser has access to the lexical representations of those words, including grammatical category and information about the complements each word can select.[3]

In addition to those two assumptions, Berwick & Weinberg (1984) make two proposals about the architecture of the parsing mechanism. Their parser is deterministic, which means that it processes sentences in a left-to-right manner, and it lacks a backtracking mechanism. If the parser builds the wrong structure, it must start all over again from the beginning of the sentence. Second, the parser has an input buffer with three cells. It takes the information in the first cell as input, but it is allowed to look ahead at the information of the next two cells.

Berwick & Weinberg (1984) assume a deterministic parser with a limited look-ahead buffer for three reasons. In this way, the fact that humans parse in linear time is mimicked. Second, the limited look-ahead option enables the parser to eliminate some local ambiguities. Finally, determinism and a limited look-ahead buffer are used to mimic the limitations in processing capacity observed in humans.

Berwick & Weinberg's (1984) model consists of two modules, a syntactic processor and a propositional list. The syntactic processor contains an input buffer where words come in. As soon as words can be combined to a phrase, the syntactic processor stores that phrase in a so-called active node stack. There, the phrase is extended as long as the incoming words can still be attached to that phrase. When a new word cannot be integrated into the currently active phrase, that phrase is

[3] The assumption that words have already been segmented before entering the syntactic stage is actually not compatible with Caplan & Waters's (1999) assumption that processing up to syntactic analysis occurs within the interpretive stage. This issue will be re-addressed later in this section.

transferred to a pushdown stack, and the new word serves as the beginning of a new phrase in the active node stack.

Phrases (or parts thereof) that are stored in the different stacks are combined as soon as they can constitute an entire clause. Once the clause has been built, syntactic analysis is complete.

Then, the clause is transferred to a second processing stage, the propositional list. The assumption of two processing stages offers a functional account of the distinction in government and binding theory between reflexives and pronouns. Antecedents of traces and reflexives are detected during syntactic processing, whereas the antecedents of PRO and pronouns are established during the semantic stage (i.e., in the propositional list). Reflexives must always be bound within their governing category, in this case the embedded clause, whereas pronouns must be free within that category.

This is shown below: in sentence (2a), the reflexive *himself* refers to the subject of the embedded clause, *who.* However, the pronoun *him* in (2b) cannot refer to a constituent within the clause; instead it refers to the matrix subject, *John.* In the examples below, words that have the same referent, are co-indexed with the same subscript letter, in this case *i* and *j:*

(2) a. John talks to the man_i who_i is shaving himself_i
b. John_j talks to the man_i who_i is shaving him_j

Waters et al.'s Adaptation of the Berwick & Weinberg Parser. Waters et al. (1987) assume that two WM resource pools are addressed during sentence comprehension. One resource pool is deployed during the first processing stage of the Berwick & Weinberg (1984) model (during the construction of syntactic structure), and the second resource pool is used during the second stage, when propositional content is checked. During that checking procedure, a phonological back-up representation is used. Waters et al. assume that in that stage "reference to this form of lexical representation might initiate enough of a parse to establish the first thematic roles and other sentential features which are licensed by the parser, thereby also establishing which are erroneously inferred pragmatically" (Waters et al., 1987, p. 534).[4]

[4] In this sense, Waters et al.'s proposal is reminiscent of the gardenpath theory (Frazier, 1979).

This proposal differs from the one made by Baddeley et al. (1987) in two respects. Baddeley et al. assume that a back-up is kept in the phonological store and that it is used for syntactic processing as well as other processes. Waters et al. (1987) assume an articulatory back-up which is addressed only after the syntactic processing stage. And second, Baddeley et al. assume one central executive rather than a partition of central-executive resources.

Waters et al. (1987) refer to the second stage as the *postsyntactic* stage. In a later version of their model (Caplan & Waters, 1999), the processing occurring at this stage is referred to as *postinterpretive* processing. The reason for this terminology shift is the fact that Caplan & Waters extended the Waters et al. model. The extension consists of the assumption that the first stage not only involves syntactic analysis, but also processes that precede it, such as phoneme detection, word recognition, etc. When discussing the models proposed by Caplan & Waters and Waters et al., I will use the terms *postinterpretive* and *postsyntactic;* in other contexts, I will use the theory-neutral term *propositional,* which refers to the experimental support for this second stage.

The Waters et al. (1987) model predicts that phonological-loop loads should interact with postsyntactic manipulations. However, postsyntactic complexity should not interact with syntactic complexity, and articulatory-suppression manipulations should not produce such interactions either.

Re-examining Evidence for Separate Resources

Waters et al.'s Design. After this overview of the most prominent models of WM and sentence processing, I will discuss the experiment on which Waters et al. (1987) based their conclusions. Waters et al. tested the first or *syntactic* stage of their model by comparing sentences containing subject- (SR) and object-relative (OR) clauses (3a and b respectively; examples taken from King & Just, 1991, p. 582):[5]

(3) a. The reporter e_i $that_i$ t_i attacked the senator admitted the error
b. The reporter e_i $that_i$ the senator attacked t_i admitted the error

---

[5] Within linguistic theory, it is assumed that the relative pronoun *that* is coindexed with an empty wh-element that moves from an underlying position to clause-initial position. The empty wh-element is indicated with an *e.* It is taken to leave a trace behind at the position before movement (e.g., Chomsky, 1981, 1995). This trace is indicated with a *t.* The fact that the trace, the empty wh-element, and the relative pronoun refer to the same entity is denoted with a subscript *i*.

There are good reasons for looking at exactly this comparison: the two sentence types differ in terms of syntactic complexity and processing difficulty. The differences in processing difficulty have been accounted for in terms of differences in WM load. King & Just (1991, p. 581) attribute this complexity difference to the fact that the filler, *that*, has to be stored longer in the object relative than in its subject-relative counterpart (cf. Frazier's, 1978 active filler strategy and Wanner & Maratsos, 1978). As can be seen in (3) above, the filler and the position it refers to are adjacent in (3a), but not in (3b), where the filler must be kept active until after the embedded verb *attacked.*

Waters et al. (1987) examined the second or postsyntactic stage of the parsing model by comparing sentences with one proposition (1P) and sentences with two propositions (2P). They did this because in the Berwick & Weinberg (1984) model, which they adopted, clauses are transferred to a propositional list as soon as they are complete. When in the propositional list, the clauses are not available to the syntactic processor. Therefore, their model predicts that the number-of-actions manipulation should affect the second, but not the first stage. Combining the two factors (Clause Type and Propositional Complexity) yields four sentence conditions: 1PSR (cf. 4a), 1POR (cf. 4b), 2PSR (cf. 4c), and 2POR (cf. 4d):[6]

(4) a. It was the thief that broke into the warehouse
b. It was the broken clock that the jeweller adjusted
c. The man hit the landlord that requested the money
d. The meat that the butcher cut delighted the customer

The third factor in their design was Articulatory Suppression. Waters et al. (1987) had three memory load conditions: no load, articulatory suppression (counting from 1 to 6 out loud), and finger tapping. The choice of the articulatory-suppression condition stems from the Baddeley (1986) model. As mentioned above, the model consists of a central executive and (at least) two slave systems. One of these is the phonological loop, which consists of a phonological store and an articulatory rehearsal mechanism. In the phonological store, acoustic traces are maintained; these traces can be refreshed via inner speech in the articulatory rehearsal mechanism.

Articulatory suppression is a secondary task that consists of uttering irrelevant speech sounds. Because uttering the speech sounds prevents full functioning of the

[6] Technically speaking, sentences (4a and b) are not relative clauses, but cleft sentences; for the sake of convenience, these sentences will be referred to as one-propostion subject- and object-relative sentences.

articulatory rehearsal mechanism, phonological-loop functioning in WM is hindered. This should reduce the availability of the loop during sentence processing.

If interactions of Articulatory Suppression and one of the other two factors are found, one can conclude that the phonological loop is used by that specific module. However, comparing a no-load condition to an articulatory-suppression condition introduces a potential concern: one cannot be sure about the nature of the suppression effects. They can result from the fact that the phonological loop is taxed, but they may also reflect mere motor activity (because participants have to move their mouths) or distraction by a secondary task.

This is the reason why Waters et al. (1987) had a third condition, finger tapping. Finger tapping is as complex as counting in terms of motor activity. (In both the counting and the tapping condition, participants were instructed to count or tap as fast as possible.) Having these three load conditions offers the opportunity to rule out the movement or second-task confound.

If the tapping condition is similar to the no-load condition, but differs from the counting condition, one can conclude that the effects of counting are due to articulatory suppression and not to movement per se. However, if tapping and counting produce comparable effects, the effects probably reflect motor activity. The finger tapping condition produced the same results as the no-load condition, whereas the counting condition differed from the other two.

Waters et al. (1987) found that Clause Type and Propositional Complexity failed to interact. Additionally, Propositional Complexity interacted with Articulatory Suppression, but Clause Type, a syntactic factor did not. They concluded that this data pattern offers support for a two-stage processing model of sentence comprehension along the lines of Berwick & Weinberg (1984).

Waters et al. (1987) claim that inner speech does not play a role in the first processing stage, but that it is needed for the second stage, in which propositional content is checked. They argue that in order to evaluate the propositional content of a sentence, people use a phonological back-up representation, which has been kept active via rehearsal.

Re-evaluating Waters et al.'s Conclusions. Waters et al.'s (1987) interaction of Propositional Complexity x Articulatory Suppression may seem to point to a role for inner speech in this process. However, without further specification, this assumption seems slightly farfetched. What role could a low-level phonological representation possibly play in propositional content checking? It seems unlikely that readers would rely on low-level information such as a phonological representation of the sentence when looking for antecedents for pronouns.

Because the involvement of inner speech in propositional content checking is questionable from a theoretical point of view, it is interesting to take a closer look at the experimental design used by Waters et al. (1987). A striking feature of their design is that all one-proposition sentences (clefts), like (4a), repeated here as (5a), contain by their very nature only right-embedded clauses. The two-proposition sentences, like sentence (4d), repeated here as (5b), however, contain center-embedded clauses as well. The relative clause *that the butcher cut* appears in the middle of the main clause and not at the end:

(5) a. It was the broken clock that the jeweller adjusted
b. The meat that the butcher cut delighted the customer

Center-embedded structures have long been known to be fairly complex (Gunter, 1995; Miller & Isard, 1964). Processing of these constructions is more memory demanding, because the main-clause subject cannot be integrated immediately: it has to be kept active until the embedded clause has been processed. This may be done by the phonological-loop component of Baddeley's (1986) WM model.

Right-embedded structures are easier, because they can be understood one clause at a time. Due to the use of center-embedded structures, it may well be that the propositional effects found by Waters et al. (1987) do not reflect postsyntactic processing, but processing of the center embeddings.

The goal of the current study is twofold. First, to investigate which part of the Waters et al. (1987) results must be accounted for in terms of the sort of embedded structures they used and second, to look at the role of the phonological loop in sentence processing.

Experiment 1

Introduction

The fact that Waters et al. (1987, exp. 3) used center-embedded structures implies that the propositional effect, which they labeled as a sentence-level semantic effect, may in fact be syntactic in nature. In order to pursue this issue, a replication of their study was conducted in Dutch. This replication differs from the Waters et al. experiment in several respects:

1. All relative clauses were right embedded to avoid the confounding just discussed;
2. Dutch materials were used;
3. the main clause noun phrases (NPs) of the two-proposition sentences were personal pronouns instead of full NPs to provide a better match in terms of length;
4. all sentences were reversible; and
5. there was no tapping condition.

Each of these differences will be discussed in the following paragraphs.

Only Right-Embedded Structures. The reason for including right-embedded structures only was that in the Waters et al. (1987) study, propositional complexity and center embedding were confounded. Since the goal of the current experiment was to look at the nature and size of the propositional effect, the experiment compared one-proposition sentences, which are right embedded anyhow, with right-embedded two-proposition sentences.

Using Dutch Materials. Second, using Dutch instead of English materials introduced an additional structural complication: in Dutch relative clauses, the verb appears at the end. This had consequences for the comparison of subject- and object-relative structures: due to the position of the verb, Dutch subject and object relatives have the same surface orders, as can be seen in (6a and b):

(6) a. Het is de senator die de reporters belaagt
It is the senator that the reporters attacks
'It is the senator who attacks the reporters'

b. Het is de senator die de reporters belagen
It is the senator that the reporters attack
'It is the senator who the reporters attack'

When Dutch people read or hear a relative clause, they initially pursue the simpler and more frequent subject-relative analysis. If it becomes apparent at the embedded verb that the sentence is an object relative, the sentence must be reanalyzed (Frazier & Flores d'Arcais, 1989; Kaan, 1997). Dutch readers only build the more complex object relative after they have found out that the subject-relative analysis is incorrect (i.e., after they have gardenpathed). The cue for disambiguation is agreement of the subject and the verb: if the verb agrees with the first NP, as in (6a), the sentence is a subject relative. If it agrees with the second NP, as in (6b), the sentence is an object relative.

This difference between Dutch and English has consequences for the strategies and the kind of information readers may use in order to interpret sentences. In English, where subject and object relatives have different surface structures, word order is an important cue for disambiguation, whereas Dutch readers have to rely on a more fine-grained cue, the morphological affixes of the verb and the nouns.

This may make the reader more dependent on a phonological representation. Secondly, interactions only occur if the system is maximally loaded. Therefore, increased syntactic load may improve chances of finding interactions between syntactic variables and phonological memory load.

Using Personal Pronouns. The third difference between the current experiment and Waters et al.'s (1987) was that personal pronouns were used as main-clause subjects in the two-proposition sentences instead of full NPs. This was done in order to provide a better match between one- and two-proposition sentences in terms of sentence length and frequency of the first NP. In Waters et al.'s examples, all one-proposition sentences begin with a pronoun, *it*, and all two-proposition sentences begin with an overt NP, consisting of at least two words. This was primarily necessitated by the center-embedded relative clauses they used. Using pronouns provides a better match in terms of number of words per sentence.

In addition, personal pronouns have a high frequency of usage, just like the dummy pronoun *it*, which occurs in all one-proposition sentences. In the two-proposition sentences, only the pronouns *I* and *you* were used because they are a nice match to the dummy pronoun *it* in terms of referential complexity. According to Gibson (1998), integrations are more demanding when new discourse referents need to be reactivated. Given that speaker and listener are already defined in the discourse setting, the pronouns *I* and *you* are easier to integrate.

Using Reversible Structures. Fourth, reversible constructions were used instead of the nonreversible ones used in the Waters et al. (1987) experiment. In reversible structures, the two participants are equally likely to perform the action of the sentence. For example, in (7a), *the boy is watching the dog*, but dogs can watch boys equally well. In (7b), a nonreversible sentence, the situation is different however: *the boy* is likely to watch *the cake*, but *the cake* cannot watch *the boy:*

(7) a. The boy is watching the dog
b. The boy is watching the cake

Reversible structures were chosen for a number of reasons. First, they are harder to process than nonreversible sentences. Nonreversible structures give Dutch readers hints as to who does what to whom before the verb is encountered (Lamers, 2001; Mak, 2001):

(8) Het is de senator die de interviews geeft
It is the senator that the interviews gives
'It is the senator who gives the interviews'

Since animate nouns are generally better agents than inanimate ones, Dutch readers may already guess the thematic relation between *the senator* and *the interviews* before they have even seen the verb. Reversible sentences do not offer that possibility, and therefore readers have to rely solely on morphological information during syntactic analysis.

One reason for making those constructions harder has to do with Sternberg's (1969) interaction logic. Interaction effects only occur when a system is taxed sufficiently. Therefore, using the harder reversible structures increases the chances of finding an interaction and decreases the chance of a Type II error.

Second, using reversible structures may make participants more likely to address a phonological back-up representation during syntactic processing. Caramazza, Basili, Koller, & Berndt (1981) proposed that the phonological loop is involved in the processing of reversible sentences, but that it is not, or to a lesser degree, active during comprehension of nonreversible sentences.

The Waters et al. (1987) model predicts no difference between reversible and nonreversible sentences with respect to the stages at which the sentences are processed. Both sentence types contain the same linguistic complexity manipulation, relative-clause complexity, which should be dealt with in the first stage of the model.

Despite the advantages of using reversible instead of nonreversible sentences, there is a potential concern with the use of reversible sentences in anomaly-judgment experiments. With nonreversible sentences, one can use θ-role inversions as implausible distractors. In (9), a plausible sentence, (9a), and its θ-role inversion, (9b) are shown:[7]

(9) a. This is the man who is washing the dishes
b. # These are the dishes that are washing the man

The advantage of using θ-role inversions as distractors is that one forces one's participants to completely perform the syntactic analysis of the sentence. Participants can only judge the sentences using syntactic information.

[7] The pound sign (#) is used in this book for sentences that are syntactically correct, but semantically unacceptable.

However, with reversible sentences different distractors must be used. The θ-role inversions of a reversible target, which is plausible, will also yield a plausible sentence. This is shown in (10):

(10) a. This is the man who is kissing the woman
b. This is the woman who is kissing the man

Instead of θ-role inversions, lexical distractors must be used, as shown in (11):

(11) # This is the man who is devouring the woman

The use of lexical distractors may enhance the task strategy that participants judge a sentence on the basis of lexical fit alone and that they fail to syntactically analyze the sentence (Caplan, personal communication, April, 1999). This potential concern must be kept in mind when looking at the data from the current experiment.

Lack of a Tapping Condition. Finally, there was no tapping condition in this experiment. Adding an extra level to the factor Articulatory Suppression would increase the number of conditions by four; in which case many more materials would have to be included without any extra benefit. Including a tapping condition was not considered necessary, since Waters et al. (1987) have already demonstrated that a tapping condition does not differ from processing in a no-load condition.

Method

Subjects. Thirty-two undergraduate students at the University of Groningen served as paid participants. All participants had normal or corrected to normal vision and were native speakers of Dutch. Fourteen participants were male; 18 were female. Their mean age was 20.1 years; their age range was between 17 and 26.

Materials. This experiment orthogonally varied Clause Type, Propositional Complexity, and Articulatory Suppression. Clause Type was varied by contrasting subject- and object-relative clauses within items. In these materials, the object relatives were derived from the subject relatives by changing the number of the embedded verb. In the examples in (6), repeated here as (12a and b), we have already seen that if the embedded verb agrees with the first NP, the embedded clause is subject relative. If it agrees with the second NP, the embedded clause is object relative:

(12) a. Het is de senator die de reporters belaagt
It is the senator that the reporters attacks
'It is the senator who attacks the reporters'

b. Het is de senator die de reporters belagen
It is the senator that the reporters attack
'It is the senator who the reporters attack'

The factor Propositional Complexity was manipulated between items by opposing sentences with one proposition and sentences with two propositions (1P and 2P). This was done in order to manipulate the second stage processing in the Berwick & Weinberg (1984) model. Combination of the two factors yielded the following conditions: 1PSR (cf. 13a), 1POR (cf. 13b), 2PSR, (cf. 13c), and 2POR (cf. 13d):

(13) a Het was de zendeling die de rechercheurs haatte
It was the missionary that the investigators $\text{hated}_{\text{sg.}}$
'It was the missionary who hated the investigators'

b. Het was de zendeling die de rechercheurs haatten
It was the missionary that the investigators $\text{hated}_{\text{pl.}}$
'It was the missionary who the investigators hated'

c. Ik bevrijdde de arts die de architecten verborg
I freed the doctor that the architects $\text{hid}_{\text{sg.}}$
'I freed the doctor who hid the architects'

d. Ik bevrijdde de arts die de architecten verborgen
I freed the doctor that the architects $\text{hid}_{\text{pl.}}$
'I freed the doctor who the architects hid'

There were also two articulatory-suppression conditions: no load and counting from 1 to 6 out loud. All sentence types were presented both in the no load and in the counting condition.

In order to avoid artefacts, the experimental conditions were matched for the number of letters per word position, frequency of usage per word position (taken from the CELEX corpus; cf. Burnage, 1990), sentence plausibility, and reversibility of the relative clause.

Sentence plausibility was rated in a pretest on a scale from 1 *(very bad)* to 5 *(excellent)* by 40 students of the University of Groningen who served as paid subjects; none of them were included in the actual experiment. Reversibility of noun-verb-noun combinations was rated in a different pretest as –1 (*N1 is a better agent*), 0 *(both are equally likely as agents),* or 1 *(N2 is a better agent).* The means for plausibility and reversibility are presented in Table 1.

Table 1
Mean Plausibility and Reversibility per Experimental Condition

Sentence type	Mean plausibility	Mean reversibility
1PSR	3.49	.039
1POR	3.63	.039
2PSR	2.89	.003
2POR	3.00	.003

Note. In a pretest, sentence plausibility was rated on a scale from 1 *(very bad)* to 5 *(excellent)*. Reversibility for noun-verb-noun combinations of all 48 test sentences was rated in a different pretest as –1 *(N1 is a better agent)*, 0 *(both are equally likely as agents)*, or 1 *(N2 is a better agent)*.

Because the object relatives were derived from the subject relatives by changing the number of the verb, subject and object relatives are matched perfectly in terms of word frequency per position, plausibility, and reversibility.

Subject- and object-relative pairs do not have a perfect match with respect to verb length in the embedded clause: plural verbs are longer than singular ones. This length difference is usually one or two letters, as can be seen in (13) above: *haatten (hated)* in (13b) is a plural and is longer than its singular counterpart, *haatte (hated)* in (13a).

To control for this, half of the sentences contained clauses in which the first NP was singular and the second plural; the other half contained clauses with the opposite singular/plural distinction. Thus, the length difference could not favor subject relatives over object-relative clauses or vice versa: the two conditions are well matched for sentence length in number of syllables.

The sentences with one and two propositions were matched as far as possible. Per word position, these sentence types were controlled for in terms of word length and frequency of usage. Further, the sentences were matched in terms of reversibility of the relative clause and plausibility. In all, there were 48 experimental sentence sets divided over 8 conditions; all targets were semantically plausible.

Deriving object relatives from subject relatives implies that the experiment contains pairs of sentences that are almost identical. In order to rule out influences due to the fact that participants had already seen one version of a particular sentence, the materials were divided over two lists. They were divided in such a way that if a particular sentence occurred as a subject relative in list 1, it would occur as an object relative in list 2 and vice versa.

Each list, in its turn was divided in two halves. This was because there were two articulatory-suppression conditions in the experiment: one half was read without articulatory suppression, while the other was read in the counting condition. The order of presentation of the two halves was counterbalanced across participants, as was the order in which the participants had to do the articulatory-suppression conditions.

Dividing the materials up in lists and list halves and varying the order of the articulatory-suppression conditions yielded eight participant groups. The characteristics of these groups are summarized in Table 2. Participants were assigned to lists randomly.

In addition to the targets, there were 96 distractor sentences. Twenty-four of these had the same structures as the targets, but were semantically anomalous. These anomalous fillers were created by picking plausible sentences and replacing the appropriate verbs with anomalous verbs, as in (14):

(14) # It is the drug addict who opens the clergymen

The remaining 72 distractors had all sorts of constructions. Twenty-four of these were semantically implausible; the other 48 were semantically correct.

Table 2
Characteristics of Each Participant Group in Experiment 1.

Group	Participants	List	Order of halves	Task order
1	1, 9, 17, 25	1	A1	CO1
2	2, 10, 18, 26	1	A1	NL1
3	3, 11, 19, 27	1	B1	CO1
4	4, 12, 20, 28	1	B1	NL1
5	5, 13, 21, 29	2	A1	CO1
6	6, 14, 22, 30	2	A1	NL1
7	7, 15, 23, 31	2	B1	CO1
8	8, 16, 24, 32	2	B1	NL1

Note. List assignment and order of list halves and of task condition are given for each participant in Experiment 1. A1 = list half A presented first; B1 = list half B presented first; CO1 = counting condition presented first; NL1 = no-load condition presented first.

The ratio of implausible vs. implausible sentences in the experiment was thus two plausible sentences for every implausible sentence. The original Waters et al. (1987) experiment had a ratio of 1 : 1. The reason for having a ratio of 2 : 1 in the current experiment was the amount of material. Thus, less materials could be used, which lead to shorter durations of the test sessions. This keeps participants more alert. However, it may also introduce additional effects. Participants may be more inclined to consider a sentence correct due to the fact that a *correct* judgment must be made more frequently than an *incorrect* judgment.

Procedure. The study used a whole-sentence anomaly-judgment paradigm as in Waters et al. (1987). This worked as follows: the participant saw an asterisk on the left-hand side of a computer monitor. By pressing the *correct* button, the participant replaced it with a complete sentence. As soon as the participant had read that sentence, (s)he had to decide whether the sentence was semantically plausible or not. Participants received feedback after each judgment. Presentation of the next sentence started when participants pressed the *next sentence* button.

The experiment consisted of two blocks: a no-load block and a counting block. In both blocks, participants had to do the plausibility task, but during the counting block, they also had to count from 1 to 6 out loud. Between the blocks with the different load conditions, participants had a short break.

Results

RT Analysis. RT data from the 32 participants were submitted to standard descriptive and inferential analysis. First, the raw data of the two list halves of each participant were checked for outliers. The lower absolute cut-off was set to 200 ms, under the assumption that it takes at least so much time to process a sentence. Shorter RTs were taken to reflect errors with the response boxes unrelated to sentence processing. The upper absolute cut-off was 9,000 ms.

For each participant, means and standard deviations were computed per list half in order to determine the relative cut-offs. The lower relative cut-off was the mean RT minus 2.5 standard deviation; the upper relative cut-off was the mean RT plus 2.5 standard deviation. Data points that were below the lower cut-offs (both absolute and relative) were replaced by the relative lower cut-off, values that exceeded the higher cut-off points were replaced by the upper relative cut-off.

After outlier detection, means and standard deviations per condition were calculated over subjects and items. The mean RTs of the different experimental conditions are given in Figure 1.[8]

For the subject analysis (*F1*), an analysis of variance (ANOVA) was performed, which used Clause Type (subject vs. object relative), Propositional Complexity (one proposition vs. two propositions), and Articulatory Suppression (no load vs. counting) as within-subject factors and List, List-Half Order, and Task Order as between-subject factors. For the item analysis (*F2*), an ANOVA was performed using Clause Type and Articulatory Suppression as within-item factors and Propositional Complexity and Item Group as between-item factors.

---

[8] In order to facilitate comparison of these data with the RTs of Experiments 2 to 5b, the scale of each figure was set between 2,000 and 6,000 ms.

In this study, effects are considered significant when they are significant in both the *F1* and the *F2* analysis. In addition, effects that were significant in one and approaching significance in the other analysis (with $p < .1$) will also be mentioned.

The ANOVA showed a main effect of Propositional Complexity, *F1* (1, 30) = 11.49, $p < .05$; *F2* (1, 46) = 12.40, $p < .001$: one-proposition sentences were read more quickly than two-proposition sentences. There was also a main effect of Articulatory Suppression, *F1* (1, 30) = 6.96, $p < .05$; *F2* (1, 46) = 12.30, $p < .05$: sentences were read more quickly in the counting condition than in the no-load condition. There was no main effect of Clause Type.

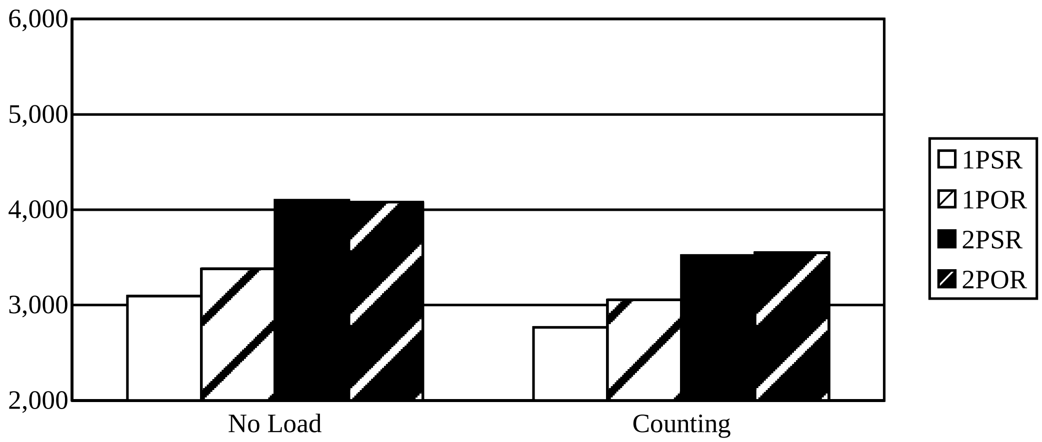


Figure 1. Mean RTs in ms for one-proposition subject relatives (1PSR), one-proposition object relatives (1POR), two-proposition subject relatives (2PSR), and two-proposition object relatives (2POR) in the no load (left) and counting conditions (right).

The interaction of Clause Type x Propositional Complexity was significant in the *F1* analysis and approached significance in the *F2* analysis, *F1* (1, 30) = 4.79, $p < .05$; *F2* (1, 46) = 3.27, $p = .08$. Simple effects of Clause Type were present in the one-, *F1* (1, 30) = 1190, $p < .01$; *F2* (1, 20) = 13.54, $p < .01$, but not in the two-proposition sentences, *F1* (1, 30) = 1.46, $p = .24$; *F2* (1, 20) = 2.11, $p = .16$. No other significant interactions were found.

Error Analysis. Error data were submitted to the same analyses as the RT data. Mean error rates for each condition are given in Figure 2.[9] The ANOVA showed a significant main effect of Propositional Complexity, *F1* (1, 27) = 23.47, $p < .001$; *F2* (1, 40) = 13.07, $p < .001$: two-proposition sentences yielded more errors than one-

[9] In order to facilitate comparison of these data with the error proportions of Experiments 2 to 5b, the scale of each figure was set between the error proportions of 0 and .50.

proposition sentences. The error analysis produced no other significant main effects or interactions.

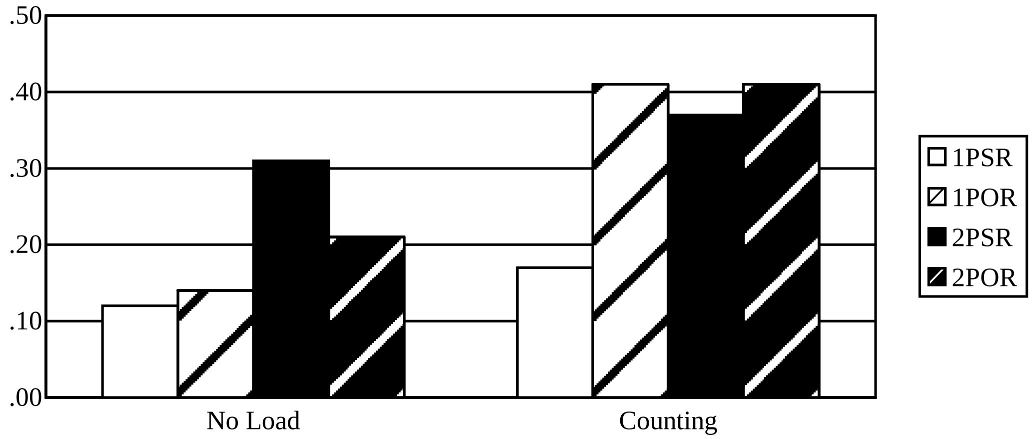


Figure 2. Mean error proportions for one-proposition subject relatives (1PSR), one-proposition object relatives (1POR), two-proposition subject relatives (2PSR), and two-proposition object relatives (2POR) in the no load (left) and counting conditions (right).

Discussion

This experiment was based on an experiment by Waters et al. (1987); it manipulated Clause Type (subject vs. object relatives), Propositional Complexity (one vs. two propositions), and Articulatory Suppression (no load vs. counting out loud). Some of the two-proposition sentences Waters et al. used contained center-embedded structures, whereas all one-proposition sentences were right embedded. Their use of center-embedded structures may have caused (part of) the effects of propositional complexity. Therefore, the current experiment used right-embedded structures only.

Examination of the Main Effects. There was a main effect of Propositional Complexity with higher RTs and error rates for two-action sentences. This is probably due to the fact that in two-action sentences, more propositions must be linked and interpreted.

This propositional effect is more interpretible than the effect found by Waters et al. (1987). The current study used personal pronouns as main-clause subjects of the two-action sentences, whereas Waters et al. used full NPs. The one-action sentences in Waters et al.'s experiment had the dummy pronoun *it* as subject. Comparing full NPs with pronouns introduces a length confound. This length difference was not present in the current experiment because of the use of the personal pronouns *ik (I)* and *jij (you)* and cannot explain the effect.

Additionally, the pronouns *I* and *you* and the dummy pronoun *it* are more comparable in terms of referential complexity. Gibson (1998) proposes that syntactic integrations are more demanding when new discourse referents need to be reactivated. Since each discourse setting involves a speaker/writer (or messenger) and a listener/reader (or recipient), the pronouns *I* and *you* are already specified and therefore easier to integrate. The dummy pronoun *it* is not a new discourse referent either, because it has no θ-role.

The most interesting main effect in the RT analysis was one of Articulatory Suppression, with faster RTs for the articulatory-suppression condition. This is contrary to what one would expect and not in line with the effect found by Waters et al. (1987): in their study, readers were slower under articulatory suppression.

This reversed load effect seems to suggest that readers go more quickly when confronted with the degeneration of the memory trace in the articulatory-suppression conditions in Dutch. This difference between Dutch and English readers may be related to the structures of the two languages. As already pointed out, the verb occurs in final position in Dutch embedded clauses. This means that in both subject- and object-relative clauses, θ-role assignment and integration of the different constituents must be delayed until the end of the clause. If addressing a phonological back-up is necessary during this process, speeding up could prevent problems.

Evidence for a Response Bias. One objection against such an explanation is that one would expect this speeding up to lead to decreased accuracy: under articulatory suppression, participants read more quickly, but they are likely to do so at the expense of accuracy. This is not what we observe in the analyses: the articulatory-suppression conditions do not produce significantly higher error proportions than the no-load conditions.

However, a closer look at the data is necessary given that the experiment was set up in such a way that there might have been a response bias. Each sentence was preceded by a warning signal that the participant had to react to by pressing the *correct* button. After that, the sentence appeared and participants judged its semantic acceptability. Thus, the button sequence for a correct sentence would be: *next sentence, correct,* and *correct* (i.e., sticking to the *correct* button), but the sequence for an incorrect sentence would be *next sentence, correct,* and *incorrect* (i.e., switching buttons).

A second reason for suspecting that a response bias may be present in the current experiment is that two plausible sentences appeared for every implausible sentence.

Therefore, participants may be more inclined to judge a sentence as correct than as incorrect, especially in the more demanding articulatory-suppression condition.

If this is the case, one would expect higher error proportions for the semantically implausible sentences under articulatory suppression than in the no-load conditions. Since all target sentences were semantically acceptable, this possibility could only be explored by looking at the responses to filler sentences. A Wilcoxon Matched Pairs Signed Rank Test was carried out in order to test this. It compared the proportion of errors for the implausible sentences in the no-load condition with the error proportion for the implausible sentences of the articulatory-suppression condition.

There was a significantly higher proportion of errors for filler sentences in the articulatory-suppression condition than in the no-load condition ($Z = -2.29$, $p < .05$). This seems to substantiate the suggestion made above that participants want to diminish processing demands by trying to reach the clause-final embedded verb as soon as possible when under articulatory suppression.

This finding provides some indication of the strategies readers use when faced with the articulatory-suppression task, but it does not reveal anything about the role of the phonological loop in sentence comprehension. The lack of an interaction of Propositional Complexity x Articulatory Suppression in Experiment 1 suggests that the phonological loop has nothing to do with propositional analysis. Instead, it seems to play some role in the processing of center-embedded structures, but based on these data alone, further specification of that role remains speculative.

A Non-additive Interaction of Clause Type x Propositional Complexity. A third effect in the RT analysis was a two-way interaction of Clause Type x Propositional Complexity. Subject relatives were read more quickly than object relatives in the one-proposition condition only; no effects of Clause Type were present in the two-proposition sentences. The reason for faster RTs for subject relatives is probably that Dutch object relatives are initially processed as subject relatives, since subject relatives are structurally simpler. Most important is that this interaction is non-additive in nature; therefore, it is not possible to say anything about the independence of the two factors by applying Sternberg's (1969) interaction logic.

We must keep in mind that a potential drawback of Experiment 1 was the use of reversible sentences. Because the plausibility of the sentences could be judged on the basis of the lexical content of the words in the sentence, participants may not have carried out the syntactic analysis. This possibility cannot be ruled out entirely on the basis of the outcome of Experiment 1. The fact that effects of Clause Type were obtained for the one-proposition sentences suggests that participants syntactically analyzed the one-proposition sentences. Since no such effect was seen for two-proposition sentences, they may have failed to analyze these syntactically. However, they did at least some syntactic analysis during the task; therefore, a no-syntax strategy in the participants of Experiment 1 does not seem likely.

Implications for the Waters et al. Model. The first striking difference between the current experiment and Waters et al. (1987, exp. 3) is that the current experiment failed to replicate the superadditive interaction of Propositional Complexity x Articulatory Suppression, *F1* (1, 30) = .32, $p = .58$; *F2* (1,46) = .39, $p = .54$. This seems to suggest that Waters et al.'s interaction is not an interaction of Propositional Complexity x Articulatory Suppression, but an interaction of Embedding Type x Articulatory Suppression. This implies that their assumption about a role for the phonological loop in propositional processing may need to be reconsidered since the phonological loop seems to be addressed during syntactic rather than propositional processing.

Note that although the current experiment failed to replicate the interaction of Propositional Complexity x Articulatory Suppression, it did replicate the main effect of Propositional Complexity. This suggests that Waters et al.'s (1987) interaction of Propositional Complexity x Articulatory Suppression is more readily explained as an embedding effect but that the main effect of Propositional Complexity is not (or not only) to be explained by this confound.

A second observation with respect to the Waters et al. (1987) model is that Propositional Complexity and Clause Type produced a nonadditive interaction. This means that it is impossible to apply Sternberg's (1969) interaction logic. Therefore, one cannot decide between a single- and a separate-resource hypothesis based on these data: at present, this issue remains open for further exploration.

The fact that I argue that Waters et al.'s (1987) interaction is due to using center-embedded structures is based the fact that the current experiment did not find this interaction. One is not allowed to draw strong conclusions based on a null result. In order to give more substance to this claim, one will have to find an interaction of Embedding Type x Articulatory Suppression. Therefore, a follow-up experiment is needed. That experiment is discussed in the next section.

Experiment 2

Introduction

Experiment 2 was conducted in order to investigate two issues. Primarily, I wanted to test the idea that articulatory rehearsal is deployed during the processing of center-embedded structures. Second, the experiment was conducted to test whether center embedding, which increases syntactic complexity, interacted with another type of syntactic complexity, relative-clause complexity. This issue is relevant with respect to the assumption of a two-stage parsing model like Berwick & Weinberg's (1984). Both types of complexity are syntactic in nature and may therefore be expected to interact. Therefore, Clause Type was also included as a factor.

The experiment compared the effects of no load vs. counting, right- vs. center-embedded clauses, and subject- vs. object-relative clauses. All sentences were reversible, in order to facilitate comparison between this and the previous experiment.

Because the use of reversible materials potentially enhances a strategy in which participants fail to do the syntactic analysis of the sentences, the syntactic effects will be closely monitored and discussed.

Method

Subjects. Thirty-two undergraduate students of the University of Groningen served as paid participants. Inclusion criteria were the same as in Experiment 1. Eight participants were male; 24 were female. Their mean age was 21.7 years and the age range was between 18 and 35. No participants that had participated in Experiment 1 were included.

Materials. This experiment used an orthogonal design. It varied the following within subject and item factors: Clause Type (subject vs. object relatives; SR vs. OR), Embedding Type (right- vs. center-embedded structures; RE vs. CE), and Articulatory Suppression (no load vs. counting from 1 to 6).

Materials were constructed in such a way that both the noun-verb-noun combination in the main clause and the one in the embedded clause were reversible. The noun-verb-noun combination in the main clause had to be reversible for the construction of center embeddings out of right-embeddings.

As in Experiment 1, object-relative clauses were derived from their subject-relative counterparts by changing the number of the embedded verb. All embedded clauses had one singular and one plural NP. If the verb agreed with the first NP of the embedded clause, the clause was subject relative; if it agreed with the clause's second NP, the clause was object relative. As in Experiment 1, half of the embedded clauses had first NPs that were singular; the other half had plural first NPs. This was done in order to control for sentence length.

How the center embeddings were derived can be seen below. In (15a and b), the NP without relative clause appears first; in (15c and d) order of the NPs is reversed. Combining the factors Clause Type and Embedding Type yielded the following experimental conditions: RESR (cf. 15a), REOR (cf. 15b), CESR (cf. 15c), and CEOR (cf. 15d):

(15) a. De studenten prijzen de fotograaf die de uitgevers accepteert
The students praise the photographer that the publishers accepts
'The students praise the photographer who accepts the publishers'

b. De studenten prijzen de fotograaf die de uitgevers accepteren
The students praise the photographer that the publishers accept
'The students praise the photographer who the publishers accept'

c. De fotograaf die de uitgevers accepteert, prijst de studenten
The photographer that the publishers accepts, praises the students
'The photographer who accepts the publishers praises the students'

d. De fotograaf die de uitgevers accepteren, prijst de studenten
The photographer that the publishers accept, praises the students
'The photographer who the publishers accept praises the students'

All main clauses had one plural and one singular NP. This was done because, although it is a marked construction, the main-clause object can grammatically appear in first position as a topic. This occurs when the object is a focal NP, as can be seen in the second sentence in (16):

(16) Wie straft de leraar precies?
Who punishes the teacher exactly?
'Who does the teacher punish exactly?'

De leerlingen straft de leraar
The pupils punishes the teacher
'The teacher punishes the pupils'

This focal reading is normally not preferred and the singular/plural distinction was made in this experiment in order to rule such a reading out entirely.

The experimental conditions were matched for the number of letters per word position, frequency of usage per word position (taken from the CELEX corpus; cf. Burnage, 1990), sentence plausibility, reversibility of the main clause, and reversibility of the embedded clause. Sentence plausibility was rated in a questionnaire on a scale from 1 to 5 (resp. *very bad* and *excellent*). The noun-verb-noun combinations were rated for reversibility in a different pretest, as –1 *(N1 is preferred as agent over N2),* 1 *(N2 is preferred as agent),* or 0 *(no difference),* as in Experiment 1. The means for plausibility and reversibility are presented in Table 3.

Table 3
Mean Plausibility and Reversibility per Experimental Condition

Sentence type	Mean plausibility	Mean reversibility
RESR	2.93	.13
REOR	2.44	-.13
CESR	3.11	.13
CEOR	2.42	-.13

Note. In a pretest, sentence plausibility was rated on a scale from 1 *(very bad)* to 5 *(excellent)*. Reversibility for noun-verb-noun combinations was rated in a different pretest as –1 *(N1 is a better agent),* 0 *(both are equally likely as agents),* or 1 *(N2 is a better agent).*

Because plural verbs are longer in terms of number of letters, materials were constructed in such a way that half of the main clauses contained singular and half of them plural verbs. For the embedded clauses, the same approach was taken.

Ninety-six distractor sentences were also included. There were 24 semantically anomalous distractors with the same syntactic structure as the targets. Anomalies were created by replacing a plausible verb with one that did not fit in the context of the sentence. The remaining 72 fillers had all sorts of constructions; 48 of those were semantically plausible, but the remaining 24 were not.

Because the sentences in the different conditions were obtained by changing the order of the main-clause NPs or the number of the embedded verb, each word combination occurred in four conditions. In order to avoid repetition artefacts, sentences were assigned to four lists. This was done in such a way that no participant would see more than one version of the same combination.

As in Experiment 1, lists were divided in two halves, one for the no-load condition and one for the counting condition. Order of presentation of the list halves and of the articulatory-suppression conditions was counterbalanced across participants. Assignment to lists and counterbalancing of list halves and articulatory-suppression conditions yielded 16 participant groups. Participants were assigned to groups randomly. An overview of group characteristics is given in Table 4.

Table 4
Characteristics of Each Participant Group in Experiment 2.

Group	Participants	List	Order of halves	Task order
1	1, 17	1	A1	CO1
2	2, 18	1	A1	NL1
3	3, 19	1	B1	CO1
4	4, 20	1	B1	NL1
5	5, 21	2	A1	CO1
6	6, 22	2	A1	NL1
7	7, 23	2	B1	CO1
8	8, 24	2	B1	NL1
9	9, 25	3	A1	CO1
10	10, 26	3	A1	NL1
11	11, 27	3	B1	CO1
12	12, 28	3	B1	NL1
13	13, 29	4	A1	CO1
14	14, 30	4	A1	NL1
15	15, 31	4	B1	CO1
16	16, 32	4	B1	NL1

Note. List assignment, and order of list halves and of task condition are given for each participant in Experiment 2. A1 = list half A presented first; B1 = list half B presented first; CO1 = counting condition presented first; NL1 = no-load condition presented first.

Procedure. The procedure was identical to the procedure followed in Experiment 1.

Results

RT Analysis. The RTs were checked for outliers and analyzed using the same procedure as in Experiment 1. Mean RTs per condition are given in Figure 3.[10]

For the subject analysis (*F1*), an analysis of variance (ANOVA) was performed with the following within-subject factors: Clause Type (subject vs. object relative), Embedding Type (right vs. center embedded), and Articulatory Suppression (no load vs. counting). List, List-Half Order, and Task Order were between-subject factors. For the item analysis (*F2*), an ANOVA was performed which had Clause Type, Embedding Type, and Articulatory Suppression as within-item factors and Item Group as a between-item factor.

The ANOVA showed main effects of Clause Type, $F1\ (1, 16) = 6.22$, $p < .05$; $F2\ (1, 36) = 4.94$, $p < .05$, and of Articulatory Suppression, $F1\ (1, 16) = 5.32$, $p < .05$; *F2*

[10] In order to facilitate comparison of these data with the RTs of Experiments 1, 3, 4, 5a, and 5b, the scale of each figure was set between 2,000 and 6,000 ms.

(1, 36) = 17.11, $p < .05$. Subject-relative clauses were read more quickly than object-relative clauses and as in Experiment 1, readers were faster in the counting condition relative to the no-load condition. No main effect of Embedding Type was found.

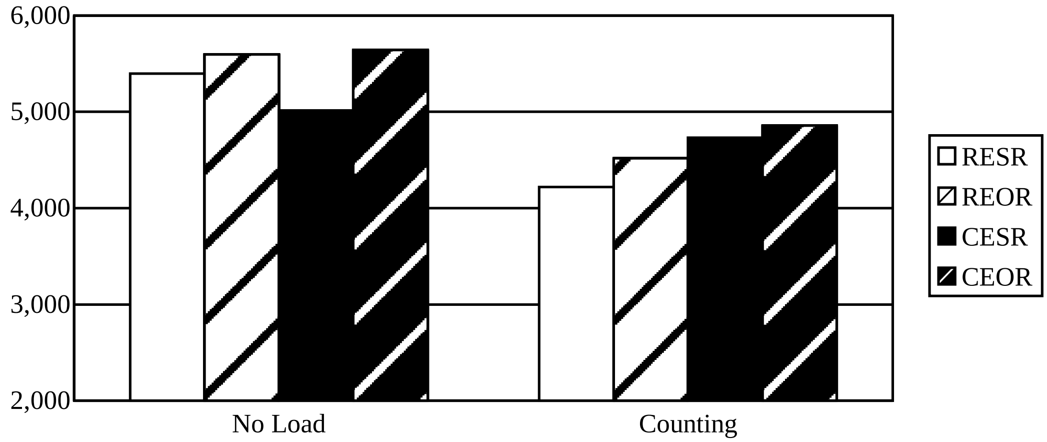


Figure 3. Mean RTs in ms for right-embedded subject relatives (RESR), right-embedded object relatives (REOR), center-embedded subject relatives (CESR), and center-embedded object relatives (CEOR) in the no load (left) and counting conditions (right).

The two-way interaction of Embedding Type x Articulatory Suppression was significant, *F1* (1, 16) = 5.87, $p < .05$; *F2* (1, 36) = 4.80, $p < .05$. It turned out that no simple effects of Embedding Type were found in the no-load condition analysis, *F1* (1, 16) = .91, $p = .36$; *F2* (1, 38) = .44, $p = .51$. In the counting condition however, center-embedded structures were read more slowly than right-embedded clauses. This difference was significant in the subject analysis and marginally significant in the item analysis, *F1* (1, 16) = 7.19, $p < .05$; *F2* (1, 38) = 3.94, $p = .07$. No other interactions were found.

Error Analysis. The error data were also submitted to the same statistical analysis as the RT data. Mean error proportions per condition are given in Figure 4.[11] The ANOVA showed a main effect of Clause Type: subject relatives produced higher error rates than object relatives, *F1* (1, 16) = 4.80, $p < .05$; *F2* (1, 40) = 5.17, $p < .05$. This finding should be interpreted in the context of an interaction of Clause Type x Articulatory Suppression, *F1* (1, 16) = 9.47, $p < .05$; *F2* (1, 40) = 5.74, $p < .05$. In the counting condition only, subject relatives produced significantly more errors than object relatives, *F1* (1, 16) = 10.65, $p < .01$; *F2* (1, 44) = 9.45, $p < .01$; the no-load

[11] In order to facilitate comparison of these data with the error proportions of Experiments 1, 3, 4, 5a, and 5b, the scale of each figure was set between the error proportions of 0 and .50.

condition did not show any effect of Clause Type, $F1\ (1, 16) = .55$, $p = .49$; $F2\ (1, 44) = 9.45$, $p < .01$. No other effects were found.

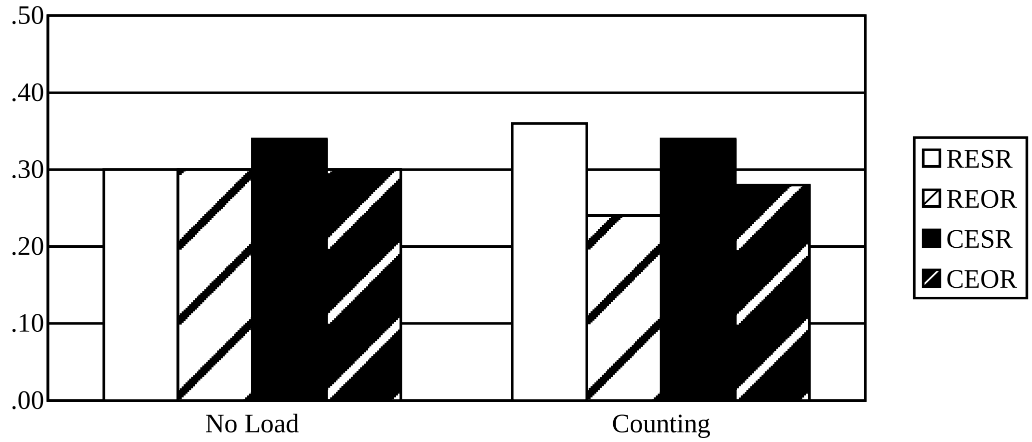


Figure 4. Mean error proportions for right-embedded subject relatives (RESR), right-embedded object relatives (REOR), center-embedded subject relatives (CESR), and center-embedded object relatives (CEOR) in the no load (left) and counting conditions (right).

As in Experiment 1, a Wilcoxon Matched Pairs Signed Rank Test was carried out comparing the proportion of errors for the implausible sentences in the no-load condition with the error proportion for the implausible sentences of the articulatory-suppression condition. There was a higher error proportion for filler sentences in the articulatory-suppression condition than in the no-load condition. This difference was marginally significant ($Z = -1.92$, $p = .055$).

Discussion

The experiment replicated the reversed load effect of Articulatory Suppression found in Experiment 1: participants were faster under articulatory suppression than in the no-load condition. This finding seems to substantiate the idea that in Dutch, participants react differently to articulatory suppression than in English. I suggest that in Dutch, readers may reach the final verb of the embedded clause as quickly as possible when under articulatory suppression to prevent effects of degeneration of the phonological representation inducing negative effects on processing.

A second effect that could be observed in the RT analysis is a two-way interaction of Embedding Type x Articulatory Suppression: center-embedded structures are read more slowly then right-embedded clauses under articulatory suppression only. This finding suggests that during the processing of center-embedded structures, the phonological loop is addressed. It also supports the hypothesis formulated in the discussion subsection of Experiment 1 that Waters et al.'s (1987) interaction of Propositional Complexity x Articulatory Suppression is in fact an interaction of Embedding Type x Articulatory Suppression.

Third, the RT analysis produced a main effect of Clause Type: subject relatives were responded to more quickly than object relatives. This is probably because Dutch readers initially tend to pursue the easier subject-relative reading of the sentence for both subject- and object-relative sentences. In the case of object relatives, they find out that they are wrong at the end of the embedded clause, when verb information becomes available. Then they have to reanalyze the sentence, which costs extra processing time (Frazier & Flores d'Arcais, 1989).

In the error rates, there was a main effect of Clause Type as well, this time with subject relatives yielding higher error proportions than object relatives. This finding should be interpreted in the context of a two-way interaction of Clause Type x Articulatory Suppression, with higher error proportions for the subject relatives in the articulatory-suppression condition only.

This interaction suggests that subjects may address a back-up representation stored in the phonological loop to check the agreement features of the subject and verb in the embedded clause. The poorer accuracy in subject relatives may reflect the fact that for a successful check of the agreement features, readers need a longer stretch of the back-up representation in subject than in object relatives, as shown in (17a and b):

(17) a. De studenten prijzen de *fotograaf* die de uitgevers *accepteert*
The students praise the photographer that the publishers accepts
'The students praise the photographer who accepts the publishers'

b. De studenten prijzen de fotograaf die de *uitgevers accepteren*
The students praise the photographer that the publishers accept
'The students praise the photographer who the publishers accept'

The fact that this effect shows up in the error rates only may indicate that subjects do not carry out this checking procedure in a conscious way.

The current experiment yielded pleasantly large syntactic complexity effects. Therefore, the assumption that syntactic analysis is not carried out in anomaly judgment tasks because decisions can be made on the basis of lexical fit is not supported by the findings of the current experiment. Thus, we can remove one possible explanation for the lack of interactions in Experiment 1.

General Discussion

In this section, the implications of Experiments 1 and 2 for the single vs. separate resource discussion will be evaluated. More importantly, an alternative specification of the use of WM resources in sentence processing will be proposed. First, the implications for the different WM theories will be dealt with; then the alternative model will be explained.

The Single vs. Separate Resource Issue

As already mentioned in the introduction, there are two opposing views with respect to this issue. Just & Carpenter (1992) advocate a single-resource theory, which predicts that all manipulations should interact with each other. However, Waters et al. (1987) and Caplan & Waters (1999) propose a partition of resources predicting the following:

1. Syntactic manipulations should interact with each other;
2. postsyntactic manipulations and Articulatory Suppression should interact with each other; and
3. syntactic manipulations should interact with neither postsyntactic processing nor Articulatory Suppression.

Table 5
Superadditive Interactions in Experiments 1 and 2 and Theoretical Predictions

Effect	Embedding Type	Clause Type	Articulatory Suppression
	Effects found in Experiments 1 and 2		
Propositional Complexity	not tested	uncertain	absent
Embedding Type		absent	present
Clause Type			present
	Interactions predicted by Just & Carpenter (1992)		
Propositional Complexity	yes	yes	yes
Embedding Type		yes	yes[a]
Clause Type			yes[a]
	Interactions predicted by Caplan & Waters (1999)		
Propositional Complexity	no	no	yes
Embedding Type		yes	no
Clause Type			no

Note. This table gives an overview of all possible interactions, along with the predictions of the different theories. *Present* means that a superadditive interaction was found in the current study; *absent* that it was not. *Uncertain* means that there were nonadditive interactions. *Not tested* means that a particular combination did not occur. *Yes* indicates that an interaction is predicted by a theory; *no* indicates that it is not.
[a]predictions that are born out.

Table 5 provides an overview of effects found in Experiments 1 and 2, along with the predictions made by the two theories.

Comparing the effects found with the predictions made by the single-resource theory by Just & Carpenter (1992) suggests that the assumption of a single resource pool for language processing is too strong. The single-resource theory correctly predicts two-way interactions of Embedding Type x Articulatory Suppression and of Clause Type x Articulatory Suppression. However, it incorrectly predicts two-way interactions of Propositional Complexity x Articulatory Suppression and of Clause Type x Embedding Type. The lack of the latter two interactions argues for a partition of WM resources along some line or other and against a single-resource hypothesis.

However, the model proposed by Caplan & Waters (1999) is not in line with the current data either. They predict that Propositional Complexity should interact with Articulatory Suppression, which is not the case.[12]

Furthermore, they predict that the following interactions should not occur: Clause Type x Articulatory Suppression and Embedding Type x Articulatory Suppression. These two interactions were present in the data of the current study. These three findings suggest that Waters et al.'s (1987) specification of the role of the phonological loop in sentence processing needs revision.

Finally, they predict an interaction of Clause Type x Embedding Type. Within the Berwick & Weinberg (1984) parser, which Waters et al. (1987) adopted, the matrix subject of a center-embedded sentence is kept active in a pushdown stack of the syntactic processing component. That component is at the same time dealing with the relative-clause complexity. Therefore, the two factors should produce an interaction, which is not found in the data.

About the Nature of the Phonological Back-Up Representation

At present, the nature of the phonological back-up representation is not entirely clear. On the basis of the data from Experiment 1, three hypotheses about that nature can be formulated:

---

[12] Remember that Waters et al. (1987) assumed that a phonological back-up representation is addressed during propositional analysis based on a two-way interaction of Articulatory Suppression x Propositional Complexity. However, their two-proposition sentences were confounded because they contained center-embedded structures. All interactions with Articulatory Suppression found in the current study, which did not have this confound, interact at the syntactic level instead of the propositional level.

1. The phonological back-up is used for the rehearsal of words that cannot be integrated immediately (cf. Wanner & Maratsos's, 1978 HOLD hypothesis);
2. the phonological back-up is addressed during long-distance θ-role assignment (cf. Waters et al., 1987); and
3. the phonological back-up is addressed during long-distance subject-verb integration (cf. Gibson, 1998).

The first hypothesis can be ruled out immediately on the grounds that Clause Type interacted with Articulatory Suppression. If unintegrated material were stored in the back-up, one would not expect an interaction of Clause Type x Articulatory Suppression, since subject- and object-relative clauses have the same amount of unintegrated material before the verb is encountered.

The second and the third hypothesis are very similar in the sense that they both predict interactions with articulatory suppression when subject and verb of an embedded clause must be integrated over a long distance. This is because subject-verb integration and θ-role assignment occur at the same position. At present, one cannot distinguish between the two hypotheses; this issue will be re-addressed in Chapter 5.

Toward a New Model of WM and Sentence Comprehension

As an alternative to the proposals discussed above, I will present a separate-resource model that shares a number of features with Caplan & Waters's (1999) model. The two proposals share the assumption of two processing stages and of a role for a phonological back-up in sentence comprehension. The specification of the back-up's role and its interactions with the two modules differs between the two theories however.

It is proposed that in the syntactic processor, both storage of unintegrated phrases and integration of these phrases occur. Integration contributes most to syntactic complexity effects and during this integration, a phonological back-up is addressed for the checking of agreement features.[13]

First, the assumption of two processing stages will be motivated. The second part will discuss the processes occurring in the syntactic processor. This discussion will focus on the contribution of integration cost and storage demands to the size of syntactic complexity effects. Next, the function of the phonological back-up will be discussed.

[13] The back-up may be used for other purposes as well, but at least for this purpose.

Two Processing Stages. Like the Berwick & Weinberg (1984) model, the current theory assumes two processing stages, one for syntactic processing and one for propositional processing. There are a number of reasons for positing this dichotomy. In Experiment 1, there was a nonadditive interaction between Clause Type and Propositional Complexity; therefore, these data can not be used as conclusive evidence either in favor or against a dichotomy. However, Waters et al. (1987) found that Clause Type, which is dealt with at the syntactic stage, was independent from Propositional Complexity. The independence of the two factors has also been reported for Dutch by Withaar (2001a).

Additional evidence for the assumption of two processing stages stems from the difference between Experiments 1 and 2 in this chapter. Experiment 1 showed that Propositional Complexity is not dealt with by addressing the phonological loop, whereas Experiment 2 produced a two-way interaction of Embedding Type, a first stage-syntactic factor, with Articulatory Suppression. This interaction suggests the involvement of the phonological loop during (some aspects of) syntactic processing.

Integration Cost or Storage Demands? It is interesting to look at what the independence of these two complexities tells us about the way the syntactic processor works. If maintaining structure in the syntactic processor is the factor that contributes most to the complexity of center embeddings, one would expect an interaction between the Clause Type and Embedding Type.[14] The lack of this interaction suggests that maintaining the matrix subject is not the major contributing factor to the complexity effect, but that integration plays a more prominent role.

The fact that maintaining an unintegrated matrix subject in the center-embedded conditions does not produce a large complexity effect is in line with Gibson's (1998) syntactic prediction locality theory (SPLT). In this theory, sentence complexity effects are accounted for in terms of the number of active predictions in a sentence and the number of new discourse referents over which a particular integration takes place. Gibson also proposes that the prediction of a matrix verb is associated with a 0 prediction cost. Thus, there is no difference in terms of the number of active predictions between right- and center-embedded clauses in Gibson's theory.

The complexity difference between right- and center-embedded constructions is accounted for in SPLT in terms of integration difficulty:

(18) a. The students praise the photographer who accepts the publishers
b. The photographer who accepts the publishers praises the students

[14] The idea that maintenance of structure contributes significantly to the syntactic complexity of center-embedded structures is in line with Berwick & Weinberg's (1984) idea of *active node stacks* (cf. p. 15).

In the right-embedded sentence (18a), the integration of the matrix subject *the students* and the matrix verb *praise* is an integration over two new discourse referents (*students* and *praise*). In the center-embedded version, sentence (18b), integration of the matrix subject *the photographer* and the matrix verb *praises* is an integration over four active discourse referents (*photographer, accept, publishers,* and *praise*).

Furthermore, the interaction of Clause Type x Articulatory Suppression also suggests that integration largely contributes to the syntactic complexity effects in the current study. Subject relatives were responded to more quickly than object-relative clauses, but yielded higher error proportions. As can be seen below, a larger part of the back-up is needed for subject relatives than for object relatives. The distance between *fotograaf (photographer)* and *accepteert (accepts)* in (19a) is larger than the distance between *uitgevers (publishers)* and *accepteren (accept)* in (19b):

(19) a. De studenten prijzen de fotograaf die de uitgevers accepteert
The students praise the photographer that the publishers accepts
'The students praise the photographer who accepts the publishers'

b. De studenten prijzen de fotograaf die de uitgevers accepteren
The students praise the photographer that the publishers accept
'The students praise the photographer who the publishers accept'

Although it may be that integration processes cause most of the syntactic complexity effects found in the data, there are reasons not to reject the possibility that storage demands may play a role too. If storage of unintegrated structure were the major contributor to the syntactic effects, one would expect an interaction of Clause Type x Embedding Type in Experiment 2. This is not found in that experiment, but it is important to note that the influence of Embedding Type was too small to produce a main effect. The fact that the Embedding Type effect itself is rather small is important with respect to Sternberg's (1969) interaction logic. One may argue that there was no interaction of Clause Type x Embedding Type, because combining the two effects was not taxing the processing system enough to produce an interaction.

Other evidence for active storage of unintegrated material stems from a PET study by Stowe, Withaar, Wijers, Broere, & Paans (2002). In that study, Stowe et al. compared four different experimental conditions:

1. Word lists (content and function words presented in such a way that no two consecutive words could form a phrase);
2. simple sentences (declarative sentences without embeddings);
3. syntactically complex sentences (structures containing list-like sequences, such as center embeddings); and

4. syntactically most complex sentences (between-category ambiguities in which two options are assumed to be processed simultaneously [Frazier & Rayner, 1987]).

They reported a significant activation in the left inferior frontal gyrus during the processing of these materials. This activation tested whether there was a brain area that was involved in the storage of both lexical words and phrase structure. The logic was that if there is such an area, it should become more active when either the number of phrases or the number of words to be stored increases.

In that case, simple sentences should be easiest, since there are only few phrases and no unintegrated words to be stored. Word lists should be more complex since they cannot be stored as phrases and should be stored individually. Syntactically complex sentences should be comparably difficult, but in this case, it is not the number of words that is responsible for the complexity, but the number of phrases to be stored. The ambiguous sentences, which require simultaneous storage of two phrase-structure representations, should be most active. This data pattern seems to suggest that the left inferior frontal gyrus may be the locus where syntactic processes such as mentioned in the current study are carried out.

In summary, the syntactic processor seems to carry out both the storage of unintegrated elements and the integration of these elements. It seems that in the current study, integration causes most of the syntactic effects, especially in interaction with the phonological back-up representation.

The Phonological Back-up: Function. Concluding that a phonological back-up representation is used during integration processes does not suffice. It is necessary to specify the nature of the use of this back-up representation. Since all sentences in the experiment were reversible and Dutch subject- and object-relative clauses have similar surface order, syntactic analysis hinges on correct processing of subject-verb agreement. It seems likely that during the integration of nouns and verbs, the back-up is addressed for the checking of number information of the subject and the verb in the sentence.

The difficulty of this checking procedure seems to be largely determined by the distance between the subject and the verb. Center-embedded sentences are read more slowly under articulatory suppression than right-embedded sentences. This finding was found both for Dutch (Experiment 2) and for English (Waters et al., 1987, exp. 3): the findings of Experiment 1 and 2 in this chapter suggest that the super-additive interaction of Propositional Complexity x Articulatory Suppression found by Waters et al. is in fact an interaction of Embedding Type x Articulatory Suppression.

These findings suggest that during the integration of a subject and a verb, the phonological back-up representation is addressed. This representation may be consulted in order to check the agreement information of both words. (Note that subjects and verbs must agree in number and person in both English and Dutch.)

This checking procedure seems to be particularly demanding when (a) a large stretch of the back-up must be consulted for this feature check (i.e., when subject and verb are far apart) and (b) the back-up is degraded as a result of articulatory suppression. Remember: center-embedded sentences, which require long-distance integration of the matrix subjects and verbs, were read more slowly than right-embedded sentences, which do not. This effect was present under articulatory suppression only.

A second phenomenon that indicates problems with long-distance subject-verb feature checking under articulatory suppression is that Dutch subject-relative clauses yielded higher error proportions than object-relatives in the articulatory-suppression condition of Experiment 2. The checking of agreement features requires a shorter stretch of the back-up representation for Dutch object relatives, where subject and verb are adjacent, than for Dutch subject relatives, where they are not.

The fact that readers of Dutch seem to suffer from degradation of the phonological back-up under articulatory suppression can be concluded from the fact that performance on implausible sentences deteriorated under articulatory suppression. Given that this is so, we must interpret the faster RTs for the articulatory-suppression conditions than for the no-load conditions as an attempt to avoid the negative consequences of articulatory suppression.

These findings also suggests that participants need the back-up representation for feature checking. They must speed up if they want the back-up to be available at the end of the sentence, because Dutch embedded clauses are verb final and the back-up may decay more rapidly under articulatory suppression.

In English, where embedded clauses do not frequently require such long-distance integrations, subjects seem to slow down under articulatory suppression because of the extra demands of the articulatory-suppression task.

Summary. In this chapter, a modification of the Separate Language Interpretation Resource Hypothesis (Caplan & Waters, 1999) is proposed. The revised model assumes two processing stages, one for syntactic and one for propositional processing. In the first stage, a syntactic representation up to the clause level is built up. During this process, a phonological back-up representation is addressed for the checking of subject-verb agreement features. This checking procedure becomes complicated when features must be checked over a long distance and when the quality of the back-up has deteriorated and large parts of the degraded back-up must be consulted.

Chapter 4
Reversibility and the Use of the Phonological Loop

Introduction

It is generally assumed that working memory (WM) plays a role in sentence comprehension (Baddeley, 1986; Berwick & Weinberg, 1984; Caplan & Waters, 1999; Gibson, 1998; Just & Carpenter, 1992; Martin, 1993; Stowe, Withaar, Wijers, Broere, & Paans, 2002; Vos, 1999). Although there is consensus about the fact that a WM system is needed for language comprehension, there is disagreement about its exact role and architecture.

In the discussion about the organization of the WM system that is involved in language comprehension, some researchers argue that it consists of a single module (e.g., Just & Carpenter, 1992; King & Just, 1991). However, there are those who argue for a partition of resources (e.g., Baddeley, Vallar, & Wilson, 1987; Martin & Romani, 1994; Waters, Caplan, & Hildebrandt, 1987).

The current chapter ties in with this discussion since it evaluates the evidence for a partition of WM resources put forward by Waters et al. (1987, exp. 3). It will report two follow-ups to Waters et al.'s experiment and to Experiments 1 and 2 in Chapter 3.

Experiments 1 and 2 used reversible structures, whereas the Waters et al. (1987) experiment used nonreversible sentences. Experiments 3 and 4 in the current chapter are follow-ups to Waters et al. that used nonreversible structures.

Outline

This introduction has a certain overlap with Chapter 2 and with the introductory section of Chapter 3. Readers who are familiar with the Waters et al. (1987) model and Experiments 1 and 2 may want to proceed directly to Experiment 3 on page 79.

The first subsection of the current introduction is devoted to Waters et al. (1987, exp. 3). It begins with a summary of their data; this summary is followed by an outline of the model they developed based on these data.

The second subsection will discuss some critiques of Waters et al.'s (1987) study. First, the experiment potentially confounded Propositional Complexity, as tested by contrasting one- vs. two-action sentences, and Embedding Type. Some of their two-, but none of their one-action sentences contained center-embedded sentences. Center-embedded structures are known to be more complex than right-embedded clauses (e.g., Gunter, 1995; Miller & Isard, 1964). Second, their account of the data in Waters et al. (1987, exp. 3) is hard to reconcile with neuro-imaging data presented by Caplan, Alpert, & Waters (1998).

The third subsection of this introduction is dedicated to Experiments 1 and 2 in Chapter 3, which were follow-ups to Waters et al. (1987). These experiments differed from the original study in three important respects:

1. Experiment 1 contained right-embedded structures only, whereas Waters et al. had included center-embedded constructions in some of their experimental conditions while Experiment 2 contrasted right- vs. center-embedded structures;
2. the experiments used Dutch instead of English materials; and
3. reversible structures were used instead of the nonreversible constructions used in the Waters et al. experiment.

Especially the last difference may have consequences for the comparability of Waters et al. (1987, exp. 3) and Experiments 1 and 2. These consequences will be discussed in the fourth subsection of this introduction.

A Separate-Resource Model

Waters et al.'s experiment (1987, exp. 3) was conducted in order to determine how many WM resources were addressed during sentence comprehension. Waters et al. used the whole-sentence anomaly-judgment paradigm, in which sentences appear as a whole on a computer screen and participants have to judge the plausibility of these sentences.

Their experiment orthogonally varied the following factors: Clause Type (subject- vs. object-relative clauses, SR vs. OR), Propositional Complexity (one- vs. two-proposition sentences, 1P vs. 2P), and Articulatory Suppression (no load vs. counting from 1 to 6 out loud). Examples of the different sentence conditions are given below: 1PSR (cf. 1a), 1POR (cf. 1b), 2PSR (cf. 1c), 2POR (cf. 1d):[1]

(1) a. It was the thief that broke into the warehouse
b. It was the broken clock that the jeweller adjusted
c. The man hit the landlord that requested the money
d. The meat that the butcher cut delighted the customer

Waters et al. (1987) found main effects of all three factors. Subject relatives were responded to more quickly and accurately than object relatives; one-proposition sentences produced shorter response latencies and fewer errors than two-proposition sentences; and readers were slower in the counting condition than in the no-load condition. Additionally, Waters et al. reported that Postsyntactic Complexity interacted with Articulatory Suppression: two-proposition sentences were more affected by concurrent counting than one-proposition sentences in both reaction time (RT) and error data.

On the basis of these results, Waters et al. (1987) proposed that two WM resource pools are deployed during sentence processing. The first resource pool is dedicated to syntactic processing, and the second one is concerned with propositional content checking. Waters et al. assumed that during the propositional content check, a phonological back-up representation is addressed because of the interaction of and Propositional Complexity x Articulatory Suppression. For a more detailed description of this model, the reader is referred to page 25.

The assumption that a phonological back-up representation is used during the postsyntactic stage stems from Sternberg's (1969) interaction logic. Sternberg proposed that if two factors produce additive effects, they are processed using separate resources. If they produce a superadditive effect, however, they share a common resource pool. Some interactions are non-additive in nature; in these cases, the interaction logic cannot be applied. For a more detailed coverage of the interaction logic, the reader is referred to page 18 of this book.

Shortcomings of the Separate-Resource Theory

Waters et al.'s (1987) account of the findings from Waters et al. (exp. 3) has a number of shortcomings, both from a theoretical and an empirical point of view. A

[1] Technically speaking, sentences (1a and b) are not relative clauses, but cleft sentences; for the sake of convenience, these sentences will be referred to as one-propostion subject- and object-relative sentences.

theoretical objection against the model concerns the use of a phonological back-up representation in postsyntactic processing. In the model, processing proceeds from lower- to higher-level representations: phonemes are analyzed and chunked to words; words are grouped together in syntactic phrases; and phrases in sentences. Why would readers analyze a sentence, combine the information to a syntactic representation, and then go all the way back to phonology for propositional analysis? This idea seems counterintuitive, unparsimonious, and inefficient.

In addition to this theoretical challenge to the model, there are two empirical issues. First, some of the two-, but none of the one-action sentences contained center-embedded structures (cf. 1b vs. 1d). Center-embedded sentences have been known to be hard to process from a syntactic point of view. In other words, the interaction of Propositional Complexity x Articulatory Suppression may in fact be an interaction of Embedding Type x Articulatory Suppression. This suggestion is reinforced by the fact that Experiment 1 reported in Chapter 3 produced independent effects of Propositional Complexity and Articulatory Suppression and that Experiment 2 showed a two-way interaction of Embedding Type x Articulatory Suppression.

A second empirical concern about Waters et al.'s (1987) model is that it is hard to reconcile with a neuro-imaging study conducted by Caplan, Alpert, & Waters (1998, exp. 2). In that experiment, increased regional cerebral bloodflow was found in posterior regions for the one vs. two-action manipulation. These areas are different from the regions found to be active during phonological-loop activity by Paulesu, Frith, & Frackowiak (1993).

Table 1 gives an overview of Talairach & Tournoux (1988) stereotactic coordinates of propositional processing as found by Caplan et al. (1998) and of the phonological-loop activations reported by Paulesu et al. (1993). If the phonological loop were to play a role in propositional processing, one would have expected at least some overlap between the two studies; as can be seen in Table 1, this is not the case.[2]

However, their neuro-imaging results do show activation of Broca's area during syntactic processing as measured by contrasting right- and center-embedded structures. This area is also active during articulatory rehearsal in the Paulesu et al. (1993) study. This suggests a role for the phonological loop in syntactic processing rather than propositional processing. However, we must stress that the Talairach & Tournoux (1988) coordinates of Broca's area differ between the Paulesu et al. and the Caplan et al. (1998) study.

---

[2] Given the fact that stereotactic coordinates refer to a three-dimensional space, both the *x*, *y*, and *z* coordinates must overlap for two activations to be located at the same location in the brain.

Table 1
Brain Areas Activated by Syntactic and Propositional Processing and by Phonological Storage and Articulatory Rehearsal

Study	Region (Brodmann's area)	Stereotactic coordinates		
		x (mm)	*y* (mm)	*z* (mm)
	Propositional Processing			
Caplan et al. (1998)	left occipital (BA 37)	-32	-66	0
Caplan et al. (1998)	left inferior temporal (BA 18)	-22	-84	-4
Caplan et al. (1998)	right inferior temporal (BA 19)	46	-40	-4
Caplan et al. (1998)	medial temporal (BA 19/39)	-52	-64	20
	Syntactic Processing			
Caplan et al. (1998)	medial frontal gyrus (BA 8)	10	5	52
Caplan et al. (1998)	cingulate gyrus (BA 24)	-2	6	40
Caplan et al. (1998)	Broca's area (BA 44)	-42	18	24
	Phonological Store			
Paulesu et al. (1993)	left supra marginal gyrus (BA 40)	-44	-32	24
	Articulatory Rehearsal			
Paulesu et al. (1993)	Broca's area (BA 44)	-64	2	16

Note. Stereotactic coordinates refer to the maximal activation indicated by the highest *Z* score in a particular cerebral structure. Distances refer to the stereotactic space defined by Talairach & Tournoux (1988).

Evidence for Separate Resources Reconsidered

Because of the weak theoretical basis of the involvement of the phonological loop in propositional processing, the potential confound of propositional processing by embedding in Waters et al.'s (1987) materials, and incompatibility of the model with neuro-imaging data reported by Caplan et al. (1998) and Paulesu et al. (1993), Experiments 1 and 2 reported in Chapter 3 were conducted.

Experiment 1. Experiment 1 was a Dutch replication of the original Waters et al. (1987) study that used right-embedded structures only. A second difference between the Waters et al. experiment and Experiment 1 was that Waters et al. used nonreversible sentences, whereas in Experiment 1, reversible sentences were used. This was done in order to increase the syntactic complexity of the sentences and to reduce the chance of a Type II error.

The experiment replicated the main effect of Propositional Complexity in both the RT and error analysis. In addition, there was a marginally significant non-additive interaction of Propositional Complexity x Clause Type (subject- vs. object-relative clauses) in the RT analysis. Due to the non-additive nature of the interaction, it was not possible to draw conclusions as to the (in)dependence of syntactic and propositional processing.

In addition to these effects, there was a main effect of Articulatory Suppression in the RT analysis. It is noteworthy that this effect was opposite to the one found by Waters et al. (1987). In the replication, participants were faster in the counting condition than in the no-load condition.

Experiment 2. Experiment 2 orthogonally varied Clause Type (subject- vs. object-relative sentences), Embedding Type (right vs. center), and Articulatory Suppression (no load vs. counting from 1 to 6 out loud) using reversible structures as well. There were main effects of Clause Type (as in Experiment 1, subject relatives were read more quickly than object relatives) and of Articulatory Suppression (as in Experiment 1, the counting condition produced faster RTs). Additionally, there was an interaction of Embedding Type x Articulatory Suppression: center-embedded structures yielded longer response latencies than right-embedded sentences in the counting condition, but effects of Embedding Type were absent in the no-load condition.

Implications. These two experiments yielded three important results: faster RTs in the articulatory-suppression conditions, presence of an interaction of Embedding Type x Articulatory Suppression in Experiment 2, and absence of an interaction of Propositional Complexity x Articulatory Suppression in Experiment 1.

In Chapter 3, it was proposed that readers of Dutch are faster in the counting conditions of the two experiments due to strategic factors. In Dutch embedded clauses, the verb appears at the end (cf. *accepteert [accepts]* in [2]):

(2) De studenten prijzen de fotograaf die de uitgevers accepteert
The students praise the photographer that the publishers accepts
'The students praise the photographer who accepts the publishers'

Therefore, readers must delay structural integration processes until the end of the embedded clause. Faster RTs for targets and higher error rates for implausible fillers in the counting conditions suggested that readers want to finish processing the unintegrated structure as soon as possible and try to find the verb quickly under articulatory suppression. It was suggested that this was to avoid degeneration of a phonological back-up representation that is addressed during subject-verb integration.

The other two main findings from Experiments 1 and 2 have consequences for the model presented by Waters et al. (1987). First, the presence of an interaction of Embedding Type x Articulatory Suppression in Experiment 2 suggests that the phonological loop plays a role in syntactic processing. Caplan & Waters (1999) and

Waters et al. explicitly predict that these interactions should not occur: under their hypothesis, interactions of Propositional Complexity x Articulatory Suppression ought to be found instead. Therefore, the lack of an interaction of Propositional Complexity x Articulatory Suppression in Experiment 1 pose a further problem to the Waters et al. model. This interaction does not support any role for the phonological loop in propositional processing whatsoever.

Summary. The results of the Experiments 1 and 2 argue for a partition of WM resources different from the one proposed by Waters et al. (1987). Especially the role of phonological representations in sentence processing needs to be reconsidered. The fact that Articulatory Suppression interacts with Embedding Type suggests that phonological representations are used during integrative processes at the syntactic stage, but not during propositional processing.

Reconsiderations Reconsidered

Reversible vs. Nonreversible Sentences. However, one should not jump to conclusions. Waters et al. (1987) used a whole-sentence anomaly-judgment task in which nonreversible sentences were used, whereas the materials in Experiments 1 and 2 were reversible. In Waters et al.'s experiment, implausible sentences were θ-role inversions of plausible sentences, as in (3b):[3]

(3) a. It was the broken clock that the jeweller adjusted
b. # It was the broken clock that adjusted the jeweller

In Experiments 1 and 2, which used reversible sentences, θ-role inversions could not be used as implausible distractors, because inverting θ-roles in reversible sentences yields a plausible sentence, as can be witnessed in (4):

(4) a. It was the senator who attacked the lawyer
b. It was the senator who the lawyer attacked

Instead, implausible sentences were created by replacing the embedded verb with another verb that was not compatible with the rest of the clause, as demonstrated in (5):

[3] The pound sign (#) is used in this book for sentences that are syntactically correct, but semantically unacceptable.

(5) # It was the senator who devoured the lawyer

Using reversible instead of nonreversible sentences does not necessarily weaken the conclusions drawn from Experiments 1 and 2: Waters et al.'s (1987) model predicts interactions of Propositional Complexity x Articulatory Suppression, and not of Clause Type x Articulatory Suppression. Experiments 1 and 2 showed exactly the opposite pattern. This finding poses a problem to the Waters et al. model.

Although the data from Experiments 1 and 2 suggest that Waters et al.'s (1987) model needs revision, a direct comparison between these experiments and the Waters et al. experiment is somewhat problematic. As can be seen in examples (3) and (4), the judgment tasks differed between the two studies. This may have consequences for complexity of the tasks.

The whole-sentence anomaly-judgment task has two subcomponents: (1) readers must parse the sentence that is being judged and (2) once the sentence is understood, the reader must make a judgment. Reversible and nonreversible sentences provide different cues for the two subcomponents of the judgment task. Baddeley et al. (1987) and Caplan & Waters (1999) both suggest that articulatory suppression inhibits plausibility judgment. Let us consider the difference between reversible and nonreversible sentences for each subcomponent.

Consequences for Parsing. Readers have a number of cues that they can use in order to parse a sentence correctly. The availability of these cues is language dependent. Experiments 1 and 2 used Dutch materials. Given the fact that Dutch has different structural properties than English, it is useful to explain these structural properties.

Dutch relative clauses have a different word order than English ones. In Dutch relative clauses, the verb appears at the end of the clause. Subject- and object-relative clauses are distinguished in a different way in Dutch than in English. In English, word order is the major cue, as can be seen in (6a), a subject relative and in (6b), an object relative:

(6) a. The reporter that attacked the senator admitted the error
b. The reporter that the senator attacked admitted the error

If *that* is followed by a verb, the clause is subject relative, if it is followed by a noun, it is object or oblique relative.

Dutch subject- and object-relative clauses both have a noun-noun-verb order. Therefore, Dutch readers must rely on different cues in order to decide what kind of sentence they are reading: frequency information, subject-verb agreement, animacy information, and (to a lesser extent) word order.

Subject-verb agreement is the most critical cue: if the embedded verb agrees with the first noun phrase (NP) of the relative clause, the sentence is subject relative, if it agrees with the second, the clause is object relative.

In addition, animacy information provides an important cue: animate NPs are more likely to be the subject of a sentence than inanimate ones in combination with most verbs.

Frequency information plays a role in the sense that subject relatives are much more frequent than object relatives. Therefore, there is a frequency bias for subject over object relatives.

Finally, word order, in combination with the frequency information just discussed, provides an additional cue: if the clause starts with an inanimate NP, chances are high that the clause is object relative.

When Dutch participants must parse a nonreversible sentence, they can use all four cues. When they have to parse reversible sentences, however, word order and animacy information do not provide additional cues. As a consequence, readers must rely on frequency information and subject-verb agreement when parsing reversible sentences.

Although subject-verb agreement is the most critical cue, the morphosyntactic information expressing subject-verb agreement is phonologically nonsalient. In Chapter 3, it was proposed that articulatory suppression degrades a phonological back-up representation that is addressed during integration of the subject and the verb of a clause. It may well be that some parts of the back-up representation are more vulnerable to degradation than others. If that is the case, phonologically nonsalient elements are likelier to be subject to decay than salient parts of the back-up. Thus, it may be that particular interactions with Articulatory Suppression were enhanced by using reversible instead of nonreversible sentences.

Consequences for the Judgment Task. Now let us examine the consequences of reversibility for the judgment task. Waters et al. (1987) used θ-role inversions as semantically anomalous sentences, whereas in Experiments 1 and 2, lexical distractors were used. This means that in order to make the judgment, participants must rely on different cues in the two studies.

In Waters et al.'s (1987) experiment, they must rely crucially on the outcome of syntactic analysis in order to make a judgment. In Experiments 1 and 2, participants had to use a different cue: compatibility of the embedded verb with the two NPs of the embedded clause. In fact, it may even have been possible for the subjects in

Experiment 1 and 2 to perform well without doing the syntactic analysis. They may have judged the sentences by checking whether the words in these sentences could possibly constitute a plausible sentence without making the effort to analyze it syntactically.

The suggestion that subjects may have skipped the syntactic analysis in Experiments 1 and 2 is far from convincing, however. If no syntactic analysis were made, there should have been no effects of syntactic complexity. This prediction is not borne out by the data: the two experiments produced effects of Clause Type, and Experiment 2 produced effects of Embedding Type. Although there is only a weak basis for believing that subjects may have omitted or rushed through the syntactic analysis, this possibility cannot be ruled out entirely.

In order to investigate whether differences between Experiments 1 and 2 on the one hand and Waters et al.'s (1987) on the other hand, Experiment 3 was conducted. It was a replication of Waters et al. (exp. 3) and of Experiment 1.

Experiment 3

Introduction

This experiment is a replication of Waters et al. (1987, exp. 3). It differs from the latter experiment in two respects: only right-embedded structures were used, and the materials were in Dutch. It is also similar to Experiment 1 in Chapter 3, but is different from that experiment in the sense that nonreversible sentences were used in the current experiment instead of reversible structures.

Method

Subjects. Thirty-two undergraduate students at the University of Groningen participated in the experiment; they were paid participants. All participants had normal or corrected to normal vision and were native speakers of Dutch. Fourteen participants were male; 18 were female. The mean age was 20.2 years old; the participants' age range was between 18 and 27. No participants that had participated in Experiments 1 or 2 of Chapter 3 were tested in the current experiment.

Materials. The current experiment varied the following factors: Clause Type (subject- vs. object-relative clauses), Propositional Complexity (1 vs. 2 propositions), and Articulatory Suppression (no load vs. counting from 1 to 6 out loud). The Clause Type manipulation was within items: object relatives were derived from the subject relatives by changing the order of the nouns. The factor Propositional Complexity was varied between items: half of the items were one-proposition sentences; the other half consisted of two-proposition sentences. Half of all sentences were presented in a no-load block; the other half in a counting condition.

The subject- and object-relative clauses were created by changing the word order and the relative pronoun in the sentence, as shown in (7a and b):

(7) a. Dit zijn de politici die het voorstel afwezen
These are the politicians that the proposal rejected
'These are the politicians who rejected the proposal'

b. Dit is het voorstel dat de politici afwezen
This is the proposal that the politicians rejected
'This is the proposal who the politicians rejected'

In the current experiment, all object-relative clauses start with inanimate relative pronouns and all subject relatives start with animate relative pronouns. This can be seen in (7a and b) above: in (7a), a subject-relative clause, the antecedent of the relative pronoun, *de politici (the politicians),* is animate. In (7b), which is object relative, the first NP of the embedded clause is inanimate: *het voorstel (the proposal).*

Note that subject and object relatives may be ambiguous because they have similar surface structures, as demonstrated in (8):

(8) Dit is de man die de vrouw kust
This is the man that the woman kisses
'This is the man who the woman kisses/This is the man who the woman kisses'

As mentioned before, if the verb agrees with the first NP, the sentence is subject relative, and with object relatives, the verb agrees with the second (subject) NP. In example (8), the verb agrees with both the first and the second NP, and therefore both the subject relative and the object-relative reading are possible.

When comparing subject- and object-relative clauses, one should make sure that all structures are unambiguous. The current experiment uses the whole-sentence anomaly-judgment paradigm. Therefore, implausible sentences should be included in the experiment. In the Waters et al. (1987) materials, implausible sentences were θ-role inversions of plausible sentences, as shown in (9):

(9) a. It was the broken clock that the jeweller adjusted
b. # It was the jeweller that the broken clock adjusted

It is important to keep differences between the Waters et al. (1987) study and the current replication as minimal as possible. Therefore, the same anomalies are used in the current experiment. There is a complication with θ-role inversions in Dutch relative clauses, however. If the two NPs have the same number, disambiguation of the subject- and object-relative versions of a sentence is not possible, as can be seen in (10):

(10) a. Dit is de man die het koekje opeet
This is the man that the cookie eats
'This is the man who eats the cookies'

b. Dit is het koekje dat de man opeet
This is the cookie that the man eats
'This is the cookie that the man eats'

Unlike in English, changing the order of the words is not effective for the construction of implausible sentences. Changing the word order in Dutch only generates the object-relative counterpart of a subject-relative clause. Therefore, one has to use both subject-verb agreement and a singular-plural distinction between the two nouns of the embedded clause as disambiguating cues:

(11) a. Dit is de man die de koekjes opeet
This is the man that the cookies $\text{eats}_{\text{sg.}}$
'This is the man who eats the cookies'

b. # Dit zijn de koekjes die de man opeten
These are the cookies that the man $\text{eats}_{\text{sg.}}$
'These are the cookies that eat the man'

The factor Propositional Complexity was manipulated between items: half of the items were one-proposition sentences. The other half consisted of two-proposition sentences. Examples of the different experimental conditions are given below: 1PSR (cf. 12a), 1POR (cf. 12b), 2PSR (cf. 12c), and 2POR (cf. 12d):

(12) a. Dit zijn de politici die het voorstel afwezen
These are the politicians that the proposal $\text{rejected}_{\text{pl.}}$
'These are the politicians who rejected the proposal'

b. Dit is het voorstel dat de politici afwezen
This is the proposal that the politicians $\text{rejected}_{\text{pl.}}$
'This is the proposal that the politicians rejected'

c. De experts beoordeelden de juwelier die de horloges repareerde
The experts judged the jeweler that the watches $repaired_{sg.}$
'The experts judged the jeweler who repaired the watches'

d. De expert beoordeelde de horloges die de juwelier repareerde
The experts judged the watches that the jeweler $repaired_{sg.}$
'The experts judged the watches that the jeweler repaired'

In all, there were 48 target sentences divided over 8 conditions (all sentence types occurred in a no load and a counting block). All target sentences were semantically plausible.

Materials were matched between conditions for the number of letters and frequency of usage of the first NP, the second NP, and the verb (taken from the CELEX corpus; cf. Burnage, 1990) of the embedded clause and for sentence plausibility. The mean plausibility of each experimental condition is presented in Table 2.

Table 2
Mean Plausibility per Experimental Condition

Sentence type	Mean plausibility
1PSR	4.36
1POR	4.14
2PSR	3.66
2POR	3.57

Note. In a pretest, sentence plausibility was rated on a scale from 1 *(very bad)* to 5 *(excellent)*. Reversibility for noun-verb-noun combinations was rated in a different pretest as –1 *(N1 is a better agent),* 0 *(both are equally likely as agents),* or 1 *(N2 is a better agent).*

Because the object relatives were derived from the subject relatives, each sentence occurred in two versions. In order to avoid repetition effects arising from participants seeing the same sentence twice, sentences were divided over two lists. Sentences that occurred in list 1 as subject relatives would occur in list 2 as object relatives, and vice versa. Participants were assigned to lists randomly.

Each list was divided in two halves. This was done because each participant had a counting and a no-load block. The order in which the list halves were presented was counterbalanced across participants.

The order of the articulatory-suppression condition was also counterbalanced across participants. Materials were presented in two blocks. In one block, participants had to perform the judgment task without a concurrent task, in the other, they had to judge sentences while counting from 1 to 6 out loud. Table 3 gives an overview of the characteristics of each of the participant groups that resulted from the various counterbalancings.

Table 3
Characteristics of Each Participant Group in Experiment 3.

Group	Participants	List	Order of halves	Task order
1	1, 9, 17, 25	1	A1	CO1
2	2, 10, 18, 26	1	A1	NL1
3	3, 11, 19, 27	1	B1	CO1
4	4, 12, 20, 28	1	B1	NL1
5	5, 13, 21, 29	2	A1	CO1
6	6, 14, 22, 30	2	A1	NL1
7	7, 15, 23, 31	2	B1	CO1
8	8, 16, 24, 32	2	B1	NL1

Note. List assignment, and order of list halves and of task condition are given for each participant in Experiment 3. A1 = list half A presented first; B1 = list half B presented first; CO1 = counting condition presented first; NL1 = no-load condition presented first.

In addition to the experimental sentences, 168 distractor sentences were created. Twenty-four of these had the same structures as the target sentences, but were semantically implausible. The implausible sentences were θ-role inversions of sentences that were semantically plausible. Additionally, 96 distractor sentences served as experimental targets for Experiments 4 and 5b (cf. pp. 87 and 114); these sentences were all semantically plausible. Another 48 distractor sentences were θ-role inversions of the targets for Experiments 4 and 5b; these were implausible.

Materials were presented in two blocks. In one block, participants had to perform the task without a concurrent task; in the other, they had to judge sentences while counting from 1 to 6 out loud.

Procedure. Participants had to read sentences from a computer screen. Sentences were presented at once and remained on the screen until a judgment was made. Participants could respond by pressing a button labeled *correct* if the sentence was correct or by pressing an *incorrect* button if the sentence was not. Each sentence was preceded by a warning signal, an asterisk, which disappeared after the participant had pressed the *correct* button. After each judgment, participants received feedback as to whether they had made the right judgment. Materials were presented in two blocks: a no load and a counting block. Each block was preceded by a practice session so that participants would get used to the judgment task and the articulatory-suppression conditions.

Results

RT Analysis. The RT data from all participants were submitted to standard descriptive and inferential analysis. First, the raw data of the two list halves of each participant were checked for outliers. RTs shorter than 200 ms were taken to reflect errors with the response boxes unrelated to sentence processing and were not included in the analysis. The absolute upper cut-off was 9,000 ms.

Means and standard deviations were computed per list half in order to determine the relative cut-offs for each participant. The relative lower cut-off was the mean RT minus 2.5 standard deviation; the relative upper cut-off was the mean RT plus 2.5 standard deviation. Data points that were below the lower cut-offs (both absolute and relative) were replaced by the relative lower cut-off; values that exceeded the higher cut-off points were replaced by the upper relative cut-off criterion.

After outlier detection, means and standard deviations per condition were calculated over subjects and items. Mean RTs per condition are given in Figure 1.[4]

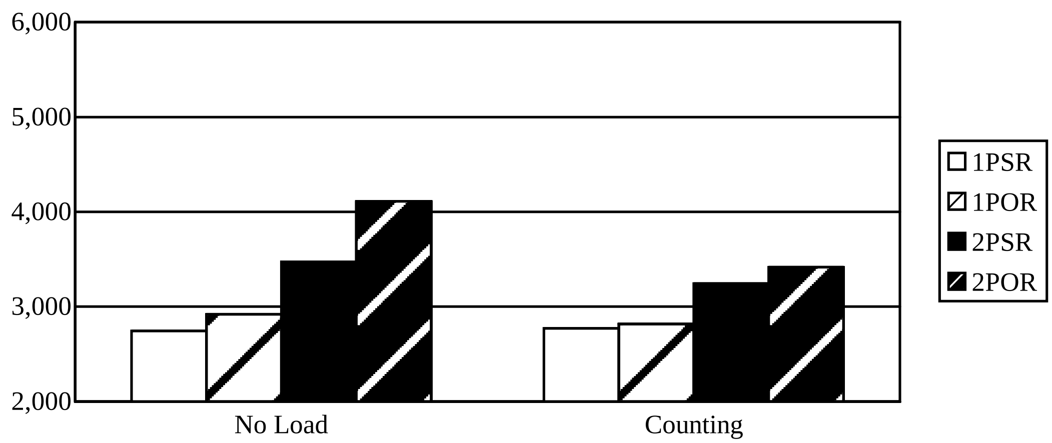


Figure 1. Mean RTs in ms for one-proposition subject relatives (1PSR), one-proposition object relatives (1POR), two-proposition subject relatives (2PSR), and two-proposition object relatives (2POR) in the no load (left) and counting conditions (right).

For the subject analysis (*F1*), an analysis of variance (ANOVA) was performed, which used Clause Type (subject relative vs. object relative), Propositional Complexity (1 proposition vs. 2 propositions), and Articulatory Suppression (no load vs. counting) as within-subject factors and List-Half Order, Task Order, and List as between-subject factors. For the item analysis (*F2*), an ANOVA was performed using Clause Type and Articulatory Suppression as within-item factors and Propositional Complexity and Item Group as between-item factors. Here, effects are taken to be significant when they are significant in both the *F1* and the *F2* analysis. Effects that

[4] In order to facilitate comparison of these data with the RTs of Experiments 1, 2, 4, 5a, and 5b, the scale of each figure was set between 2,000 and 6,000 ms.

were significant in one and approaching significance in the other analysis (with $p <$ .1) will also be mentioned.

The ANOVA revealed a main effect of Clause Type: subject relatives were read more quickly than object relatives, $F1$ (1, 16) = 5.35, $p < .05$; $F2$ (1, 35) = 4.48, $p < .05$. There was also a main effect of Propositional Complexity, $F1$ (1, 16) = 65.05, $p < .001$; $F2$ (1, 35) = 42.76, $p < .001$: one-proposition sentences yielded shorter response times than two-proposition sentences. Articulatory Suppression yielded a main effect that was significant in the analysis over items and approached significance in the subject analysis, $F1$ (1, 16) = 3.96, $p = .06$; $F2$ (1, 35) = 5.49, $p < .05$. No other main effects or interactions were found.

Error Analysis. The error data were analyzed in the same way as the RT data. Mean error proportions per condition are given in Figure 2.[5]

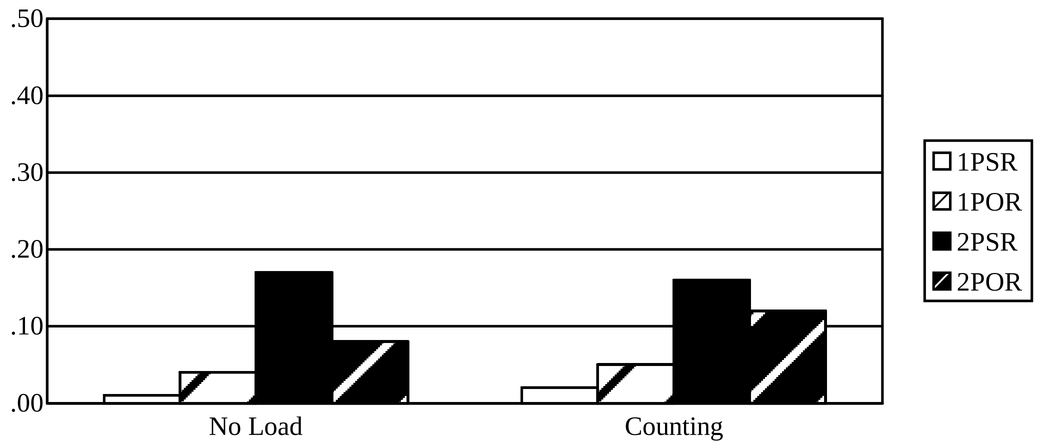


Figure 2. Mean error proportions for one-proposition subject relatives (1PSR), one-proposition object relatives (1POR), two-proposition subject relatives (2PSR), and two-proposition object relatives (2POR) in the no load (left) and counting conditions (right).

The ANOVA revealed a significant main effect of Propositional Complexity, $F1$ (1, 16) = 64.00, $p < .001$; $F2$ (1, 36) = 7.73, $p < .01$, with fewer errors for one-proposition sentences than for two-proposition sentences. Also, there was a significant interaction of Clause Type x Propositional Complexity. In the one-proposition sentences, object-relative clauses produced more errors than subject-relative clauses, $F1$ (1, 16) = 5.56, $p < .05$; $F2$ (1, 23) = 4.39, $p < .05$. However, in the two-proposition sentences, there was a tendency in the opposite direction, i.e., with more errors for subject relatives,

[5] In order to facilitate comparison of these data with the error proportions of Experiments 1, 2, 4, 5a, and 5b, the scale of each figure was set between the error proportions of 0 and .50.

F1 (1, 16) = 6.37, $p < .05$; *F2* (1, 23) = 2.49, $p = .128$. No other main effects or interactions were found. As in Chapter 3, a Wilcoxon Matched Pairs Signed Rank Test was carried out in order to investigate whether faster RTs in the counting condition were accompanied by higher error rates for filler in the articulatory-suppression condition. The test compared the proportion of errors for the implausible sentences in the no-load condition with the error proportion for the implausible sentences of the articulatory-suppression condition. There was a significantly higher proportion of errors for filler sentences in the articulatory-suppression condition than in the no-load condition ($Z = -4.47$, $p < .01$). No significant interactions were found.

Discussion

The current experiment was based on Waters et al. (1987, exp. 3) and on Experiment 1. Waters et al. included center-embedded materials in some of their conditions, whereas Experiment 1 used only right-embedded structures. A second difference between the two studies was that Experiment 1 used reversible sentences, whereas the original Waters et al. study used nonreversible sentences. The current experiment replicated the effect of Propositional Complexity found in Experiment 1 and in the Waters et al. experiment: one-proposition sentences were responded to more quickly and more accurately than two-proposition sentences.

There was also a main effect of Clause Type in the RT analysis: subject-relative clauses yielded shorter response latencies than object-relative constructions. This is in line with the proposal made by Frazier & Flores ’d Arcais (1989) and Kaan (1997) that readers of Dutch tend to prefer the subject- over the object-relative reading of a sentence. In Experiment 1, there was also an effect of Clause Type, but this was found for the one-proposition sentences only. In that experiment, the interaction seemed to reflect a ceiling effect, the idea being that the two-proposition sentences took so long that the effects of Clause Type did not emerge here. The fact that in the current experiment, the effect shows up in both the one and the two-proposition sentences, combined with faster overall RTs, supports this idea.

Additionally, the effect of Articulatory Suppression was similar to that of both Experiments 1 and 2. This finding is yet another replication of the reversed load effect; this effect is probably the result of a strategy that readers apply when faced with the challenges of articulatory suppression. They try to reach the verb of the embedded clause as soon as possible, because that is the place where agreement features need to be checked during integration of the different sentence parts. (Remember that Dutch embedded clauses are verb final.)

As far as interactions between the different factors are concerned: crucially, there was no interaction of Propositional Complexity x Articulatory Suppression in Experiment 3, RTs: $F1$ (1, 16) = .26, p = .62; $F2$ (1, 35) = .67, p = .65. This finding suggests, like the lack of such an interaction in Experiment 1, that these factors are independent, contrary to what Waters et al. (1987) claim. In the error rates, there was a two-way interaction of Clause Type x Propositional Complexity. Subject-relative clauses yielded lower error rates than object relatives in one-proposition sentences, but in the two-proposition, there was an opposite tendency. This tendency was significant in the subject analysis, but far from significant in the item analysis. The exact interpretation of this interaction remains open for speculation.

Experiment 4

Introduction

The fact that there was no interaction of Propositional Complexity x Articulatory Suppression with nonreversible sentences suggests two things. First, the interaction of Propositional Complexity x Articulatory Suppression found by Waters et al. (1987, exp. 3) is probably an interaction of Embedding Type x Articulatory Suppression. This suggestion is backed up by the independent effects of Propositional Complexity and Articulatory Suppression in Experiment 1 and a two-way interaction of Embedding Type x Articulatory Suppression in Experiment 2. Second, the lack of a two-way interaction of Propositional Complexity x Articulatory Suppression in Experiment 1 cannot be attributed to the use of reversible clauses.

The first suggestion is based on a null result however. It would be more convincing if one could back it up with a two-way interaction of Embedding Type x Articulatory Suppression with nonreversible materials. The current experiment was conducted in order to see whether the independent effects of Clause Type and Embedding Type and the two-way interaction of Embedding Type x Articulatory Suppression in Experiment 2 could be replicated with nonreversible materials.

Method

Subjects. The same participants were used as in Experiment 3.

Materials. This experiment had an orthogonal design. It varied the following within subject and item factors: Clause Type (subject vs. object relatives; SR vs. OR), Embedding Type (right- vs. center-embedded structures; RE vs. CE), and Articulatory Suppression (no load vs. counting from 1 to 6).

The sentences had a reversible noun-verb-noun combination in the main clause and a nonreversible noun-verb-noun combination in the embedded clause. As in Experiment 3, object relatives were derived from subject relatives by changing the word order in the sentence. In the subject-relative clauses, the animate NP preceded the inanimate one; in the object relatives, the order was the other way round, as can be seen in (13). In order to derive center-embedded constructions from right-embedded structures, word order was changed.

Additionally, psych verbs were used in the center-embedded sentences. This was necessary because in the center-embedded object-relative clauses, the main-clause subject is inanimate. This strongly limits the type of verbs that can occur in the main clause of the center-embedded clauses. Example sentences of each experimental condition are given below: RESR (cf. 13a), REOR (cf. 13b), CESR (cf. 13c), and CEOR (cf. 13d):

(13) a. De columnist veroordeelde de commando's die de kaping beëindigden
The columnist condemned the commandos that the hijack ended
'The columnist condemned the commandos that ended the hijack'

b. De columnisten veroordeelden de kaping die de commando's beëindigden
The columnists condemned the hijack that the commandos ended
'The columnists condemned the hijack that the commandos ended'

c. De commando's die de kaping beëindigden, schokten de columnist
The commandos that the hijack ended, shocked the columnist
'The commandos that ended the hijack shocked the columnist'

d. De kaping die de commando's beëindigden, schokte de columnist
The hijack that the commandos ended, shocked the columnist
'The hijack that the commandos ended shocked the columnist'

All main clauses had one plural and one singular NP in order to exclude the possibility that readers would interpret a sentence as an object-first construction. These constructions are not very frequent in Dutch, but occur in cases where the object is a focal NP (cf. the second sentence in [14]):

(14) Wie straft de leraar precies?
Who punishes the teacher exactly?
'Who does the teacher punish exactly?'

De leerlingen straft de leraar
The pupils punishes the teacher
‘The teacher punishes the pupils’

Although animacy provides sufficient disambiguating information, the plural-singular difference between the two NPs was maintained in the embedded clause as well. This had to do with the kinds of distractors that were used in the current experiment.

The experimental conditions were matched for the number of letters per word position, frequency of usage per word position (cf. Burnage, 1990), and sentence plausibility. Sentence plausibility was rated in a questionnaire on a scale from 1 to 5 (resp. *very bad* and *excellent*). Because plural verbs are longer in terms of number of letters, materials were constructed in such a way that half of the main clauses contained singular and half of them plural verbs. For the embedded clauses, the same approach was taken.

Table 4
Mean Plausibility per Experimental Condition

Sentence type	Mean plausibility
RESR	3.84
REOR	3.56
CESR	3.50
CEOR	3.07

Note. In a pretest, sentence plausibility was rated on a scale from 1 *(very bad)* to 5 *(excellent)*.

Distractor sentences were also included, 168 in total. There were 24 semantically anomalous distractors with the same syntactic structure as the targets. Anomalies were created by changing the number of the embedded verb. In each embedded clause, one of the nouns was plural and the other singular. Given the fact that the subject of the sentence has to agree with the verb, changing the number information of the verb will result in a semantically anomalous sentence as shown in (15) below:

(15) a. De columnist veroordeelde de commando’s die de kaping beëindigden
The columnist condemned the commandos that the hijack ended
‘The columnist condemned the commandos that ended the hijack’

b. # De columnist veroordeelde de commando’s die de kaping beëindigde
The columnist condemned the commandos that the hijack ended
‘The columnist condemned the commandos that ended the hijack’

Changing the number of the embedded verb implied that there was no perfect match between subject- and object-relative clauses in terms of length. In order to get a perfect match between these two conditions in terms of length, half of the clauses had singular animate NPs and plural inanimate NPs, and the other half had the opposite singular/plural distinction.

Additionally, 96 distractor sentences served as experimental targets for Experiments 3 and 5b (cf. pp. 79 and 114); these sentences were all semantically plausible. Another 48 distractor sentences were θ-role inversions of the targets for Experiments 3, 5a, and 5b; these were implausible.

Because the sentences in the different conditions were obtained by changing the order of the main-clause NPs or the number of the embedded verb, each word combination occurs in four conditions. In order to avoid repetition artefacts, all sentences were assigned to four lists. This was done in such a way that no participant would see more than one version of the same combination.

As in the previous experiment, lists were divided in two halves, one for the no load and one for the counting condition. Order of presentation of the list halves and of the articulatory-suppression conditions was counterbalanced across participants. Assignment to lists and counterbalancing of list halves and articulatory-suppression conditions yielded 16 participant groups. Participants were assigned to groups randomly. An overview of group characteristics is given in Table 5.

Table 5
Characteristics of Each Participant Group in Experiment 4.

Group	Participants	List	Order of halves	Task order
1	1, 17	1	A1	CO1
2	2, 18	1	A1	NL1
3	3, 19	1	B1	CO1
4	4, 20	1	B1	NL1
5	5, 21	2	A1	CO1
6	6, 22	2	A1	NL1
7	7, 23	2	B1	CO1
8	8, 24	2	B1	NL1
9	9, 25	3	A1	CO1
10	10, 26	3	A1	NL1
11	11, 27	3	B1	CO1
12	12, 28	3	B1	NL1
13	13, 29	4	A1	CO1
14	14, 30	4	A1	NL1
15	15, 31	4	B1	CO1
16	16, 32	4	B1	NL1

Note. List assignment, and order of list halves and of task condition are given for each participant in Experiment 4. A1 = list half A presented first; B1 = list half B presented first; CO1 = counting condition presented first; NL1 = no-load condition presented first.

Procedure. The procedure was identical to the procedure followed in the previous experiment.

Results

RT Analysis. The RTs of the current experiment were checked for outliers and analyzed using the same procedure as in Experiment 3. Mean RTs for all experimental conditions are given in Figure 3.[6] The ANOVA showed a significant main effect of Clause Type, *F1* (1, 16) = 26.55, $p < .001$; *F2* (1, 42) = 13.14, $p < .01$: object-relative clauses were read more quickly than subject-relative constructions. There was also a main effect of Articulatory Suppression, *F1* (1, 16) = 27.47, $p < .001$; *F2* (1, 42) = 51.99, $p < .001$, with shorter response latencies for the counting condition. No main effect of Embedding Type was found, *F1* (1, 16) = 2.18, $p = .16$; *F2* (1, 42) = 2.35, $p = .13$.

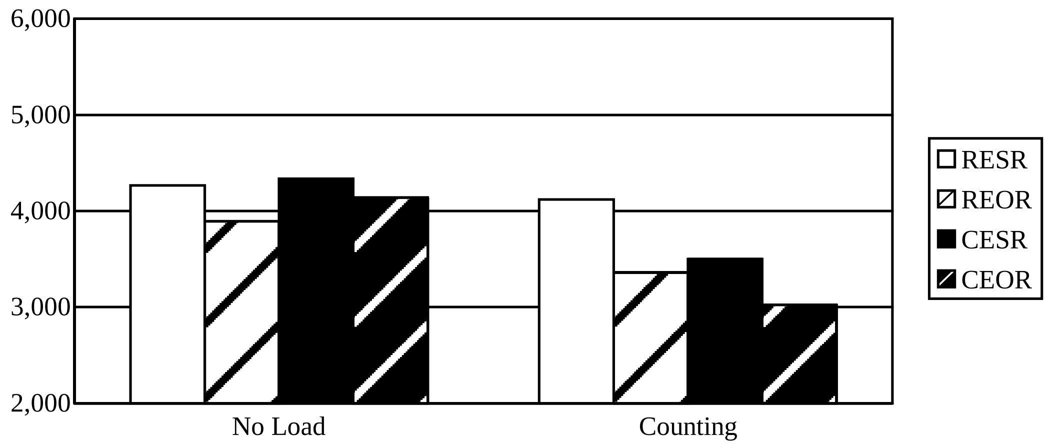


Figure 3. Mean RTs for ms right-embedded subject relatives (RESR), right-embedded object relatives (REOR), center-embedded subject relatives (CESR), and center-embedded object relatives (CEOR) in the no load (left) and counting conditions (right).

In addition to these main effects, there was a two-way interaction of Embedding Type x Articulatory Suppression, *F1* (1, 16) = 10.85, $p < .01$; *F2* (1, 42) = 6.18, $p < .05$. There was a simple effect of Embedding Type in the counting condition: center-embedded sentences yielded shorter RTs, *F1* (1, 16) = 8.14, $p < .05$; *F2* (1, 42) = 9.18, $p < .05$. There were no effects of Embedding Type in the no-load condition, *F1* (1, 16) = 1.8, $p = .19$; *F2* (1, 42) = .57, $p = .45$; no other significant main effects or interactions were found.

---

[6] In order to facilitate comparison of these data with the RTs of Experiments 1, 2 ,3, 5a, and 5b, the scale of each figure was set between 2,000 and 6,000 ms.

Error Analysis. The error data were submitted to the same statistical procedures as the RT data. Mean error proportions per condition are given in Figure 4.[7] In the error analysis, the ANOVA revealed a main effect of Clause Type: subject-relative constructions yielded fewer errors than object-relative sentences, $F1$ (1, 16) = 55.12, $p < .001$; $F2$ (1, 42) = 24.98, $p < .001$). There was also a main effect of Embedding Type, $F1$ (1, 16) = 8.66, $p < .05$; $F2$ (1, 42) = 11.63, $p < .01$), with higher accuracy for right embeddings than for center-embedded constructions. Third, there was a main effect of Articulatory Suppression with decreased accuracy for the counting condition, $F1$ (1, 16) = 14.11, $p < .01$; $F2$ (1, 42) = 20.27, $p < .001$). In addition to these main effects, there was a two-way interaction of Embedding Type x Articulatory Suppression, $F1$ (1, 16) = 5.60, $p < .05$; $F2$ (1, 42) = 5.95, $p < .05$). It was due to the fact that the effects of Embedding Type were significant under articulatory suppression, $F1$ (1, 16) = 8.17, $p < .05$; $F2$ (1, 42) = 12.23, $p < .01$, but not in the no-load condition, $F1$ (1, 16) = 3.32, $p = .09$; $F2$ (1, 42) = 2.35, $p = .13$.

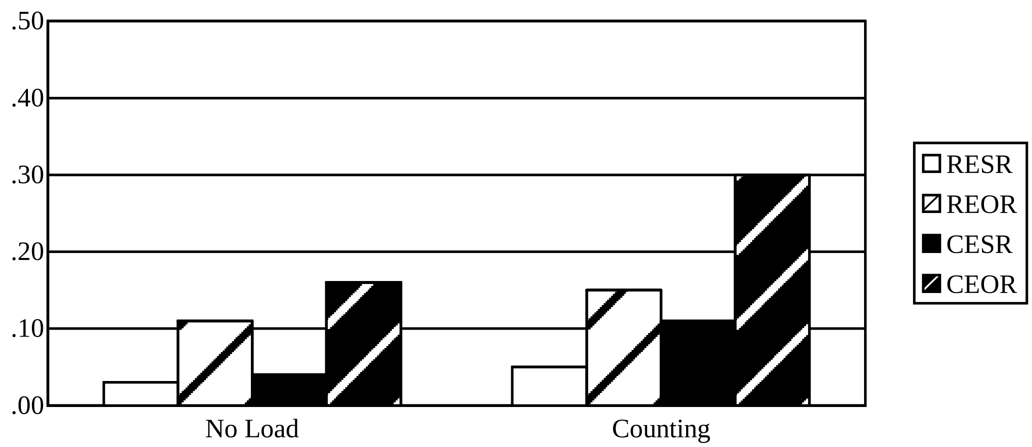


Figure 4. Mean error proportions for right-embedded subject relatives (RESR), right-embedded object relatives (REOR), center-embedded subject relatives (CESR), and center-embedded object relatives (CEOR) in the no load (left) and counting conditions (right).

Discussion

In the current subsection, the effects found will be discussed. Given the fact that so many speed-accuracy trade-offs were found, I will discuss the RT and error data of the effects together.

---

[7] In order to facilitate comparison of these data with the error proportions of Experiments 1, 2 ,3, 5a, and 5b, the scale of each figure was set between the error proportions of 0 and .50.

Articulatory Suppression. Subjects were faster in the counting condition than in the no-load condition. This effect was also found in Experiments 1 and 2. Participants seem to try to find the verb as quickly as possible and in Dutch relative clauses, the verb appears in final position. This causes them to speed up; higher error rates show that participants chose for speed, but did so at the cost of accuracy in Experiment 4.

Although previous experiments have not shown increased errors, it was suggested that these might have been masked by a response bias. The current result confirms that articulatory suppression does lead to processing difficulties and the speeding up of responses

Clause Type. Subject relatives yielded longer response latencies than object relatives, but error rates were higher for object- than for subject-relative clauses. Subject-relative clauses are known to be easier to process than object-relative clauses (e.g., Gibson, 1998; King & Just, 1991). Therefore, lower error proportions for subject relatives probably reflect the fact that participants have less difficulty with the structure building of subject- than of object-relative clauses.

For this reason, the longer RTs for subject-relative clauses are unlikely to be due to processing difficulty in the classical sense. But although subject-relative clauses are easier from a structural point of view, feature checking must occur over a longer distance in subject- than in object-relative clauses. This may lead to slower RTs for subject-relative clauses than for object-relative clauses.

Embedding Type. Worse performance was found for center-embedded clauses than for right-embedded ones. Given the fact that there was a two-way interaction of Embedding Type x Articulatory Suppression, this finding must be seen in the light of that interaction. No effects of Embedding Type were present in the no-load condition, but in the counting condition, center-embedded clauses were read more quickly and less accurately.

This finding suggests that participants maintain the main-clause subject in the phonological loop until it can be integrated. Alternatively, they may address a back-up representation in the phonological loop when the main-clause subject must be integrated with the main-clause verb. Articulatory suppression seems to make it harder to address that back-up. This is either because articulatory suppression impedes active rehearsal of the unintegrated matrix subject or because it degrades the back-up, thus complicating the retrieval of the back-up for reference during integration of the matrix subject and its verb.

Comparison with Experiments 1 to 3. In the next section, the implications of Experiments 3 and 4 will be discussed, along with a comparison between Experiments 3 and 4 on the one hand and Experiments 1 and 2 on the other hand.

General Discussion

Experiments 3 and 4, which were conducted as follow-ups to Waters et al. (1987, exp. 3), yielded data patterns that differ considerably from Waters et al.'s results in a number of respects. In this section, I will discuss these differences and the consequences of these data for the WM resource debate. A second aspect that will be dealt with in this section is the fact that the Dutch materials used in Experiments 3 and 4 seem to evoke different strategies in participants than the English materials in Waters et al.'s experiment.

Consequences for the WM Resource Debate

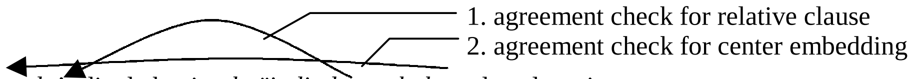
The results found in Experiment 3 and 4 have a number of implications for the WM resource debate, especially for the models proposed by Just & Carpenter (1992) and by Caplan & Waters (1999) and Waters et al. (1987).

The Involvement of the Phonological Loop. The first finding of importance was that Embedding Type x Articulatory Suppression yielded a superadditive two-way interaction in Experiment 4, but that Propositional Complexity and Articulatory Suppression were independent in Experiment 3. This suggests that the interaction of Propositional Complexity x Articulatory Suppression in the Waters et al. (1987) experiment was in fact an interaction of Embedding Type x Articulatory Suppression.

Furthermore, Experiment 4 failed to produce a two-way interaction of Clause Type x Embedding Type. Under the application of Sternberg's (1969) interaction logic, this suggests that the two factors are processed by different WM resource pools.

However, it seems more likely that the independent effects arise because processing complexity effects do not arise simultaneously. For example, one form of complexity is feature checking according to the model developed in Chapter 3.

The feature check between the subject and the verb of the embedded clause and of the main-clause subject-verb agreement features do not occur at the same time. Integration of the embedded subject and verb occurs earlier, because the embedded verb precedes the main verb. Once this integration has been accomplished and subject-verb agreement features of the embedded subject and verb have been checked, the integration and feature checking of the main clause can be carried out. This is illustrated in (16) below:



(16) De *commando's* die de kaping *beëindigden, schokten* de columnist
The commandos that the hijack ended, shocked the columnist
'The commandos that ended the hijack shocked the columnist'

One vs. Multiple WM Resource Pools. The fact that the RT analysis of Experiment 3 produced independent effects of Clause Type and Propositional Complexity supports the assumption of two processing stages made by Caplan & Waters (1999) and Waters et al. (1987). However, this finding poses a problem for the single-resource theory proposed by Just & Carpenter (1992). According to this theory, all kinds of linguistic complexity effects should interact given the fact that they are processed by a single WM system.

Influences of Reversibility. The findings of Experiments 1 and 2 reported in Chapter 3 and of Experiments 3 and 4 show that the use of reversible or nonreversible sentences has no consequences with respect to the partition of WM resources proposed by Waters et al. (1987). However, the fact that Dutch sentences were used has some interesting implications for the strategies that participants used. These will be discussed in the next subsection.

The Use of the Phonological Loop in Syntactic Processing

Now that it has been established that phonological WM representations play a role in syntactic processing, it is necessary to specify that role in more detail. We can do this by looking at the effects of and interactions with Articulatory Suppression, a factor known to burden the phonological loop (Baddeley, 1986; Gupta & MacWhinney, 1995). In the current subsection, a proposal of the role of phonological WM in sentence processing is made. After that, this proposal will be motivated using the data from Experiments 3 and 4 reported in this chapter and from Experiment 2 in Chapter 3.

The Role of Phonological Memory in Syntactic Processing. The phonological loop is assumed to store a phonological back-up of the sentence that is being analyzed (Baddeley et al., 1987; Waters et al., 1987). The syntactic processor addresses this back-up when the subject and verb of a clause or sentence must be integrated in order to check whether subject and verb agree in person and number. Subject-verb integration and the checking of subject-verb agreement features are taken to be demanding when the elements to be integrated are far apart. Given that Articulatory Suppression is taken to hamper phonological-loop functioning, Articulatory Suppression has a detrimental effect in sentence processing when long-distance checking is necessary.

Motivating the Assumptions I: The Main Effect of Articulatory Suppression. The model sketched above is based on a number of effects. Let us first consider the main effect of Articulatory Suppression in the Waters et al. (1987) study and in Experiments 3 and 4. Articulatory Suppression is taken to affect the quality and decay rate of a phonological WM representation (Gupta & MacWhinney, 1995).

If reliance on large parts of this phonological WM representation is important for comprehension, readers will benefit from processing a sentence quickly enough to still have some phonological representation left at the end of the sentence. This is the case for Dutch, a language with verb-final relative clauses, but not for English, a verb-second language.

The speeding of Dutch readers is likely to go at the expense of accuracy. This pattern is exactly what we observe in Experiments 3 and 4. Both experiments produced faster RTs for the articulatory-suppression conditions. Additionally, Experiment 3 showed worse performance on implausible fillers in the articulatory-suppression condition than in the no-load condition, although the decreased accuracy was probably masked for plausible sentences by a *yes*-response bias. Experiment 4 produced a speed-accuracy trade-off with faster RTs and higher error rates for the articulatory-suppression condition.

However, speeding up is not useful for readers of English, because English is not verb final. Therefore, the demands of an additional task in the articulatory-suppression conditions are expected to produce a slow-down. This is exactly what Waters et al. (1987) found: slower RTs for the articulatory suppression than for the no-load conditions.

Motivating the Assumptions II: Interactions with Articulatory Suppression. The idea that readers of Dutch benefit from speeding up when under articulatory suppression is backed up by a two-way interaction of Clause Type x Articulatory Suppression in Experiment 2 reported in Chapter 3.

That experiment used reversible materials as shown in (17). It turned out that subject-relative clauses such as (17a) yielded more errors than object-relative clauses such as (17b) in that particular experiment:

(17) a. De studenten prijzen de fotograaf die de uitgevers accepteert
The students praise the photographer that the publishers accepts
'The students praise the photographer who accepts the publishers'

b. De studenten prijzen de fotograaf die de uitgevers accepteren
The students praise the photographer that the publishers accept
'The students praise the photographer who the publishers accept'

This effect was present under articulatory suppression only. This finding confirms that performance deteriorates due to the burden of Articulatory Suppression when agreement information must be checked over a long distance.

The fact that readers must expend more effort to check subject-verb agreement features over a long distance is also backed up by two-way interactions of Embedding Type x Articulatory Suppression that were found for both English and Dutch.

Waters et al. (1987) reported a two-way interaction of Propositional Complexity x Articulatory Suppression with a larger effect of Propositional Complexity in the articulatory-suppression conditions. However, their two-proposition sentences contained center embeddings, but their one-proposition sentences did not, as can be seen in (18):

(18) a. It was the broken clock that the jeweller adjusted
b. The meat that the butcher cut delighted the customer

Therefore, their interaction can in fact best be explained as a two-way interaction of Embedding Type x Articulatory Suppression with slower RTs for center- than for right-embedded structures in the articulatory-suppression conditions only.

Two-way interactions of Embedding Type x Articulatory Suppression with slower RTs for center- than for right-embedded structures in the articulatory-suppression conditions only were also found for Dutch. There was one in Experiment 2 reported in Chapter 3 and one in Experiment 4 reported in this chapter.

Summary. The findings from Experiments 3 and 4 substantiate the model presented at the end of Chapter 3. This model assumes to processing stages (following Berwick & Weinberg, 1984): one for syntactic and one for propositional processing. The assumption of two processing stages stems from the fact that the RT analysis of Experiment 3 produced independent effects of Clause Type, a complexity at the syntactic level and of Propositional Complexity, a manipulation which is taken to affect sentence level semantic processing.

Furthermore, the model assumes that during the syntactic stage, but not during the propositional stage, a phonological back-up representation is addressed (contra Waters et al., 1987). This checking is assumed to occur during the long-distance integration of subjects and verbs, presumably for feature checking purposes. The idea that a phonological back-up is addressed during the syntactic stage, but not the propositional stage stems from the following observations:

1. Articulatory Suppression yielded slower RTs in English (cf. Waters et al., exp. 3), but faster RTs in Dutch and the main difference between the two languages is word order, a syntactic phenomenon;
2. in both English and Dutch, Articulatory Suppression produces superadditive interactions with Embedding Type, a syntactic factor;[8] and
3. in sentences that used right-embedded structures only, additive effects of Articulatory Suppression and Propositional Complexity.

---

[8] Note that Waters et al. (1987, exp. 3) found a superadditive interaction of Propositional Complexity x Articulatory Suppression, but that they used potentially confounded data because they had included center-embedded structures in their two-, but not in their one-proposition sentences; thus their two-way interaction can be interpreted as an interaction of Embedding Type x Articulatory Suppression.

Chapter 5
The Phonological Loop and Integration of Adjectives

Introduction

In Chapters 3 and 4, four whole-sentence anomaly-judgment experiments were reported that investigated the use of different working memory (WM) resources during sentence comprehension. It is generally accepted that some form of WM is addressed when people read or listen to sentences (e.g., Baddeley, 1986; Caplan & Waters, 1999; Gibson, 1998; Just & Carpenter, 1992). WM is taken to comprise both storage of information and computation of the relations between the elements stored (e.g., Baddeley; Gibson; Just & Carpenter; Salthouse 1994). Thus, the computation of sentence structure is taken to be part of WM.

The findings of Experiments 1 to 4 in these chapters argue for a role of a phonological back-up representation in syntactic processing. The goal of the experiments reported in the current chapter is to further specify the role of such a back-up representation.

Experiments 1 to 4 were conducted to identify possible subcomponents of WM that were used during sentence comprehension. In order to decide whether certain processes were carried out by a single or multiple WM systems, Sternberg's (1969) interaction logic was used. Sternberg proposed that if two effects produce additive effects, they may be processed by different cognitive systems, but that if they produce a superadditive interaction, they are processed by a single cognitive system. For a more detailed discussion of the interaction logic, the reader is referred to pages 10 and 33.

A Two-Stage Model of WM and Sentence Comprehension

On the basis of Experiments 1 to 4 in the Chapters 3 and 4, a sentence comprehension model was proposed that consists of two stages, one for syntactic analysis and one for propositional analysis. The assumption of two processing stages was taken over from Berwick & Weinberg (1984), Waters, Caplan, & Hildebrandt (1987), and Caplan & Waters (1999). The model proposed in the Chapters 3 and 4 assumes that during the syntactic stage, a phonological back-up representation is addressed under some circumstances.

Motivating the Assumption of Two Processing Stages. Two processing stages were proposed. This was done because syntactic complexity and propositional complexity produced independent effects in Experiment 3. Syntactic complexity was tested by comparing subject- and object-relative sentences (cf. 1a and b respectively); propositional complexity was measured by contrasting one- and two-action sentences (cf. 1a and c respectively):

(1) a. These are the politicians that rejected the proposal
  b. This is the proposal that the politicians rejected
  c. The journalist interviewed the politicians that rejected the proposal

The Role of a Phonological Back-Up in Syntactic Processing

Experiments 2 and 4 produced super-additive interactions of Embedding Type (right vs. center) by Articulatory Suppression (no load vs. counting out loud). No effects of Embedding Type were found in the no-load condition, but in the articulatory-suppression condition, center-embedded sentences were read more slowly than right-embedded constructions. Articulatory suppression is taken to hamper inner speech. Therefore, the interaction of Embedding Type x Articulatory Suppression points to a role for inner speech during the processing of center-embedded constructions.

Additionally, Articulatory Suppression interacted with Clause Type (subject- vs. object-relative clauses) in Experiment 2: under articulatory suppression, subject relatives produced higher error rates than object relatives, while in Experiment 4, subject relatives had longer RTs than object relatives. Both of these effects seemed to stem from the distance of subject-verb integration in these conditions. It is noteworthy that Experiments 1 to 4 used Dutch materials.

In Dutch, the subject/object-relative contrast is expressed differently than in English. Therefore, it is useful to discuss briefly how this distinction is made in Dutch. Dutch subject- and object-relative structures have identical surface structures, as shown in (2), a sentence that can be read as both a subject- and an object-relative clause:

(2) Dit is de man die de vrouw kust
This is the man that the woman kisses
'This is the man who the woman kisses/This is the man who the woman kisses'

If the embedded verb agrees with the first noun phrase (NP) of its clause, that clause is subject relative. However, if it agrees with the second NP of the clause, it is object relative; and if it agrees with both, the sentence is ambiguous.

Subject-verb integration occurs over a longer distance in subject relatives than in object relatives. In the subject-relative reading of (2), *man (man)* and *kust (kisses)* must be integrated and in the object-relative reading, *vrouw (woman)* and *kust (kisses)* must be integrated. The latter integration is an integration of adjacent words whereas the first integration must be accomplished over a longer distance.

Because Dutch subject- and object-relative sentences have similar surface orders, they were disambiguated in Experiments 1 to 4 using number: the embedded clauses contained one singular and one plural NP as demonstrated in (3):

(3) De studenten prijzen de fotograaf die de uitgevers accepteert
The students praise the photographer that the publishers accepts
'The students praise the photographer who accepts the publishers'

The embedded subject, *fotograaf (photographer)* is singular and the embedded direct object, *uitgevers (publishers)*, is plural.

The Suppression Task, Subject-Verb Agreement, and Θ-Role Assignment

A striking observation about the interactions found in the Experiments 2 and 4 is that articulatory suppression causes most difficulty in conditions that require long-distance integrations of subject and verb. As can be seen in example (2), distinguishing between the subject- and object-relative reading of the embedded clause hinges on the correct processing of number information, a phonologically nonsalient cue.

It may be that the interaction of Clause Type x Articulatory Suppression in Experiment 2 is due to the use of a phonological back-up representation during the processing of the embedded clause (cf. Baddeley, Vallar, & Wilson, 1987). The phonologically nonsalient elements of that back-up representation are most prone to decay.

However, what is the exact nature of the back-up's involvement during long-distance integrations? There are two hypotheses about the nature of this involvement.

First, it may play a role in the agreement checking for the establishment of subject-verb agreement relations. Second, it may be involved in θ-role assignment.

During subject-verb agreement, two processes are carried out. First, the syntactic relation between an NP and a verb must be established. The processor must determine that a particular NP is the subject of a particular verb. It can do this by checking the subject-verb agreement information. Second, the verb assigns certain roles to the NPs of a sentence. For example, the verb *kiss* assigns two roles: there is one person that is doing the kissing, the *agent*, and one person that is kissed, the *theme*. These roles are usually referred to as θ-roles.

Given that the embedded verbs in Experiments 1 to 4 had agreement relations with their subjects and were θ-role assigners at the same time, one cannot distinguish between these two hypotheses on the basis of these experiments. An important aim of Experiments 5a and b is to investigate whether interactions with Articulatory Suppression can also be found in sentences that require long-distance θ-role assignments, but have local subject-verb agreement checking.

An example of such a sentence is given in (4):

subject-verb agreement θ-role assignment

(4) De *actrice heeft*, terwijl het publiek zat te wachten, haar haar *gekamd*
The actress has, while the audience sat to wait, her hair combed
'The actress combed her hair while the audience was waiting'

The subject-verb agreement relationship is between the matrix subject, *de actrice (the actress)*, and the adjacent auxiliary *heeft (has)*. The θ-roles in the main clause are assigned by the main verb *gekamd (combed)* which appears in final position.

The Influence of Articulatory Suppression on Integration in General

The second issue is due to the fact that in Experiments 1 to 4, all long-distance integrations were subject-verb integrations. Interactions with Articulatory Suppression suggest that inner speech is deployed during subject-verb integration. However, one cannot tell on the basis of the data presented in Chapters 3 and 4 whether inner speech is deployed specifically during these subject-verb integrations, or that it is used for long-distance integration in general.

Interactions found in Experiment 2. In Experiment 2, there was an interaction of Clause Type x Articulatory Suppression with higher error rates for subject- than for object-relative sentences in the articulatory-suppression condition only. As shown in

(2) above, repeated here as (5), integration of subject and verb occurs over a longer distance in subject- than in object-relative constructions:

subject-verb integration for subject relatives
subject-verb integration for object relatives

(5) Dit is de *man* die de *vrouw kust*
This is the man that the woman kisses
'This is the man who the woman kisses/This is the man who the woman kisses'

In order to investigate whether interactions with Articulatory Suppression occurred with other types of integrations as well, we decided to study the relation of inner speech and a different type of integration that has been found to burden WM. Adjective-noun integration is such an integration type.

Adjective-Noun Integration and WM Resources. Martin & Romani (1994) conducted a whole-sentence anomaly-judgment task in which they found that judgments were worse for sentences with delayed adjective-noun integrations, as in (6a), than for sentences with immediate adjective-noun integration, as in (6b):

(6) a. A fluffy, small, surprised kitten came out of the room
b. A fluffy kitten came out of the room

In (6a) integration of the adjective *fluffy* must be delayed over two other adjectives before it can be integrated to its head noun, *kitten.* In (6b), *fluffy* and *kitten* are adjacent and can be integrated immediately.

Martin & Romani (1994) conducted their experiment because they wanted to look at semantic integration processes and its associated WM demands. At this point, I will not discuss their proposal about the role of WM in these processes. Martin & Romani's proposal will be addressed in the General Discussion section on page 118.

Martin & Romani's (1994) findings with respect to delayed integration are in line with the syntactic prediction locality theory proposed by Gibson (1998). Gibson assumed that sentence complexity depends on the number of new discourse referents that need to be active during the integration of certain sentence parts and on how long elements must be stored before they can be integrated.

He assumed that nouns, verbs, and, in some cases, adjectives may be considered as new discourse referents: "it may be that not only nouns and verbs cause memory cost increments for predicted syntactic categories. Adjectives and contentful prepositions may also cause memory cost increments, because they indicate predications." (Gibson, 1998, p. 25)

In Gibsonian terms, the subject-verb integration in the center-embedded conditions of Experiments 2 and 4 is an integration over five new discourse referents, as shown in (7):

(7) De commando's die de kaping beëindigden, schokten de columnist
The commandos that the hijack ended, shocked the columnist
'The commandos that ended the hijack shocked the columnist'

Within the relative clause *commando's (commandos)* must be assigned the agent role of *beëindigden (ended);* for this integration both words must be active. The word *kaping (hijack)* must be assigned the theme role, also requiring both words to be active, and finally *commando's (commandos)* must also be integrated with the matrix verb *schokten (shocked).*

However, the delayed integration condition in Martin & Romani's (1994) experiment contains an adjective-noun integration that must be established over four new discourse referents (e.g., *fluffy, small, surprised,* and *kitten* in [6a]).

Martin & Romani (1994) found that the distance between an adjective and a noun produced a larger detrimental effect if the adjectives precede the noun, as in (6a and b), than if it follows the noun. An example of a sentence with adjectives following the noun is shown in (8):

(8) A kitten that was fluffy, small, and surprised came out of the room

Martin & Romani's (1994) experiment did not include articulatory-suppression conditions. Therefore, their experiment does not offer us any insight in a possible role of the phonological loop during the integration processes that they investigated. A second aim of Experiments 5a and b is to look at precisely that issue. Therefore, we used Martin & Romani's manipulation to look at long-distance integrations other than subject-verb integration. In order to decrease chances of a Type II error, we only included the conditions that proved to burden WM most, the conditions with premodification conditions, as shown in (6a and b).

In Chapters 3 and 4 it was proposed that articulatory suppression impeded the long-distance checking of agreement features between the subject and verb of the matrix clause. Using the number-of-adjectives manipulation used by Martin & Romani (1994) with Dutch materials also offers us a nice opportunity to look at feature checking in integrations other than subject-verb integrations because Dutch not only has subject-verb agreement, but also adjective-noun agreement. How this agreement is expressed in Dutch will be discussed in the next subsection.

Adjective-Noun Agreement in Dutch

In Dutch, there is agreement between adjectives and nouns. Table 1 gives an overview of this agreement between adjectives and nouns within Dutch NPs. Dutch has a two-gender system that emerged from a three-gender system. Originally, the gender system comprised male, female, and neuter gender, but in the course of time, male and female gender have collapsed to one gender, generally referred to as *common gender*.

Table 1
Adjective-Noun Agreement in Dutch NPs

Common		Neuter	
Definite	Indefinite	Definite	Indefinite
		Singular	
de uitstekende wijn the excellent wine	een uitstekende wijn an excellent wine	het uitstekende wijntje the excellent wine$_{dim.}$	een uitstekend wijntje an excellent wine$_{dim.}$
		Plural	
de uitstekende wijnen the excellent wines	uitstekende wijnen excellent wines	de uitstekende wijntjes the excellent wines$_{dim.}$	uitstekende wijntjes excellent wines$_{dim.}$

Note. *Wijn (wine)* has common gender in Dutch. All diminutive words have neuter gender. dim. = diminutive.

If an adjective precedes a common noun, the root adjective is followed by an agreement suffix, *-e;* this is the case for both definite and indefinite NPs. In the examples given in Table 1, the root adjective is *uitstekend (excellent).* In the *common* examples, it always takes up the form *uitstekende (excellent).* Neuter nouns have different agreement with preceding adjectives: with indefinite singular nouns, the root adjective is not followed by the agreement suffix *-e,* as can be seen in the rightmost upper cell of Table 1.

Summary

The goal of the experiments in the current chapter is to investigate two issues about the role of a phonological back-up representation during subject-verb integration:

1. Whether the phonological back-up is used for the checking of subject-verb agreement information or for θ-role assignment; and
2. whether it is deployed during subject-verb integration only or during other integration processes as well.

In order to pursue these issues, it is necessary to look at sentences with local subject-verb agreement, but with long distance θ-role assignment as well as at feature checking during integrations other than subject-verb integration.

Experiment 5a

Method

Subjects. Thirty-two undergraduate students of the University of Groningen served as paid participants. Admission criteria were the same as in Experiment 1 to Experiment 4. Fourteen participants were male; 18 were female. The mean age was 20.1 years old; the participants' age range was between 17 and 26. No participants who participated in Experiments 1 to 4 were included.

Materials. The current experiment was designed to look at whether the type of processing during subject-verb integration was the same as the processing of unintegrated adjectives. Since stacked adjectives require long-distance chunking of subject and verb, it seems intuitively plausible to assume that a similar kind of processing is involved in the two tasks.

The experiment varied the factors Adjectival Load (1 vs. 3; 1A vs. 3A), Embedding Type (right vs. center embedded; RE vs. CE), and Articulatory Suppression (no load vs. counting from 1 to 6) in an orthogonal design. This yielded the following sentence conditions: RE1A (cf. 9a), RE3A (cf. 9b), CE1A (cf. 9c), and CE3A (cf. 9d):

(9) a. De barkeeper heeft de chips opgegeten terwijl de verregende klant niets doorhad
The barkeeper has the chips eaten while the drenched customer nothing noticed
'The barkeeper ate the chips while the drenched customer did not notice anything'

b. De barkeeper heeft de chips opgegeten terwijl de verregende hongerige trouwe klant niets doorhad
The barkeeper has the chips eaten while the drenched hungry loyal customer nothing noticed
'The barkeeper ate the chips while the drenched hungry loyal customer did not notice anything'

c. De barkeeper heeft, terwijl de verregende klant niets doorhad, de chips opgegeten
The barkeeper has, while the drenched customer nothing noticed, the chips eaten
'The barkeeper ate the chips while the drenched customer did not notice anything'

d. De barkeeper heeft, terwijl de verregende hongerige trouwe klant niets doorhad, de chips opgegeten
The barkeeper has, while the drenched hungry loyal customer nothing noticed, the chips eaten
'The barkeeper ate the chips while the drenched hungry loyal customer did not notice anything'

Each condition was examined in both the no-load and the counting condition.

As can be seen in the examples above, the three-adjective sentences, such as (9b and d), were derived from the one-adjective sentences (9a and c) by adding two more adjectives. The center-embedded sentences, such as (9c and d), were derived from the right-embedded versions, such as (9a and b), by moving the embedded clause to the middle of the main clause.

In order to avoid artefacts, the experimental conditions were matched for the number of letters per word position, frequency of usage per word position (taken from the CELEX corpus; cf. Burnage, 1990), and sentence plausibility. Sentence plausibility was rated in a pretest on a scale from 1 *(very bad)* to 5 *(excellent).* The mean plausibility per condition is given in Table 2.

Table 2
Mean Plausibility per Experimental Condition

Sentence type	Mean plausibility
RE1A	3.73
RE3A	3.70
CE1A	3.92
CE3A	3.75

Note. In a pretest, sentence plausibility was rated on a scale from 1 *(very bad)* to 5 *(excellent).*

In addition to the target sentences, which were all semantically plausible, 96 distractor sentences were included. These distractors consisted of 24 sentences that were structurally similar to the targets, but contained implausible adjective-noun combinations, as shown in (10):[1]

(10) # De zuster heeft, terwijl het zwemmende gordijn steeds openbleef, het kussen opgeschud
The nurse has, while the swimming curtain all-the-time open remained, the pillow plumped
'The nurse plumped the pillow while the swimming curtain remained open all the time'

[1] The pound sign (#) is used in this book for sentences that are syntactically correct, but semantically unacceptable.

Apart from these 24 sentences, 72 filler sentences had structures different from the ones that were investigated in the experiment; 48 of these were plausible; the remaining 24 had lexical verbs that were not compatible with the rest of the sentence.

Because the sentences in the different conditions were obtained by changing the word order of the main clause or by adding two more adjectives to a one-adjective sentence, each word combination occurs in four conditions. In order to avoid repetition artefacts, all sentences were assigned to four lists. This was done in such a way that no participant would see more than one version of the same combination.

Table 3
Characteristics of Each Participant Group in Experiment 5a.

Group	Participants	List	Order of halves	Task order
1	1, 17	1	A1	CO1
2	2, 18	1	A1	NL1
3	3, 19	1	B1	CO1
4	4, 20	1	B1	NL1
5	5, 21	2	A1	CO1
6	6, 22	2	A1	NL1
7	7, 23	2	B1	CO1
8	8, 24	2	B1	NL1
9	9, 25	3	A1	CO1
10	10, 26	3	A1	NL1
11	11, 27	3	B1	CO1
12	12, 28	3	B1	NL1
13	13, 29	4	A1	CO1
14	14, 30	4	A1	NL1
15	15, 31	4	B1	CO1
16	16, 32	4	B1	NL1

Note. List assignment, and order of list halves and of task condition are given for each participant in Experiment 5. A1 = list half A presented first; B1 = list half B presented first; CO1 = counting condition presented first; NL1 = no-load condition presented first.

Like in the previous experiment, lists were divided in two halves, one for the no load and one for the counting condition. Order of presentation of the list halves and of the articulatory-suppression conditions was counterbalanced across participants. Assignment to lists and counterbalancing of list halves and articulatory-suppression conditions yielded 16 participant groups. Participants were assigned to groups randomly. An overview of group characteristics is given in Table 3.

Procedure. Participants read sentences that appeared on a computer screen. Sentences were presented as a whole and stayed on the screen until the participant made a judgment. Participants could respond by pressing a *correct* button if the sentence was correct or by pressing an *incorrect* button if the sentence was not. Sentences were preceded by a warning signal (*) that disappeared after the participant had pressed the *correct* button. After each trial, participants received feedback about their judgment. Materials were presented in two blocks: a no-load block and a counting block. Each block was preceded by a practice session to make the participants familiar with the judgment task and the articulatory-suppression task.

Results

RT Analysis. The RT data from all participants were submitted to standard descriptive and inferential analysis. First, the raw data of the two list halves of each participant were checked for outliers. RTs shorter than 200 ms were taken to reflect errors with the response boxes unrelated to sentence processing and were not included in the analysis. The absolute upper cut-off was 9,000 ms. Means and standard deviations were computed per list half in order to determine the relative cut-offs for each participant. The relative lower cut-off was the mean RT minus 2.5 standard deviation; the relative upper cut-off was the mean RT plus 2.5 standard deviation. Data points that were below the lower cut-offs (both absolute and relative) were replaced by the relative lower cut-off; values that exceeded the higher cut-off points were replaced by the upper relative cut-off.

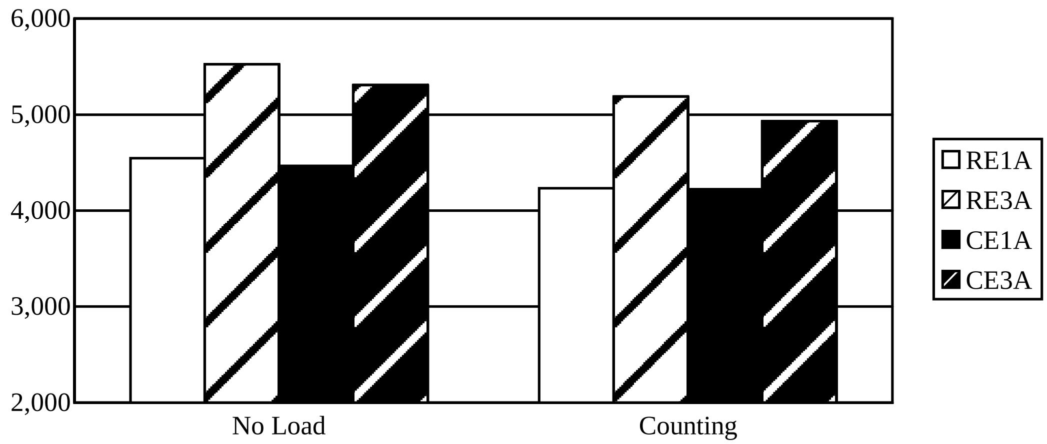


Figure 1. Mean RTs in ms for right-embedded sentences with one adjective (RE1A), right-embedded sentences with three adjectives (RE3A), center-embedded sentences with one adjective (CE1A), and center-embedded sentences with three adjectives (CE3A) in the no load (left) and counting conditions (right).

After outlier detection, means and standard deviations per condition were calculated over subjects and items. Mean RTs per condition are given in Figure 1.[2] For the subject analysis (*F1*), an analysis of variance (ANOVA) was performed with the following within-subject factors: Adjectival Load (1 vs. 3), Embedding Type (right vs. center embedded), and Articulatory Suppression (no load vs. counting). List, List-Half Order, and Task Order were the between-subject factors. For the item analysis (*F2*), an ANOVA was performed which had Adjectival Load, Embedding Type, and Articulatory Suppression as within-item factors and Item Group as a between-item factor.

The ANOVA revealed significant a main effect of Adjectival Load with faster RTs for the sentences containing one adjective than for those with three adjectives, $F1$ (1, 28) = 38.64, $p < .001$; $F2$ (1, 39) = 84.48, $p < .001$. There was also a main effect of Embedding Type, with shorter response times for center-embedded clauses than for right-embedded sentences, $F1$ (1, 28) = 6.73, $p < .05$; $F2$ (1, 39) = 5.27, $p < .05$. The ANOVA revealed no main effect of Articulatory Suppression, $F1$ (1, 28) = .77, $p = .39$; $F2$ (1, 39) = 9.80, $p < .01$.

In addition to the main effects, there was a significant three-way interaction, $F1$ (1, 28) = 5.00, $p < .05$; $F2$ (1, 39) = 4.28, $p < .05$. When divided into simple effects, right-embedded clauses were responded to more slowly than center-embedded clauses in the no-load condition, $F1$ (1, 28) = 10.71, $p < .01$; $F2$ (1, 40) = 7.59, $p < .01$, but not in the counting condition, $F1$ (1, 28) = .71, $p = .41$; $F2$ (1, 40) = .51, $p = .48$. This effect of embedding in the no load condition was observed in the one-adjective sentences, $F1$ (1, 28) = 243.38, $p < .001$; $F2$ (1, 40) = 498.04, $p < .001$, but not in the three-adjective sentences, $F1$ (1, 28) = .83, $p = .38$; $F2$ (1, 40) = 1.69, $p = .20$.

Error Analysis. The error data were analyzed in the same way as the RT data. Mean error proportions are given in Figure 2.[3] No significant effects or interactions were found.

In Experiments 1 to 3, there was a response bias in the articulatory-suppression conditions. Therefore, a Wilcoxon Matched Pairs Signed Rank Test was carried out. It compared the proportion of errors for the implausible fillers in the no-load condition with the error proportion for the implausible fillers of the articulatory-suppression condition. There was no significant difference between the counting and no-load condition ($Z = -.79$, $p = .43$).

---

[2] In order to facilitate comparison of these data with the RTs of Experiments 1 to 4 and 5b, the scale of each figure was set between 2,000 and 6,000 ms.

[3] In order to facilitate comparison of these data with the error proportions of Experiments 1 to 4 and 5b, the scale of each figure was set between the error proportions of 0 and .50.

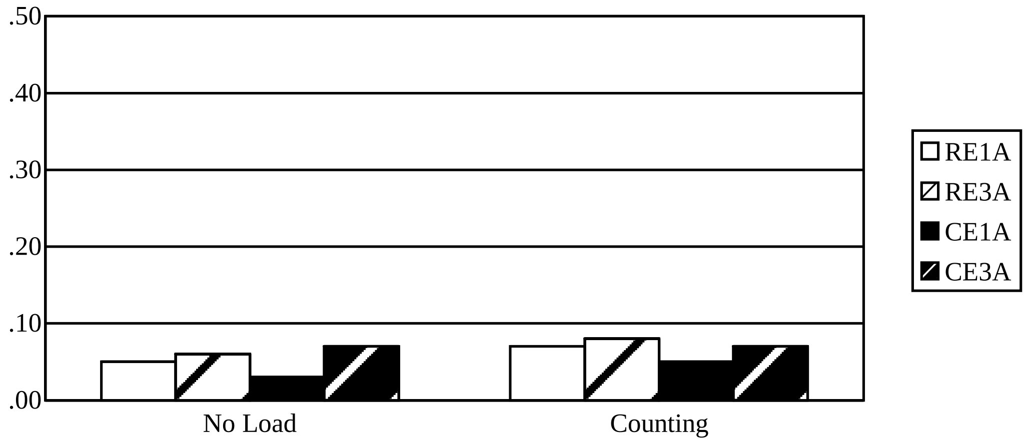


Figure 2. Mean error proportions for right-embedded sentences with one adjective (RE1A), right-embedded sentences with three adjectives (RE3A), center-embedded sentences with one adjective (CE1A), and center-embedded sentences with three adjectives (CE3A) in the no load (left) and counting conditions (right).

Discussion

Three Hypotheses. So far, three hypotheses about the role of a phonological back-up representation have been proposed:

1. The phonological back-up representation is addressed during long-distance integrations in general;
2. the phonological back-up representation is addressed during long-distance θ-role assignment; and
3. the phonological back-up representation is addressed during long-distance feature checking for subject-verb agreement.

Under the first hypothesis, a superadditive three-way interaction is predicted because both center-embedding and stacked adjectives require long-distance integration, although not necessarily at the same time (cf. Chapter 4).

The second hypothesis predicts a super-additive two-way interaction of Embedding type x Articulatory Suppression, because in three-adjective sentences, adjective-noun agreement must be checked over a long distance. However, this hypothesis does not predict a super-additive two-way interaction of Adjectival Load x Articulatory Suppression. The θ-role assignment must be established over a long distance, and therefore interactions of Embedding Type x Articulatory Suppression are expected. However, since the hypothesis states that the use of the back-up is restricted to θ-role assignment only, no interactions of Adjectival Load x Articulatory Suppression are expected to occur.

Finally, the third hypothesis predicts only main effects of Adjectival Load and of Articulatory Suppression. Although the integration of subject and verb must be accomplished over a long distance, the subject-verb agreement features can be checked locally due to the use of an auxiliary, therefore no interactions are predicted.

No Support for Any of the Hypotheses. Given the fact that the first hypothesis predicted a three-way interaction and the others did not, one can conclude that the second and third hypotheses are not supported by the current data pattern. However, although the first hypothesis did predict a superadditive three-way interaction, closer inspection of the data pattern reveals that the interaction is non-additive in nature. Therefore, the results of Experiment 5a do not shed any light on which hypothesis is most appealing.

An Explanation. In the RT analysis, there were two main effects. Sentences with three adjectives were read more slowly than sentences with one adjective. This effect can be attributed to two factors. First, three-adjective sentences have two more words than one-adjective sentences. Since the three-adjective sentences are longer, it is probably due to sentence length that they yielded higher RTs. Second, in three-adjective sentences, long-distance adjective-noun integration must be accomplished, whereas the adjective-noun integrations in the one-adjective sentences are local. As already pointed out in the introduction to the current chapter, long-distance integrations are more demanding to WM than local integrations (Gibson, 1998; Martin & Romani, 1994).

Second, there was a main effect of Embedding Type: right-embedded sentences were responded to more slowly than center-embedded ones. The nature of this effect can only be understood if one considers the three-way interaction in the RT analysis: right-embedded sentences yielded longer response latencies in the no-load condition of the one-adjective sentences only.

This finding is entirely unexpected: in Experiments 1 to 4, two-way interactions of Embedding Type x Articulatory Suppression were found, but center-embedded sentences always yielded longer response latencies. And more crucially, effects of embedding were present under articulatory suppression only. In the current experiment, an entirely opposite pattern emerges: the absence of embedding effects in the articulatory-suppression condition of the one-adjective sentences and the presence of a reversed embedding effect in the no-load condition of the one-adjective sentences.

In Chapters 3 and 4, it was proposed that during the processing of center-embedded constructions, a phonological back-up representation was deployed via inner speech. It was also proposed that articulatory suppression hinders the processing of center-embedded sentences because it impedes the full deployment of inner speech (Baddeley, 1986; Gathercole & Baddeley, 1993).

Taking these considerations into account, it is difficult to account for the simple interaction of Embedding Type x Articulatory Suppression in the one-adjective sentences in terms of processing difficulty. Therefore, it may be worth looking at the materials in more detail. How can right-embedded sentences, as in (9a), repeated here as (9a), take longer to judge than center-embedded ones, as in (11b), repeated here as (11b)?

(11) a. De barkeeper heeft de chips opgegeten terwijl de verregende klant niets doorhad
The barkeeper has the chips eaten while the drenched customer nothing noticed
'The barkeeper ate the chips while the drenched customer did not notice anything'

b. De barkeeper heeft, terwijl de verregende klant niets doorhad, de chips opgegeten
The barkeeper has, while the drenched customer nothing noticed, the chips eaten
'The barkeeper ate the chips while the drenched customer did not notice anything'

Closer inspection of these sentences suggests that the effect may be task related: in (11a), the point of judgment comes later. Remember that the distractor sentences had implausible noun-adjective combinations, as shown in (10), repeated here as (12):

(12) # De zuster heeft, terwijl het zwemmende gordijn steeds openbleef, het kussen opgeschud
The nurse has, while the swimming curtain all-the-time open remained, the pillow plumped
'The nurse plumped the pillow while the swimming curtain remained open all the time'

The noun-adjective combinations were always in the embedded clause, which comes earlier in the sentence for center-embedded structures than for right-embedded ones.

The current results suggest that in the no-load condition, readers make a judgment as soon as they establish that the adjective and the adjacent noun are incompatible. It is not exactly clear why readers apply such a strategy in the no-load condition of the one-adjective sentences only, and not in the articulatory-suppression condition, although it may be easier to use such a strategy in simpler sentences.

The fact that the effects in the current experiment seem to reflect the use of task-related strategies implies that they do not provide conclusive evidence with respect to the main goals of the experiment.

First, checking of agreement features and θ-role assignment were separated in the current experiment by using an auxiliary verb so that it could be investigated whether inner speech was used during θ-role assignment or the checking of agreement features. Due to the task effect, one cannot tell whether superadditive interactions of Embedding Type x Articulatory Suppression can be found in relative clauses containing an auxiliary.

A second reason for conducting the experiment was to find out whether the interactions of Articulatory Suppression with subject-verb integration were interactions related to the process of subject-verb integration or to integration in general. The interactions in the current experiment do not shed any light on this issue since they seem to be strategy related. Therefore, a replication of the current experiment was conducted in order to see whether the same results would occur, or whether a different pattern with less prominent task effects would emerge.

In order to prevent participants from applying the same task strategy again, different distractors were used. In the current experiment, semantically unacceptable foils contained sentences that had violations early in the sentence. This may have enhanced the task strategy of judging as soon as the violation appeared.

In Experiment 5b, different distractor types were chosen. The target sentences and foils of Experiments 3 and 4 were used as unrelated distractor sentences for the current experiment. The sentences in these experiments had the point of judgment at the very end of the sentence. By adding these sentence types to the materials of the current experiment, I tried to discourage the use of an immediate judgment strategy.

Experiment 5b

Method

Subjects. Thirty-two undergraduate students at the University of Groningen participated in the experiment; they were paid participants. All participants had normal or corrected to normal vision and were native speakers of Dutch. Fourteen participants were male; 18 were female. The mean age was 20.2 years old; the age range was between 18 and 27. These participants had also participated in Experiments 3 and 4.

Method and Procedure. The current experiment used the same target sentences as Experiment 5a. In addition to the 48 target sentences, 168 filler sentences were included. These consisted of 24 anomalous distractors (sentences like the targets, but with implausible adjective-noun combinations), the (plausible) target sentences of the

Experiments 3 and 4 (96 in all), and 48 implausible θ-role inversions of the targets for Experiments 3, 4, and 5a. The same procedure was used as in Experiment 5a.

Results

RT Analysis. The RTs of the current experiment were checked for outliers and analyzed using the same procedure as in Experiment 5a. Mean RTs are given in Figure 3.[4] For the subject analysis (*F1*), an analysis of variance (ANOVA) was performed with the following within-subject factors: Adjectival Load (1 vs. 3), Embedding Type (right vs. center embedded), and Articulatory Suppression (no load vs. counting). List, List-Half Order, and Task Order were between-subject factors. For the item analysis (*F2*), an ANOVA was performed which had Adjectival Load, Embedding Type, and Articulatory Suppression as within-item factors and Item Group as a between-item factor.

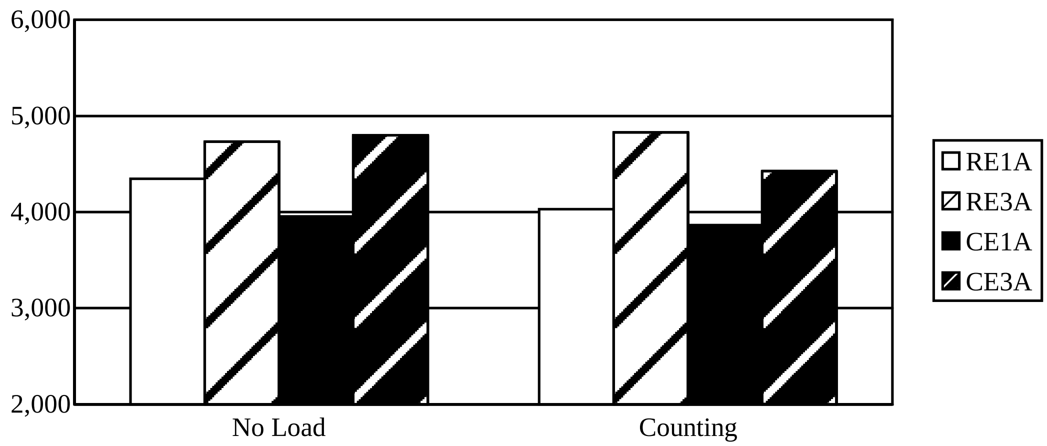


Figure 3. Mean RTs in ms for right-embedded sentences with one adjective (RE1A), right-embedded sentences with three adjectives (RE3A), center-embedded sentences with one adjective (CE1A), and center-embedded sentences with three adjectives (CE3A) in the no load (left) and counting conditions (right).

There was a significant effect of Adjectival Load, *F1* $(1, 16) = 96.99$, $p < .001$; *F2* $(1, 42) = 36.30$, $p < .001$): the sentences with one adjective were responded to more quickly than the three-adjective sentences. There was also a main effect of Articulatory Suppression with slower responses in the no-load condition than in the

---

[4] In order to facilitate comparison of these data with the RTs of Experiments 1 to 5a, the scale of each figure was set between 2,000 and 6,000 ms.

articulatory-suppression condition, *F1* (1, 16) = 8.69, $p < .001$; *F2* (1, 42) = 7.19, $p < .05$. No main effect of Embedding Type was found, *F1* (1, 16) = 1.26, $p = .28$; *F2* (1, 42) = 1.86, $p = .18$.

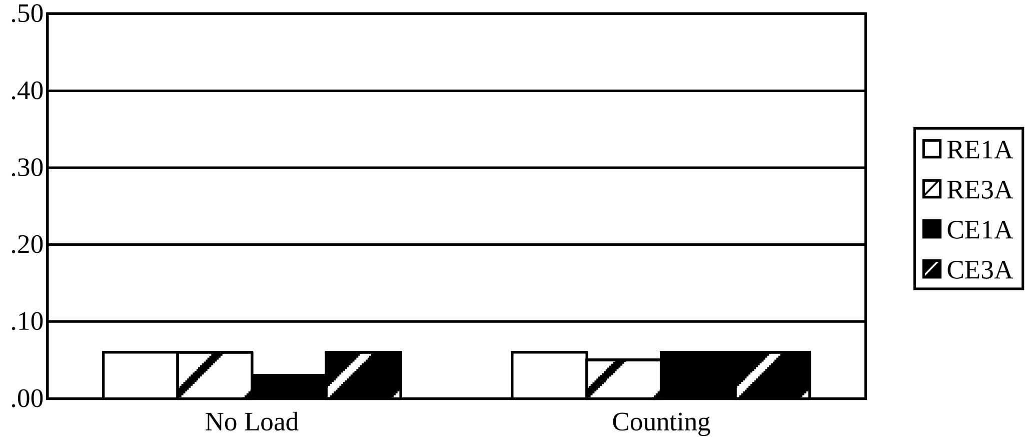


Figure 4. Mean error proportions for right-embedded sentences with one adjective (RE1A), right-embedded sentences with three adjectives (RE3A), center-embedded sentences with one adjective (CE1A), and center-embedded sentences with three adjectives (CE3A) in the no load (left) and counting conditions (right).

Error Analysis. The error data were analyzed in the same way as the RT data. Mean error proportions for the different experimental conditions are given in Figure 4.[5] No significant main effects or interactions were found.

In Experiments 1 to 4, there was a speed-accuracy trade-off with Articulatory Suppression. Therefore, a Wilcoxon Matched Pairs Signed Rank Test was carried out. It compared the proportion of errors for the implausible sentences in the no-load condition with the error proportion for the implausible sentences of the articulatory-suppression condition. There was a significantly higher proportion of errors for filler sentences in the articulatory-suppression condition than in the no-load condition ($Z = -2.46$, $p < .05$).

Discussion

The RT analysis yielded two main effects: one of Adjectival Load, with slower RTs for the three-adjective sentences, and one of Articulatory Suppression, with faster

[5] In order to facilitate comparison of these data with the error proportions of Experiments 1 to 5a, the scale of each figure was set between the error proportions of 0 and .50.

RTs for the articulatory-suppression condition. There were no interactions. These results have a number of implications.

Comparing Experiments 5a and 5b. It is important to compare the current findings with Experiment 5a. Experiment 5a produced a three-way interaction with faster RTs for center-embeddings in the no-load one-adjective sentences only. It was argued that this interaction was the result of a task strategy. In the no-load condition, participants seemed to make the judgment as soon as they assessed the plausibility of the adjective-noun combination without finishing the whole sentence.

The current experiment, which used fillers that had the violation at the end of the sentence, did not elicit this interaction. Had the three-way interaction in Experiment 5a been related to language processing or to the deployment of WM resources, these exact same results should have shown up in the current experiment. This is not the case however. Therefore, attributing these effects found in Experiment 5a to task strategies applied by readers seems justified.

Interpretation of the Effects. No effects of Embedding Type were found in Experiment 5b. In Experiments 1 to 4, interactions of Embedding Type x Articulatory Suppression were found; these interactions suggest two things. First, they suggest that during the processing of center-embedded sentences, subject-verb agreement features had to be checked and second that this checking procedure made use of a phonological back-up kept active via inner speech. In Experiment 5b, which separated subject-verb integration and θ-role assignment, no effects of Embedding Type were found. This pattern indicates that articulatory suppression impedes long-distance subject-verb integrations, but not long-distance θ-role assignment. In the previous chapters, where the main verbs had agreement relations with the subjects, interactions of Embedding Type x Articulatory Suppression were consistently found. The current experiment, which dissociated subject-verb agreement and θ-role assignment, failed to find such an interaction.

Second, the lack of an interaction of Adjectival Load x Articulatory Suppression suggests that the interactions with Articulatory Suppression found in Chapters 3 and 4 were not interactions with just any type of integrations. Instead, they seem to reflect the involvement of a phonological back-up representation during subject-verb integration.

Third, the current experiment replicated the reversed load effect that was also found in Experiments 1 to 4 reported in the previous chapters. Readers were faster in the counting condition than in the no-load condition. This supports the idea that Dutch readers speed up when under articulatory suppression, probably because they want to arrive at the clause-final embedded verb quickly so that integration can take place.[6]

General Discussion

The effects found in Chapters 3 and 4 and in Experiments 5a and b suggest that a phonological back-up representation is addressed during sentence comprehension. It seems to be used for the checking of subject-verb agreement features during subject-verb integration. This checking procedure is hampered by articulatory suppression if the agreement features must be checked over a long distance. In Experiments 5a and 5b, which used sentences that required only local checking of subject-verb agreement features, no interactions with Articulatory Suppression were found.

Evidence for a Utilitarian Deployment of the Phonological Back-Up

In Experiments 5a and b, there were no two-way interactions of Adjectival Load x Articulatory Suppression. This suggests that checking of agreement features occurs during subject-verb integration, but not during the integration of adjectives and nouns. This is interesting given the fact that in Dutch, adjectives have an agreement relation with their head noun.

Motivating the Assumption of a Utilitarian Deployment. This finding implies that readers do not use an automatic checking strategy: they seem to check agreement features only when they can benefit from it. Sentence interpretation is dependent on building the correct syntactic structure. If this assignment hinges on thorough processing of agreement features, these features are checked in the phonological back-up representation. However, the correct processing of adjective-noun agreement features is not crucial for sentence interpretation, since agreement does not choose between alternative structures. This pattern suggests that readers are utilitarian in what strategies they use during sentence comprehension.

This explanation is also in line with proposals about the role of the phonological loop presented by Baddeley et al. (1987) and Caramazza, Basili, Koller, & Berndt (1981). These researchers argue that phonological memory is addressed in cases where full syntactic analysis is required for understanding of the sentence.

---

[6] Note that although no subject- vs. object-relative comparison was tested, long-distance checking of subject-verb agreement within the relative clause was still necessary.

Alternative Explanations. An alternative reason for the lack of a two-way interaction of Adjectival Load x Articulatory Suppression may be that the adjectival complexity is not resolved by addressing a phonological back-up, but by using a lexicosemantic WM system. This idea has been advocated by Martin & Romani (1994). Their proposal about the organization of WM resources that play a role in language comprehension will be outlined below.

However the data from Experiment 5b are open to a third interpretation as well. In Experiment 5, all adjectives agreed with the head noun. Whether the adjective-noun integration requires a local or a long-distance feature checking process is not entirely clear.

The first adjective and the head noun are far apart; this suggests that a long-distance feature check may be necessary. However, the first adjective has the same agreement features as the second and the third adjective. It may well be that the agreement features of the three adjectives are checked locally. Likewise, the agreement features of the third adjective and the head noun may be checked locally.

Therefore, it may be that the lack of interactions with Adjectival Load are due to the fact that the assumed long-distance feature checking was in fact established by making multiple local comparisons.

Separate WM Resources for Lexical Semantics?

Martin & Romani (1994) proposed a WM system for language comprehension that is an extension of Baddeley's (1986) WM model. Baddeley's model consists of a central component that is involved in the temporary storage and processing of information. This module is referred to as the *central executive* of WM.

In addition to the central executive Baddeley (1986) assumes at least two so-called *slave systems,* one for phonological storage, called the *phonological loop,* and one for visual storage, called the *visuo-spatial sketchpad.* The central executive calls on these slave systems if it needs additional storage capacity. For a more detailed description of Baddeley's WM model, the reader is referred to page 20.

On the basis of a double dissociation between two patients, E.A. and A.B., Martin & Romani (1994) propose that Baddeley's (1986) WM model must be extended with a third slave system that is concerned with lexicosemantic storage. E.A. is a patient who has intact semantic storage, but disrupted phonological memory. A.B., however, is a patient with intact phonological memory and a disturbance of semantic memory.

E.A. is assumed to have an impairment of phonological WM because she is poor at rhyme probe tasks. With rhyme probe tasks, subjects listen to a list of words that is followed by a probe. The subjects must judge whether the probe rhymes with one of the words of the list that preceded the probe. Furthermore, she showed worse memory for nonwords, which must be remembered in a phonological format, than for words.

She has relatively intact semantic memory as demonstrated by her good performance on a category probe task. With category probe tasks, participants must listen to lists of words that are followed by a probe word. They must judge whether the probe word is of the same semantic category as one of the words in the list.

A.B. is a patient with an opposite pattern of impairment. He is taken to have a disturbance of semantic memory because of his poor performance on the category probe tasks. His phonological WM capacity is relatively spared as he is equally good at remembering words and nonwords. This suggests that his inner-speech loop is still functioning well.

E.A. and A.B. performed the same plausibility decision test with one- and three-adjective sentences as described in the introduction. E.A., the patient with the phonological memory impairment, showed the same pattern as the normal controls: worse performance for adjectives that preceded the noun than for adjectives that followed it. Also, the distance of adjective-noun integration had more impact in the premodification conditions than in the postmodification conditions.

This finding suggests that during the integration of adjectives and nouns, phonological memory plays no substantial role. The outcome of Experiments 5a and 5b suggest that this finding also holds for Dutch, a language that has adjective-noun agreement, unlike English.

A.B., however, performed much worse than normal controls and than E.A. on the experiment. Especially if adjectives preceded the noun, his performance decreased rapidly with distance between adjective and noun. This pattern argues for a role of a lexicosemantic WM system during the comprehension of stacked adjectives.

The pattern of performance reported by Martin & Romani (1994) and the independent effects of Adjectival Load and Embedding Type in Experiment 5b suggest that the processing of adjective-noun integration and that of subject-verb integration are carried out by separate systems. The first complexity is probably carried out by a semantic and the other by a syntactic processing system.

Summary

In this chapter, two experiments were reported that looked at the possible role of a phonological back-up representation during the processing of adjective-noun integrations and of long-distance θ-role assignments not requiring the checking of agreement features (in center-embedded constructions). Lack of two-way interactions of Embedding Type x Articulatory Suppression and of Adjectival Load x Articulatory Suppression suggest that a phonological WM system plays a prominent role in neither of these processes.

Instead, it is proposed that phonological memory is particularly important during the checking of agreement features between subject and verb, but less in cases where these features must be checked locally. The integration of adjectives and nouns seems to be mediated by a different WM system, one specifically for lexicosemantic information. However, it may also by that the lack of interactions with Adjectival Load is due to the fact that multiple local checks were carried out during the processing of stacked adjectives.

Chapter 6
On the Predictive Power of Span Measures

Introduction

In the discussion about the role of working memory (WM) in sentence comprehension, the predictive value of span measures is an old chestnut. Some researchers claim that span measures reflect WM processing capacity as tapped by all possible kinds of language tasks (e.g., Just & Carpenter, 1992; King & Just, 1991); others vehemently oppose to this view (e.g., Waters, Caplan, & Hildebrandt, 1987; Caplan & Waters, 1999).

Just & Carpenter (1992) assumed that a single WM system is involved in language processing. They also assume that people vary with respect to the capacity of that WM system. The WM capacity of people can be measured by tasks that combine storage and processing of information. Just & Carpenter claim that especially span measures that tap these two aspects are good predictors of language processing capacity.

Caplan & Waters (1999) and Waters et al. (1987) proposed a two-stage processing model in which interpretive or syntactic processing and postinterpretive or postsyntactic processing are separate. They claimed that span measures, simple and complex alike, correlate with postsyntactic, but not with syntactic processing.[1]

The current chapter links up with the debate about span measures and language processing since it is devoted to correlations of two span measures with the effects in the Experiments 1 to 5b. The span measures that will be looked at are a Dutch version of the Salthouse Listening Span Test and a nonword span test.

[1] For a more detailed description of their model, the reader is referred to page 25.

Outline

The outline of the current chapter is as follows: I will commence the chapter with an introduction. This introduction covers the use of correlation analysis, an overview of different span measures, a discussion about correlations of span measures with language comprehension skills, and a subsection on how the span data collected in Experiments 1 to 5b tie in with this debate. The introduction is succeeded by a section, in which the two span measures will be described. In the third section, an overview of correlations with the two span measures will be given. Finally, the correlation data will be related to the debate about the predictive abilities of span measures with respect to language processing.

Correlation Analysis

Correlation analysis is a means of investigating whether certain tasks are good predictors of certain cognitive processes. Suppose one wants to know whether memory span for words is a good predictor of reading comprehension.

The first thing a researcher has to do is administer the memory span. Let us say for the sake of the argument that the researcher gradually increases the number of words appearing in a list and memory span is defined as the longest list length a subject is able to recall correctly.

In order to see whether this particular memory span is a good predictor of reading efficiency, one has to administer a comprehension task with questions to all subjects who have taken the span task.

Then one can perform a correlation analysis on how well the memory span correlates with the score on the comprehension questions. The correlation value lies within a range from –1 to 1. A negative correlation means that high memory spans goes along with poor performance on the reading task; a positive correlation means that people with high memory spans are good readers. Correlation values that are close to 0 indicate that the two measures are poor predictors of each other.

One has to take two issues into account with respect to correlation analyses. First, one needs enough spread in the variables under consideration. Second, a significant correlation of two factors does not necessarily imply that one factor is caused by the other or vice versa. These two issues will be explained next.

Spread in the Data. Suppose – to return to our memory span example – that the following span sizes are possible: 2, 3, 4, 5, 6, and 7. With this particular distribution, one can look for correlations with span in three ways.

The easiest thing to do is to correlate a certain complexity measure with the span sizes obtained from the subjects in the experiment. Thus, one has six different groups of subjects (viz. subjects with span sizes 2 up to 7).

A second option is to divide the subjects into high and low-spanners, with the low-span group ranging from 2 to 4 and the high-span group ranging from 5 to 7. By doing that, one reduces the spread in the data considerably and chances of finding a significant correlation are small. (Without dividing subjects up in low and high spanners, we had six span groups; now there are only two span groups left.) By lowering one's chances of finding effects, one increases the chance of a Type I error; this may be considered a serious disadvantage of this approach.

The third way to divide up subjects in a correlation analysis is to only include those subjects who have extreme scores. In this case that would mean that one compares subjects with span sizes 2 and 3 subjects who had scores of 6 and 7. However, Figure 2 shows that the data points at the extreme ends (the round data points) usually tend to lie much closer to the regression line than the data points in the middle (the square data points). Therefore, if one chooses to apply this method, one ought to perform a statistical correction on the obtained correlation value.

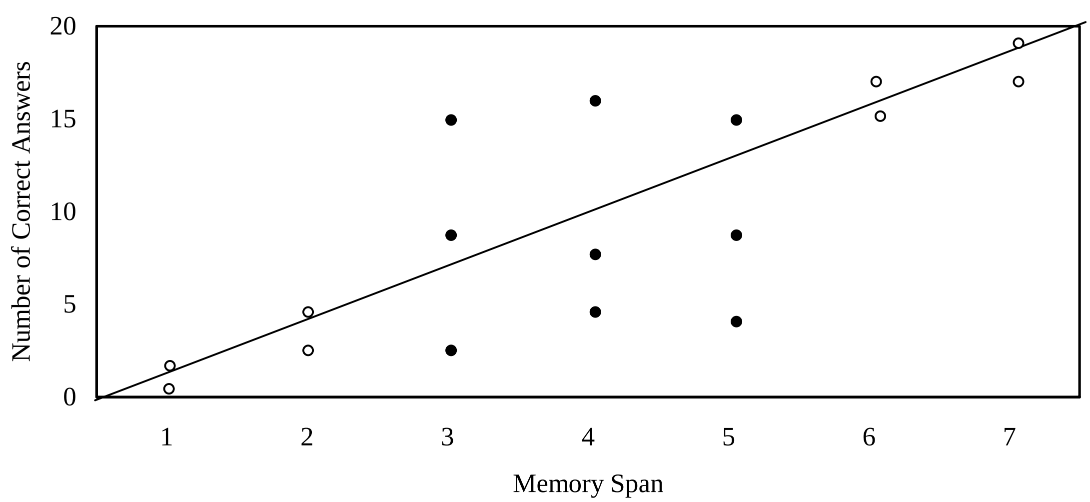


Figure 1. Scatterplot of the outcome of the hypothetical experiment mentioned above.

The Causality Issue. Apart from the issue of dividing subjects up in groups, one has to be cautious when it comes to drawing conclusions from significant correlations. When two factors correlate well with each other, it need not be the case that there is a causal relation between them. For example, the incidence rate of polio is highly correlated with the amount of soft drinks sold per month (Freedman, Purves, & Pisani, 1998). However, it is hard to find medics who claim that drinking soft drinks causes polio. It turns out that polio occurs more often when the weather is hot. Under high temperatures people tend to drink more soft drinks. Drinking soft drinks is not the cause of polio, but the two phenomena share a common cause, the hot weather. In other words, significant correlations do not necessarily imply a causal relation.

An Overview of Span Measures

Research from several fields has suggested that certain span measures correlate with language processing abilities. A number of studies have found that brain damaged patients with poor memory spans suffer from language comprehension problems (e.g., Baddeley, Vallar, & Wilson, 1987; Caplan & Hildebrandt, 1988; Caramazza, Basili, Koller, & Berndt, 1981; Martin & Romani, 1994; Saffran & Marin, 1975). In addition, several studies have found effects of memory span on language comprehension in healthy populations (e.g., Daneman & Carpenter, 1980; King & Just, 1991; Vos, 1999; Waters et al., 1987).

Within the WM literature, three types of WM tasks are distinguished. Several researchers have proposed a distinction between verbal WM tasks on the one hand and spatial WM tasks on the other (e.g., Baddeley, 1986; Gathercole & Baddeley, 1993; Shah & Miyake, 1996). In addition, distinctions have been proposed between verbal and computation spans (Daneman & Carpenter, 1980; Daneman & Merikle, 1996; Salthouse, 1994).

In this overview, verbal span measures receive most attention because they are taken to be good predictors of (certain aspects of) language comprehension (Caplan & Waters, 1999; Just & Carpenter, 1992). Within the group of verbal span measures, two subgroups can be distinguished: simple and complex span measures. These different groups of span measures will be discussed in the current subsection. The next subsection discusses what these span tasks are taken to measure. In general, it is assumed that simple span measures test peripheral storage capacity whereas complex span measures reflect the central aspects of WM. In that subsection however, it will become apparent that the distinction between simple and complex span measures is not as straightforward as initially seems (La Pointe & Engle, 1990).

Simple Verbal WM Measures. A number of verbal tasks are often used in WM research. One common measure of verbal WM is *digit span.* Digit span tasks consist of lists of increasing length that contain randomly ordered digits. The digits are read aloud to the participant; after completion of a list, the participant’s task is to recall the right digits in the right order. Digit span size is usually determined as the largest list size a participant can recall correctly.

A second WM measure, *word span*, is similar to digit span except that words are used instead of digits. These words can be drawn from a small pool of words and be re-used in the different lists in the span task or be taken from a large pool, so that no repetition of items occurs, unlike digit span.

Yet another measure of WM is *nonword span.* It was specifically designed to test memory capacity for items that cannot be remembered by their meaning since the lists consist of nonsense words, which do not have any meaning. Nonword span is taken to reflect phonological storage capacity (Baddeley, 1986; Gathercole & Baddeley, 1993).

Complex Verbal WM Measures. It is generally assumed that WM comprises both storage and processing of information (e.g., Baddeley, 1986; Caplan & Waters, 1999; Daneman & Carpenter, 1980; Gibson, 1998; Salthouse, 1993, 1994). The simple span measures mentioned above tap the storage component of WM, but not the processing aspect. Several WM measures have been designed specifically to tap both aspects of WM, these are called *complex span measures.*

One of the most widely used complex span measures is the Reading Span Task developed by Daneman & Carpenter (1980). In this task, subjects are required to read series of sentences aloud. Meanwhile, they must remember the last word of each sentence in a particular series. Sentences are presented on cards; there is one sentence on each card. At the end of a series, a blank card appears. This blank card is the participant's cue for recall: at this point (s)he has to recall all sentence final words of the series in the right serial order.

Several variants of the Reading Span Test have been developed in the last decades of the twentieth century. These span measures share the fact that they tap both the storage and processing of information. Daneman & Carpenter (1980) developed an auditory version of the Reading Span Test, the Listening Span Test. In this task, subjects must listen to sentences and recall the sentence final words of the sentences in a particular series. In the mean time, subjects must answer yes/no questions about the sentences heard.

Caplan and coworkers have people read sentences on a computer screen, some of which are semantically correct, some of which are not (Caplan & Waters, 1999; Waters et al., 1987). The participant's task is to judge each sentence and remember its last word. At the end of a series, subjects must recall these words in the order of presentation.

A third variant of the Reading Span Test requires subjects to answer comprehension questions of sentences they hear on a tape while they have to maintain the final words of the sentences in a series (Daneman & Carpenter, 1980).

Nonverbal WM Measures. In addition to complex WM measures that tap verbal WM capacity, there are also complex WM measures that tap spatial and arithmetic skills. Since these span measures fall beyond the scope of this thesis, they will not be discussed here; for a more extensive overview of these span measures, the reader is referred to Salthouse (1990, 1994).

What do Language-Related Span Tasks Measure?

The different span measures mentioned above are obviously designed to measure WM capacity. Given the plethora of sorts of span, these measures are likely to tax different aspects of WM. The current subsection will cover the claims made over the last three decades with respect to the validity of the different measures.

Most of these claims are based on WM models along the line of Baddeley's (1986). His model consists of a central processor, the *central executive*, which is involved in the storage and processing of information. It has the ability to deploy so-called *slave systems* when it needs additional storage capacity. One of these slave systems is the *phonological loop,* a kind of inner speech loop. The phonological loop itself can be subdivided into a *phonological store* and an *articulatory rehearsal mechanism.* For a more detailed description of Baddeley's model, the reader is referred to page 20.

Simple Span Measures Tap Phonological-Loop Capacity. The simple spans are generally taken to measure the phonological loop. The most pure measure of phonological-loop capacity is nonword span. This span taps maintenance of information that cannot be stored in a format other than a phonological representation because phrase structure building or semantic storage are not possible with nonsense words.

The same thing goes, but to a slightly lesser extent, for the digit span tasks. Although digits have some meaning, when presented in the context of other digits only, it is hard to use their meaning as a cue for storage. The main difference between nonword spans and digit spans is that digits can be stored in a lexical format, but nonwords cannot.

Word spans are also taken to measure phonological-loop capacity. The reason for assuming phonological-loop involvement in the storage of lists of words is that effects arising in the phonological loop, such as the word length effect, can be observed in word spans. Participants have higher spans for words with a short duration than for long words. A second reason for the assumption that word span taps the phonological loop is that this effect disappears under articulatory suppression, which is taken to be a phonological-loop load. When participants must utter irrelevant speech sounds while remembering the lists, the advantage of short words over longer items goes away (Baddeley, 1986; Gathercole & Baddeley, 1993; La Pointe & Engle, 1990).

Complex Span Measures Tap Central-Executive Capacity. Several researchers have argued that the central executive in Baddeley's (1986) model carries out processes like syntactic analysis and interpretation at the pragmatic level (e.g., Gathercole & Baddeley, 1993; Just & Carpenter, 1992; Waters et al., 1987). Under this assumption, one would expect that WM measures tapping central-executive processing are good predictors of language processing capacity.

Daneman & Carpenter (1980) claim that simple span measures are poor predictors of language skills. They found that these measures correlated poorly with measures of language competence such as Verbal Scholastic Aptitude Test (VSAT) scores. However, they found high correlations of complex span measures, such as the Reading Span Test, with VSAT scores.

The assumption that complex span measures measure central-executive processing has been adopted by many other researchers as well (e.g., Baddeley, 1986; Just & Carpenter, 1992; Salthouse, 1993; Waters et al., 1987). In addition, complex span measures have been found to be sensitive to several linguistic complexity effects that were investigated in reaction time (RT) experiments (Caplan & Waters, 1999; King & Just, 1991; Vos, 1999).

Although these researchers agree on the fact that complex spans measure central-executive processing, there is considerable debate about what aspects of central-executive processing it is sensitive to. Carpenter and colleagues argue that it is sensitive to all central-executive processing (Daneman & Carpenter, 1980; Just & Carpenter, 1992). This claim stems from Just & Carpenter's assumption that a single WM system mediates language comprehension. If there is a measure that taps this unitary system, it must correlate with all processes carried out by that system.

However, Caplan & Waters (1999), argue that these measures are sensitive to processes that occur after syntactic analysis has been accomplished, but not to the syntactic processes themselves. In their sentence processing model, there are two processing stages: one for processes that occur up to the syntactic level, the *interpretive* processes, and one for semantic integration processes that occur afterwards, the *postinterpretive* processes. The assumption of two processing stages stems from the Berwick & Weinberg (1984) parsing model. For a detailed description of Berwick & Weinberg's and Caplan & Waters's model, the reader is referred to pp. 15 and 25 respectively.

Simple vs. Complex Spans and Peripheral vs. Central Storage: Too Crude a Distinction? The overview of span measures above seems to point to a clear-cut dichotomy of simple tasks that measure peripheral storage in the phonological loop and of complex span measures that tap the central executive. However, the picture is more complicated than that.

Daneman & Carpenter (1980) indeed found poor correlations of word span and language skills, but high ones of reading span task with language skills. However, La Pointe & Engle (1990) argued that reading span tasks differ from most word span tasks in the sense that the words that have to be remembered in reading span tasks are drawn from a large pool and not repeated in other trials of the test. Word span tasks often use a limited set of words that are repeated between lists of different length.

In order to see to which extent this difference between reading span and word span tasks caused the difference in predictive power, La Pointe & Engle (1990) conducted a series of experiments. In these experiments, word span tasks were administered with words drawn from a large pool that were not repeated between trials. They found that word span correlated as well with VSAT scores as reading span when repetition is avoided.

Recap. To summarize, simple span measures are generally taken to tap peripheral storage, whereas complex spans are taken to tap the central language processing skills. However, this picture seems somewhat overgeneralized given the fact that simple span tasks, such as word span, predict language skills well when drawn from a large pool.

Span Data from Experiments 1 to 5b

Given the fact that span data play a crucial role in the debate about WM and sentence comprehension, span data were collected in Experiments 1 to 5. The current chapter reports the correlations of the effects investigated with the span measures collected in these experiments and it ties in with the discussion on span measures described above.

Collection of Span Data

Two different span tests were administered, a nonword span test and a Dutch version of the Salthouse Listening Span Test. Both span tests will be described below.

Nonword Span

Procedure. The Nonword Span Test consisted of lists of nonwords that were read out loud by the experimenter. The participants' task was to repeat a list immediately after the experimenter had finished reading it.

The lists consisted of one, two, and three-syllable nonwords. All nonwords were *pseudowords,* words that meet the phonotactic restrictions of a particular language, in this case of Dutch. The lists were matched for word length in number of phonemes and number of syllables as far as possible. Lists varied in length from two up to seven pseudowords. There were five lists of each size (five lists of two pseudowords, five lists of three, etc.).

Span size was determined as the largest set size the participant could repeat correctly. As long as subjects repeated one out of the five sets of a particular length correctly, they were allowed to continue to the lists of the next length. For example, if a participant repeated one out of five lists containing three nonwords correctly, (s)he could go on to the lists with four pseudowords. If (s)he failed on all five lists of four nonwords, his/her nonword span would be 3.

Purpose. The participants were asked to repeat the nonwords. Participants cannot store these pseudowords by using semantic or syntactic cues, because nonwords have neither meaning nor syntactic properties. Given the fact that the lists do not offer the possibility of structural integration, it is impossible to store the nonwords in some syntactic format. The only option for the participants is to store it in a phonological format in some mechanism like Baddeley's (1986) phonological loop. The nonwords are initially kept in the phonological store. After that, they may be rehearsed via the articulatory rehearsal mechanism. Therefore, nonword span can be taken to tap the capacity of phonological store and possibly that of the articulatory rehearsal mechanism.

Salthouse Listening Span Test

Procedure. With the Salthouse Listening Span Test, participants listened to lists of sentences varying in length from one to seven sentences. Their task was to remember the last word of each sentence. Meanwhile, the participants had to answer multiple-choice questions about the content of the sentences. At the end of each list, the instruction was given to write down the last word of each sentence in that particular set. Sentence lists and instructions were presented on a tape using a Hitatchi TRK-5600E tape recorder. There were three sets of each list length.

Participants were allowed to proceed to the next list size if they had performed correctly on all sentence sets of a particular size. Performance was counted as correct if participants answered the comprehension questions correctly and recalled the right sentence-final words in the right order. The span task was stopped after the list size was completed in which the participant started to make errors on the comprehension questions or on the recall task.

Span size was determined as the maximum list size the participant had done correctly plus the partial score of the next list size. For each correct set on the next list size, participants received an additional .33 score (since there were three sets in each series). For example, if a subject failed to recall a word of a set of four sentences and performed correctly on the other sets of that size, his/her score would be 3.67. The 3 represents the last set size on which the participant had a perfect performance. Additionally, two out of three sets of four were recalled correctly; therefore the score is increased with the partial score for set size 4, which is .67.

Purpose. Salthouse (1994) has argued that central-executive functioning can be measured by tests that involve two components: storage of information and computation of relations between elements. These aspects are both present in the Salthouse Listening Span Test. Therefore, in terms of Baddeley's (1986) WM model, it may be taken to measure central-executive functioning. Which language comprehension processes carried out by the central executive does the Salthouse Listening Span Test measure? The answer to this question depends on one's model of language and WM.

Just & Carpenter (1992) proposed a single-resource model of sentence comprehension. In their model, sentences are processed by a single WM store similar to Baddeley's (1986) central executive. They assume that there are individual differences in capacity of that WM store. Furthermore, they argued that complex span measures, such as their Reading Span Test and the Salthouse Listening Span Test measure the capacity of this WM system. Therefore, high correlations of

Salthouse listening span with all sorts of language complexity effects are predicted by the single-resource theory.

However, Caplan & Waters (1999), who advocate a separate-resource theory of language processing, argue that span measures tap postinterpretive processing, but not interpretive processing. Both stages are taken to be located within the central-executive component of WM. Given the assumption that this separate-resource model is a specification of central-executive processing during language processing, the Salthouse listening span can be taken as a measure of either the syntactic stage or the postsyntactic (second) stage.

In addition to these central processes, the Salthouse Listening Span Test may also tap other kinds of storage. A first candidate mechanism would be the phonological loop. It may be that during the first or second stage, an articulatory back-up is addressed. This idea has been proposed by a number of researchers (e.g., Baddeley et al., 1987; Caplan & Waters, 1999; Waters et al., 1987, Withaar, 2001b). The Salthouse Listening Span Test may measure the capacity of the phonological loop. If so, correlations of the span measure with phonological-loop tasks, such as articulatory suppression are predicted.

The second system that may be measured by the Salthouse Listening Span Test, in addition to one or both of Berwick & Weinberg's (1984) processing stages, is a semantic storage system. Martin & Romani (1994) have proposed that during sentence processing, unintegrated words are stored in a separate, semantic WM system. The storage of unintegrated words is an important aspect of the Salthouse Listening Span Test. Therefore, it is not unlikely that this task is sensitive to semantic memory loads.

Span Correlations in Experiments 1 to 5b

Correlations of the main effects and interactions investigated in the Experiments 1 to 5b with memory spans were calculated. How this was done is described below.

Data Treatment

Calculation of Effects. In Experiments 1 to 5b, the sizes of potential main effects and interactions were computed. This was done even for those effects that did not reach significance in the ANOVAs, because it may be that some effects do not reach significance due to the fact that subjects deal with them in different ways.

Participants with few WM resources may deploy other strategies than those with ample WM capacity.

Main effects were calculated by averaging the mean RTs of all conditions of the *simple* level and subtracting this average from the average of the mean RTs of all conditions of the *complex* level. For example, Experiments 1 orthogonally varied Clause Type (subject- vs. object-relative clauses; SR vs. OR), Propositional Complexity (one- vs. two-proposition sentences; 1P vs. 2P), and Articulatory Suppression (no load vs. counting; NL vs. CO). This yielded eight experimental conditions: SR1PNL, SR1PCO, OR1PNL, OR1PCO, SR2PNL, SR2PCO, OR2PNL, and OR2PCO.[2] The main effect of Articulatory Suppression was calculated using the following formula:

(1) ((SR1PNL + OR1PNL + SR2PNL + OR2PNL) / 4) – (SR1PCO + OR1PCO + SR2PCO + OR2PCO) / 4))

The other main effects were calculated in a similar way.

After calculation of the main effect, correlations were computed between memory span and the size of the main effect. A significant correlation with a particular main effect means that individuals with low memory spans react differently to that effect than high-span individuals. A significant negative correlation in the RT analysis signals that low-span individuals take much more time reading in the no-load condition than under articulatory suppression, whereas high spanners show an opposite tendency. A significant positive correlation in the RT analysis shows that high-spanners need more time to read in the no-load than in the counting condition, but that low-span individuals do so to a lesser extent.

Two-way interactions were calculated by subtracting the average of the mean RTs to the *simple* conditions from the *complex* conditions within the *simple* level of a particular factor. This subtraction is then subtracted from the same subtraction at the *complex* level of that particular factor. In (2), the calculation for the two-way interaction of Propositional Complexity x Articulatory Suppression in Experiment 1 is given:

(2) (((SR2PNL + OR2PNL) / 2) – ((SR2PCO + OR2PCO) / 2)) – ((SR1PNL + OR1PNL) / 2) – ((SR1PCO + OR1PCO) / 2)))

---

[2] These abbreviations are compounds of the ones used above: SR = subject relative; OR = object relative; 1P = one proposition; 2P = two propositions; NL = no load; CO = counting. Thus, SR1PNL is a one-proposition subject relative in the no-load condition etc.

The other two-way interactions were computed in the same manner; the three-way interactions in the experiments were computed like the two-way interactions, but then with yet another level of subtraction.

It is difficult to tell what significant correlations with two-way interactions mean just by looking at the correlation data alone. Therefore, it is important to see how subjects at the different ends of the span scale react to the different experimental conditions.

Criteria. Correlations of nonword span and Salthouse listening span with all potential main effects and interactions of the designs of the Experiments 1 to 5b were computed. If a particular span measure correlated significantly with an interaction, that correlation is reported when the span measure also correlated (almost) significantly with a simple effect in that interaction (with $p < .1$).

Significant correlations with interactions that did not meet these criteria were considered uninterpretible. These correlations will not be reported, unless the same interaction produced a significant correlation with a particular span measure in one of the other experiments reported in the Chapters 3, 4, and 5. In that case, the correlation will be reported but not explained.

For the interpretation of significant correlations with main or simple effects, a division in two or three span groups was made. Division in span groups was done to report how individuals on the different sides of the span distribution reacted to a certain complexity effect. This was done for explorative purposes only rather than as a basis for a statistical test with levels of the factor Span.

Correlations with Nonword and Salthouse Listening Span

Correlations Between Nonword and Salthouse Listening Span. For all experiments, correlations between the two span measures collected were computed.[3] Experiments 1 and 5a showed a significant correlation between the two span measures ($r = .45$, $p < .05$ and $r = .35$, $p < .05$ respectively). In the other experiments, no significant correlations were observed (Experiment2: $r = .25$, $p > .1$; Experiments 3,4, and 5b: $r = .04$, $p > .1$).[4]

---

[3] Pearson's correlations were used here; these correlations can be used only if the span data are normally distributed. The span data in this study are, as can be seen in the Appendices on p. 185.
[4] Experiments 3,4, and 5b used the same subjects; the other experiments all used different subjects.

Correlations with Complexity Effects. Table 1 gives an overview of significant and interpretable correlations of nonword span with complexity effects from Experiments 1 to 5b. All significant correlations mentioned in Table 1 were further analyzed in order to see how participants of different span sizes reacted to the variable that yielded the significant correlation with memory span. These data can be found in the appendix on span measures on p. 185. Interactions are only reported in Table 1 if they yielded a significant and interpretable correlation with span. Uninterpretible correlations with interactions are only reported when a correlation with that particular interaction yielded a significant and interpretable interaction in one of the other experiments as well.

Table 1
Correlations with Nonword Span in Experiments 1 to 5b.

	Effects						
Experiment	CT	AL	PC	ET	AS	CT x PC	ET x AS
				RT Analysis			
1	.12	-	-.04	-	-.38*	.00	-
2	-.15	-	-	.10	-.02	-	.01
3	-.25	-	.18	-	.01	-.17	-
4	-.12	-	-	-.24	-.12	-	-.18
5a	-	.15	-	.30	.02	-	-.03
5b	-	-.14	-	.15	.16	-	.20
				Error Analysis			
1	.10	-	.08	-	-.00	.10	-
2	-.15	-	-	-.07	.10	-	.03
3	-.31	-	.33	-	-.07	-.38*	-
4	-.16	-	-	-.19	-.17	-	-.15
5a	-	-.09	-	.30	.18	-	-.37*
5b	-	-.26	-	-.16	.12	-	.19

Note. The values represent *r* values (Pearson's correlation coefficients). Dashes indicate that a factor or interaction was not tested in a particular experiment. All uninterpretible two- and three-way interactions were omitted. CT = Clause Type; AL = Adjectival Load; PC = Propositional Complexity; ET = Embedding Type; AS = Articulatory Suppression.
* $p < .05$.

Table 2 gives an overview of all correlations of Salthouse listening span with the factors manipulated in Experiments 1 to 5b. The appendix on span measures on p. 185 gives an interpretation of these correlations.

Table 2
Significant Correlations with Salthouse Listening Span in Experiments 1 to 5b.

	Effects							
Experiment	CT	AL	PC	ET	AS	ET x AS	AL x ET	CT x ET x AS
					RT Analysis			
1	.01	-	.15	-	-.00	-	-	-
2	-.27	-	-	.18	.41*	-.19	-	.14
3	-.06	-	-.20	-	.07	-	-	-
4	-.15	-	-	-.08	.19	.12	-	-.45*
5a	-	-.31	-	.23	-.20	-.27	.12	-
5b	-	-.20	-	-.05	.18	.13	.35*	-
					Error Analysis			
1	.27	-	.29	-	.01	-	-	-
2	.14	-	-	.23	.02	.13	-	.01
3	.07	-	.08	-	.08	-	-	-
4	-.14	-	-	.04	.09	-.14	-	-.21
5a	-	.01	-	-.14	.29	-.40*	.35*	-
5b	-	-.38*	-	-.22	.37*	.09	.14	-

Note. The values represent *r* values (Pearson's correlation coefficients). Dashes indicate that a factor or interaction was not tested in a particular experiment. All uninterpretible two- and three-way interactions were omitted, except in those cases where a significant correlation with that interaction was found in a different experiment as well. CT = Clause Type; AL = Adjectival Load; PC = Propositional Complexity; ET = Embedding Type; AS = Articulatory Suppression.
* $p < .05$.

Discussion

In this section, the consequences of the different correlations described above will be discussed. First, the interpretation of each correlation will be covered; then, the implications of these correlations for the different WM models will be outlined. Finally, the importance of span data for the literature on WM and sentence processing will be dealt with.

Explaining the Correlations

Nonword Span. A number of significant correlations with nonword span were found. Table 1 gives an overview of these correlations.

In Table 1, one can see a significant correlation of nonword span and Articulatory Suppression in Experiment 1. This looks like an interesting result: nonword span, which is taken to measure phonological-loop capacity produces a significant correlation with a task that taps the phonological loop, articulatory suppression.

On second thought, this correlation is less promising. Figure 6 in the Appendices on p. 187 showed the effects of articulatory suppression for each subject, divided by span group. There, we can see that there is quite some difference in how participants deal with articulatory suppression; especially those in the medium group show much variation. Therefore, it is hard to tell what this correlation exactly means.

A second problem with this correlation is that it was never replicated in the other experiments, which also had the articulatory-suppression manipulation. Taking these two objections into account, I conclude that this correlation most likely reflects noise in the data and can be considered a false positive.

The second significant correlation was that of nonword span with the two-way interaction of Clause Type x Propositional Complexity. Low spanners made more errors in subject- than in object-relative clauses in the two-proposition sentences only; high spanners showed no difference between these two conditions. The interaction of Clause Type x Propositional Complexity was significant in the ANOVAs as well, and hard to interpret.

This interaction shows that not all subjects deal with this particular condition in the same way: only low-span participants seem to cause this effect. It is probably due to the fact that in subject-relative clauses, subject-verb agreement must be checked over a longer distance than object relatives. Low spanners, who have less phonological storage capacity, have trouble retrieving these features from a phonological back-up representation kept in the phonological loop; high spanner do not seem to be affected in that way.

There is a caveat with the interpretation of this particular correlation: it was not found in Experiment 1, although this experiment also had the factors Clause Type and Propositional Complexity. Therefore, one cannot really tell whether this particular correlation is due to noise in the data or not.

Finally, there was a significant correlation with the two-way interaction of Embedding Type x Articulatory Suppression in Experiment 5a. This correlation is hard to interpret, and most likely a false positive, given the fact that it was not replicated in any other experiment in this book that looked at this particular interaction.

Salthouse Listening Span. In Experiments 1 to 5b, several significant correlations with Salthouse listening span were found. Table 2 gives an overview of these correlations. The Salthouse Listening Span Test produced two significant correlations with Articulatory Suppression, in Experiments 2 and 5b. Given the fact that these correlations did not occur in the other experiments and that they go in entirely different directions in the two experiments, these interactions are likely to reflect noise in the data.

For the other correlations with the Salthouse Listening Span Test, the same thing goes: they are never found consistently. Therefore, it is hard to tell whether these correlations actually tell us something about how individuals with different memory spans deal with certain language complexities or whether the correlations themselves are just noisy.

Consequences for the Debate on Span Measures

These span data have a number of implications for the different models of WM capacity and language processing. First, the validity of the span measures used will be discussed; then the reliability will be covered.

The Validity Issue. In the introduction to the current chapter, several proposals about the systems measured by span tasks were outlined. It was proposed that the nonword span reflected phonological-loop activity, and that the Salthouse listening span reflected central-executive capacity and possibly also phonological or lexicosemantic memory.

Experiments 1 and 5a yielded significant correlations of nonword span with Salthouse listening span; the other experiments did not. This suggests that these two span measures tap different capacities. Had they measured the same capacity, systematic correlations should have been obtained across experiments. Most probably, these two tasks measure different capacities (if any), and the subjects in Experiments 1 and 5a seem to have a nonword span capacity that is about comparable with their Salthouse span capacity. However, this need not always be the case as witnessed by the lack of significant correlations between the span measures in the other experiments.

The Reliability Issue. Consistent correlations with language complexity effects were not obtained in these experiments. The span measures occasionally yielded significant correlations, but crucially, these correlations were never found in other experiments that manipulated the same factors. If the span tasks were sensitive to a particular manipulation, one would expect systematic differences across experiments. On the basis of the current data, we must conclude that the two span tasks used are of limited reliability.

The fact that few correlations with nonword span were obtained is in line with the idea proposed by Daneman & Carpenter (1980) that simple span measures are poor predictors of language processing capacity. They claim that complex span measures tap central-executive processing whereas simple span measures test something more peripheral.

If this is the case, one would expect there to be no significant correlations of simple and complex span measures. However, their suggestion that complex span measures are good predictors of language capacity is challenged by the lack of consistent correlations with Salthouse listening span. The dichotomy between simple and complex span tasks is not as sharp as Daneman & Carpenter (1980) argue it to be.

However, Caplan & Waters's (1999) proposal that both simple and complex span measures test propositional processing capacity is not supported by the current findings either: not a single correlation with propositional complexity was obtained in Experiments 1 and 3. Therefore, the claim that span measures predict postsyntactic, propositional processing is also not supported by the current data.

Some Remarks on the Use of Span Measures

In the experiments just reported, systematic correlations with simple or complex span measures were not found. However, a number of psycholinguistic experiments have been conducted that did find significant span effects (e.g., Caplan & Waters, 1999; King & Just, 1991; Vos, 1999). The fact that these studies did find effects and no effects were found in the current study may have to do with treatment of span data.

In all three studies, subjects with the lowest spans were compared to participants with highest spans. In other words, only participants at the extreme end of the span distribution were compared in these studies. If span effects can be found when one looks at the extreme ends of the distribution, why was this not done in the current study? The most important reason is that ignoring the medium group – which makes up almost 50 to 60% of the population – makes the result less generalizable. Therefore, I believe that correlation analyses, which make use of all participants, are more informative.

A second reason for being reluctant to trust studies that use WM span measures as predictors of language processing capacity (or any other capacity for that matter) is the limited reliability of span measures. Brooks & Watkins (1990, p. 1134) have shown that "estimates of a person's span fluctuate across successive tests." Even if one has explicit views on what a particular span tests, it is far from clear whether it does so systematically. Usually, span tasks of participants of sentence processing experiments are administered only once, which makes the use of span measures as predictors of language processing capacity a cautious enterprise.

Chapter 7
A Model of WM and Sentence Comprehension

Introduction

In the current chapter, a model of working memory (WM) and sentence comprehension will be introduced. The model specifies which WM components are active during sentence comprehension. The focus of the model will be on the role of the phonological loop in reading. Therefore, the processes that call on the phonological loop during comprehension receive most attention.

The current chapter is organized as follows. First, the scope of the model will be outlined. Then, a brief sketch of the model will be given. After that, the components of the model will be described. This section is followed by a listing of the predictions derived from the model. Then, the model will be related to the data from Experiments 1 to 5b. After that, the model will be discussed in relation to language processing in children suffering from specific language impairment (SLI) and in patients suffering from aphasia. Finally, some issues that need further investigation and future research will be addressed.

Chapters 3, 4, and 5 were written in such a way that they could be read independently. The current chapter is impossible to read without having read the other chapters because it refers back to effects reported in these other chapters.

Scope of the Model

The data from Experiments 1 to 5b do not fully support existing models of WM and sentence comprehension (e.g., Baddeley, Vallar, & Wilson, 1987; Caplan & Waters, 1999; Just & Carpenter, 1992; Martin & Romani, 1994). These data suggest a separation of WM resources; therefore, they are not compatible with the single-resource theory proposed by Just & Carpenter (1992) and adopted by Gibson (1998).

However, the separate-resource model proposed by Caplan & Waters (1999) is not fully supported either. The latter model assumes two processing stages, one for syntactic and one for sentence-level semantic processing, and a role for a phonological back-up in the latter, but not in the first system. Interactions of Articulatory Suppression with processes that occur during the syntactic stage suggest that Caplan & Waters's specification of the role for a phonological back-up in sentence comprehension is not correct.

The proposals made by Baddeley et al. (1987), Just & Carpenter (1992), and Martin & Romani (1994) may need revision since they do not assume separate resources within the central-executive component of WM for syntactic and propositional processing.

Instead, a different model will be presented, which shares the assumption of two processing stages with Caplan & Waters's (1999) proposal, but has a different specification of the role of a phonological back-up during sentence comprehension. The purpose of the model is to specify which WM systems are involved in sentence comprehension and under which circumstances these systems are deployed.

The model is based on experiments that looked at sentences presented in isolation. Therefore, processes that arise at the sentence or clause level receive most attention. Processes in earlier (presyntactic) and later (propositional, pragmatic, etc.) stages are less emphasized, which is not to say that WM involvement is restricted to the syntactic stage, but that WM processes in other stages fall beyond this study's scope.

The Model in a Nutshell

The model that I propose assumes two processing stages, like Berwick & Weinberg's (1984) and Caplan & Waters's (1999) models. The first stage entails syntactic processes; the second stage is responsible for sentence-level semantic processing such as the interpretation of pronouns and integration of different propositions.

These two stages are informationally encapsulated: information available to the syntactic processor is not available to the propositional processor and vice versa. The modules operate in a strictly serial way: information is first processed by the syntactic module and then passed on to the propositional processor (Berwick & Weinberg, 1984). In that sense, it proceeds from lower- to higher-level representations.

During subject-verb integration, a process carried out during the syntactic stage, a phonological back-up representation may be addressed. The phonological back-up representation is not always addressed during this process, but only in those cases where subject-verb agreement must be checked over a long distance. The checking of subject-verb agreement is vulnerable to manipulations that degrade the phonological back-up representation, such as articulatory suppression, because it requires that phonologically nonsalient morphosyntactic elements be compared over a long distance.

The Model's Components and Their Relations

The model consists of three modules, a syntactic processor, a phonological back-up representation, and a propositional store. Each of these systems will be described below. The syntactic processor and the propositional store are considered *central* modules, whereas the phonological back-up representation is taken to be *peripheral.*

The distinction between central and peripheral modules stems from Baddeley's (1986) model of WM. Baddeley assumes that WM consists of a *central executive,* which is involved in the storage and processing of information, and of peripheral *slave systems,* which may be deployed by the central executive if it needs additional storage capacity. The phonological back-up representation is taken to be stored in a system similar to one of Baddeley's slave systems, the phonological loop. For a more detailed description of Baddeley's model, the reader is referred to page 20.

The next parts deal with the function of the three modules distinguished by the model of WM and sentence comprehension outlined in this book.

The Syntactic Processor

The syntactic processor's tasks are twofold: first, it combines words into phrases, and second, it stores these phrases until they can be shunted to the propositional processor (Berwick & Weinberg, 1984). Because these tasks involve both storage and processing of information, the syntactic processor is considered a central component (i.e., a component like Baddeley's, 1986 central executive).

The syntactic processor is also involved in establishing referents for noun phrases (NPs) that must be bound in their local domain, such as reflexives (Berwick & Weinberg, 1984). In (1a), the reflexive *himself* refers to an antecedent within the same domain (i.e., the same clause), *Bill.* The fact that the two NPs refer to the same entity is indicated with a subscript *i:*

(1) a. John said that Bill_i was shaving himself_i
b. John_i said that Bill was shaving him_i

In (1b) however, the pronoun *himself* cannot refer to an NP within the same clause; instead it refers to the main-clause subject *John.* Finding referents for pronouns does not occur within the syntactic processor, but is done by the propositional processor (Berwick & Weinberg, 1984; Caplan & Waters, 1999), which will be discussed after the subsection on the phonological back-up representation.

The Phonological Back-Up Representation

In some cases, the syntactic processor addresses a phonological back-up representation of the sentence it is analyzing (Baddeley et al., 1987). This phonological back-up representation is kept active in a module similar to Baddeley's (1986) phonological store.

The syntactic processor addresses this back-up in order to check phonologically subtle cues that may influence parsing decisions at the syntactic level. In Dutch relative clauses, number information of the NPs and the verb is important for syntactic disambiguation (e.g., Kaan, 1997; Lamers, 2001; Mak, 2001; Withaar & Stowe, 1999a). In this disambiguation process, the checking of subject-verb agreement plays an important role.

Dependence of the syntactic processor on the phonological back-up becomes apparent especially when agreement information of words that are far apart must be checked. The fact that a larger part of the back-up must be consulted in case of long-distance integration makes the checking procedures for these integrations more demanding.

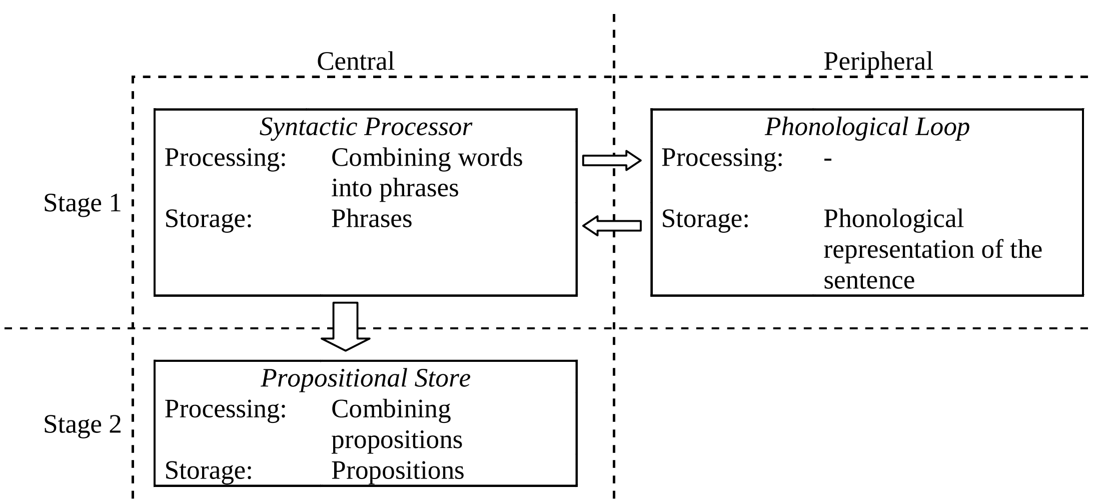


Figure 1. The model.

The Propositional Store

Once the syntactic processor has produced a complete clause, this clause is transferred to a second central module, the propositional store. Its task is to combine propositions and infer the relations between them. Furthermore, it is involved in processes such as establishing antecedents for pronouns (cf. 1b above). An overview of the contributions of the different modules in the sentence comprehension process is given in Figure 1 above.

Factors Contributing to Complexity Effects

The model just outlined makes a number of predictions. These will be discussed in the next section on page 146. However, before these predictions can be discussed, one must keep in mind that a number of factors contribute to the effects found in Experiments 1 to 5b. These factors will be considered before the predictions will be discussed. The following factors affect the data in the Experiments 1 to 5b.

First, the experiments were whole-sentence anomaly-judgment tasks. In these tasks, both reaction time (RT) and error data are collected. The model specifies which conditions are most difficult to process. However, one must take into account that complexity effects may be apparent in either the RT or the error data.

Second, the complexity of an experimental sentence is determined by the sentence's syntactic structure. Syntactic structure contributes to the complexity of the sentence because in complex sentences, a larger amount of syntactic structure must be built or items must be syntactically integrated over a longer distance.

The third factor adding to the complexity of a sentence is the distance over which subject-verb agreement must be checked in a particular clause. This factor is due to the size of the phonological segment that must be consulted. Especially in Dutch sentences, where subject-verb agreement must sometimes be checked over a long distance, this factor is an important contributor to sentence complexity. An example of a difference in integration distance found in most languages is the difference between right- and center-embedded sentences, as demonstrated in (2a and b) respectively:

(2) a. De studenten prijzen de fotograaf die de uitgevers accepteert
The students praise the photographer that the publishers accepts
'The students praise the photographer who accepts the publishers'

b. De fotograaf die de uitgevers accepteert, prijst de studenten
The photographer that the publishers accepts, praises the students
'The photographer who accepts the publishers praises the students'

In (2a), the matrix subject *studenten (students)* and verb *prijzen (praise)* are adjacent. Therefore, local feature checking suffices, which implies that at most a small part of the phonological back-up representation must be checked. In (2b), however, the feature check of the agreement features of the matrix subject and verb (in this case *fotograaf [photographer]* and *prijst [praises]*) must be accomplished over a large distance. This means that a much larger stretch of the back-up representation must be addressed.

The fourth factor contributing to the complexity of the judgment is the nature of the judgment process itself. Reversible sentences, such as (3a), repeated here as (3a), had lexical distractors, as shown in (3b):[1]

(3) a. De studenten prijzen de fotograaf die de uitgevers accepteert
The students praise the photographer that the publishers accepts
'The students praise the photographer who accepts the publishers'

b. # De studenten verslinden de fotograaf die de uitgevers accepteert
The students devour the photographer that the publishers accepts
'The students devour the photographer who accepts the publishers'

However, nonreversible sentences had θ-role inversions as distractors, as can be seen in (4a and b):

(4) a. De columnist veroordeelde de commando's die de kaping beëindigden
The columnist condemned the commandos that the hijack ended
'The columnist condemned the commandos that ended the hijack'

b. # De columnist veroordeelde de commando's die de kaping beëindigde
The columnist condemned the commandos that the hijack ended
'The columnist condemned the hijack that ended the commandos'

This cue is clearer as a basis for judgment, as witnessed by higher error rates in Experiments 1 and 2, which used reversible sentences, relative to Experiments 3 and 4, which did not. Therefore, the nonreversible sentences are expected to be easier to judge than the reversible sentences, although the decision is dependent on syntactic analysis.

---

[1] The pound sign (#) is used in this book for sentences that are syntactically correct, but semantically unacceptable.

Predictions of the Model

In the current section, the predictions of the model with respect to the different complexity manipulations used in the Experiments 1 to 5b will be discussed. In the following section, the actual effects will be related to these predictions. The model assumes two stages, one in which syntactic analysis is accomplished and one for propositional analysis. During the first stage, the syntactic processor addresses a phonological back-up representation. The recruitment of this phonological back-up may yield complex predictions.

Phonological-Loop Load

Articulatory Suppression. The first effect to be discussed in terms of predictions is the effect of articulatory suppression. Articulatory suppression hampers the accessibility of the phonological loop (e.g., Baddeley, 1986; Gathercole & Baddeley, 1993). Several researchers have proposed that during sentence comprehension, the phonological loop stores a phonological back-up representation of the sentence that is being read (e.g., Baddeley, 1986; Baddeley et al., 1987; Waters, Caplan, & Hildebrandt, 1987).

Therefore, I assume that articulatory suppression hampers recruitment of a phonological back-up representation during syntactic processing (e.g., Withaar 2001a; Withaar & Stowe, 1999a). The consequences of this detrimental influence may differ from language to language.

In Dutch, a language with verb-final relative clauses, participants must wait until the end of a clause before they can integrate the subject and verb of that clause. As already pointed out, I propose that agreement features of a clause's subject and verb are checked during this integration process. If the integration occurs over a long distance, it will be more difficult, particularly when the phonological back-up is degraded or diminished.

This suggests that if readers of Dutch want the phonological back-up to be available by the time they read the end of a clause, they must read faster under articulatory suppression. Reading faster enables them to still access the back-up in the phonological loop when they arrive at the end of the sentence. This preference for speed is likely to go at the expense of accuracy. The articulatory-suppression condition is predicted to yield faster RTs, but higher errors than the no-load condition in Dutch.

In English, where the checking of agreement features does not play an important role in determining the structure of a relative clause and usually occurs over a short distance (e.g., Baddeley et al., 1987; MacDonald & Christiansen, in press; Waters et al., 1987), other influences of articulatory suppression may be expected. There, the subject and the verb of a clause are always adjacent. Subject-verb agreement is hardly checked over a long distance in English. Therefore, readers of English have no urge to come to the end of the clause quickly when under articulatory suppression. Participants are often slowed down in a dual-task compared to a no-load condition (e.g., Baddeley, 1986; Salthouse, 1990; 1994). Given this fact, readers of English are expected to be slower under articulatory suppression than in a no-load condition.

Manipulations Affected by Phonological-Loop Load

Articulatory suppression influences some processes in the syntactic stage, but has no influence on processes in the propositional stage (Withaar & Stowe, 1999b). Therefore, the model predicts interactions of Articulatory Suppression with complexity manipulations requiring the recruitment of a phonological back-up in the syntactic stage, but no interactions with propositional manipulations, since the information represented in the phonological back-up does not seem relevant to its operations..

The model predicts interactions with phonological-loop manipulations in cases where subject-verb agreement must be checked over a long-distance. This is the case in two manipulations tested in Experiments 1 to 5b, Clause Type and Embedding Type.

Clause Type. Subject-relative clauses are structurally less complex than object-relative clauses. Therefore, one may expect higher RTs and error rates for object- than for subject-relative clauses.

However, in Dutch relative clauses, the subject and verb of the embedded clause are further apart in subject- than in object-relative clauses, as can be seen in (5a and b) below. In (5a), *politici (politicians)* and *afwezen (rejected)* are not adjacent, whereas these words are in (5b):

(5) a. Dit zijn de politici die het voorstel afwezen
These are the politicians that the proposal rejected
'These are the politicians who rejected the proposal'

b. Dit is het voorstel dat de politici afwezen
This is the proposal that the politicians rejected
'This is the proposal who the politicians rejected'

Therefore, the checking of subject-verb agreement is more demanding for subject- than for object-relative clauses. This may lead to increases in processing time and decreases in accuracy for subject-relative clauses.

In summary, on the one hand, subject-relative clauses are likely to produce shorter response latencies and/or lower error rates because they are structurally less complex. However, checking of subject-verb agreement occurs over a longer distance in these constructions; therefore, subject-relative clauses should yield longer RTs and/or higher error rates. It may depend on task strategies whether the feature checking aspect or the structure building aspect receives more prominence. I will return to this issue in the next section on page 152.

Embedding Type. Right-embedded sentences as in (6a) are easier than center-embedded sentences as in (6b) because the subject and verb of the matrix clause must be integrated over a larger distance in the latter sentence type (e.g., Frazier & Rayner, 1988; Gunter, 1995; Miller & Isard, 1964):

(6) a. De columnist veroordeelde de commando's die de kaping beëindigden
The columnist condemned the commandos that the hijack ended
'The columnist condemned the commandos that ended the hijack'

b. De commando's die de kaping beëindigden, schokten de columnist
The commandos that the hijack ended, shocked the columnist
'The commandos that ended the hijack shocked the columnist'

Given the fact that the two sentences are structurally equally complex – both sentences consist of a verb, two NPs, and an embedded clause – no complexity effects are expected to arise within the syntactic processor.[2]

Whether effects of embedding arise at all is dependent on whether subject-verb agreement features must be checked over a long distance or not. Complexity effects are expected to arise in sentence (6b), where the verb *schokten (shocked)* and the subject *commando's (commandos)* are far apart. This is not the case in (7), where subject-verb agreement is between the subject *commando's (commandos)* and the adjacent auxiliary *hebben (have)*, since the center-embedded clause does not intervene between subject and verb:

---

[2] This assumption is in line with Gibson (1998), who assumes that integration cost, and not memory load of the matrix subject is the major determinant of the complexity of the factor Embedding. However, proposals made by Marcus (1980) and Berwick & Weinberg (1984) do predict additional storage demands.

(7) De commando's hebben, toen ze de kaping beëindigden, de columnist geschokt
The commandos have, when they ended the hijack, the columnist shocked
'The commandos shocked the columnist when they ended the hijack'

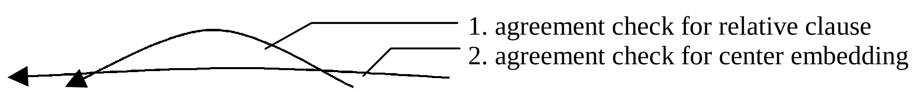
In other words: center-embedded sentences are predicted to yield longer response latencies in conditions where (a) subject-verb feature checking must occur over a long distance and (b) the phonological back-up is degraded. This implies that embedding should produce neither main effects nor in sentences like (7).

With respect to the structural aspects of the subject- vs. object-relative comparison, no interactions of Clause Type x Embedding are expected to arise within the syntactic processor because right- and center-embedded structures are equally complex. In other words, given the fact that the two levels of the factor Embedding are equally complex from a structural point of view, there ought not be a superadditive two-way interaction of Clause Type x Embedding within the syntactic processor.

It is true that the processing of center-embedded structures requires the storage of a main-clause NP that cannot be immediately integrated into the main clause. This is because the verb with which the matrix subject must be integrated comes late in the sentence. Therefore, storage demands are taken to play at least some role. However, the model assumes that the demands of long-distance integration of these two elements outweigh the temporary storage demands by far (cf. Gibson, 1998).

With respect to the recruitment of the phonological loop in sentence comprehension, no superadditive three-way interaction of Clause Type x Embedding x Articulatory Suppression is expected either. This is because the long-distance feature checking of the matrix subject and verb (in center-embeddings) and of the embedded subject and verb (in subject-relative clauses) occurs at different points in the sentence.

The feature check between the subject and the verb of the embedded clause is carried out earlier than the check of main-clause subject-verb agreement features. Integration of embedded subject and verb is taken to occur first because the embedded verb precedes the main verb. This is shown in (8):



(8) De *commando's* die de kaping *beëindigden, schokten* de columnist
The commandos that the hijack ended, shocked the columnist
'The commandos that ended the hijack shocked the columnist'

Manipulations Insensitive to Phonological-Loop Load

Adjectival Load. Sentences with stacked adjectives as in (9a) are predicted to be more demanding than sentences that have no stacking of adjectives as in (9b) because more words must be processed at both the syntactic and the semantic level:

(9) a. De barkeeper heeft de chips opgegeten terwijl de verregende klant niets doorhad
The barkeeper has the chips eaten while the drenched customer nothing noticed
'The barkeeper ate the chips while the drenched customer did not notice anything'

b. De barkeeper heeft de chips opgegeten terwijl de verregende hongerige trouwe klant niets doorhad
The barkeeper has the chips eaten while the drenched hungry loyal customer nothing noticed
'The barkeeper ate the chips while the drenched hungry loyal customer did not notice anything'

The fact that more words must be processed has implications in two respects. First, it takes more time to read the sentences with stacked adjectives, because two more words must be identified and understood. Second, these words must be integrated into a syntactic structure, a process that is likely to be time consuming as well.

Adjectival Load is predicted not to interact with Embedding Type. In center-embedded clauses, the checking of subject-verb agreement plays a crucial role in terms of syntactic disambiguation. This is not the case with the processing of stacked adjectives; therefore, no interactions of adjectival load and embedding are expected.

Propositional Complexity. Two-proposition sentences have more events than one-proposition sentences. The difference between the two conditions is shown in (10). In (10a), there is only one action: politicians reject a proposal; in (10b) there are two actions: experts judge a jeweler and the jeweler repairs watches:

(10) a. Dit zijn de politici die het voorstel afwezen
These are the politicians that the proposal rejected$_{pl.}$
'These are the politicians who rejected the proposal'

b. De experts beoordeelden de juwelier die de horloges repareerde
The experts judged the jeweler that the watches repaired$_{sg.}$
'The experts judged the jeweler who repaired the watches'

The two-proposition sentences are predicted to be more complex than the one-proposition sentences because more propositions must be stored and combined in the propositional processor. Thus, higher error rates and slower RTs are predicted for the more complex two-action sentences.

Given the fact that the model distinguishes between a syntactic and a propositional stage, no interactions are expected with complexity effects arising at the syntactic level. The only syntactic factor tested in an orthogonal design with the Propositional Complexity manipulation was Clause Type (subject- vs. object-relative clauses). For these experiments (Experiments 1 and 3), the model predicts independent effects of Propositional Complexity and Clause Type. This was the result in Experiment 3; Experiment 1 showed an non-additive interaction which had no clear interpretation.

Furthermore, the model assumes that the phonological loop is involved in syntactic, but not propositional processing. Thus, interactions of Propositional Complexity x Articulatory Suppression should not occur according to the model.

Reversibility

Although the reversibility of materials did not occur as a factor in Experiments 1 to 5b, it is useful to consider the influence of reversibility on the processing and judgment of sentences. Reversibility affects the complexity of the materials in two ways.

First, reversible sentences are harder to process because readers must rely solely on phonologically nonsalient information (number information) for structural disambiguation. Nonreversible sentences have the same information, but in addition, subjects may benefit from an important additional cue, lexical semantics. As can be seen in (11a), readers of nonreversible sentences can determine which sentence part is the subject and which is the object on the basis of world knowledge: inanimate NPs are usually worse subjects than animate NPs. Sentence (11b) does not provide this information:

(11) a. It is the man who eats the apples
b. It is the man who sees the women

Second, the use of reversible or nonreversible sentences affects the complexity of the judgment task considerably. Reversible and nonreversible had different distractors. The nonreversible sentences had θ-role inversions as distractors; the reversible sentences had lexical distractors. The latter type of distractor makes the judgment of the sentences more difficult.[3] On the basis of both phenomena, it is predicted that reversible sentences yield longer RTs than nonreversible sentences.

Therefore, the processing of relative clauses should be different in reversible and nonreversible experiments due to the attention participants must pay to different cues for processing and judgment.

Back to the Experimental Data

Now that the predictions made by the model have been outlined, it is important to go back to the data in Experiments 1 to 5b to see whether these predictions are all borne out or not. The current section will cover the obtained main effects and interactions.

At first sight, it may seem a bit odd to conduct experiments, develop a model based on these experiments, and then go back to the very same experiments in order to see whether the model makes the right predictions. The model was derived from the data and not vice versa.

However, in the Chapters 3 to 5, details of the model were specified, whereas in the current chapter, the model is presented as a whole for the first time. The fact that the model is presented fully makes it interesting to compare experiments described in the different chapters directly.

Articulatory Suppression and the Phonological Back-Up

The model predicts faster RTs and decreased accuracy for the articulatory-suppression conditions if the inflected verb appears at the end of a clause. This is the case in Dutch and German relative clauses. Faster RTs for the articulatory-suppression conditions were consistently found in Experiments 1 to 5b. In Experiment 4, these faster RTs are accompanied by higher error rates for the articulatory-suppression conditions. In the other experiments, higher error proportions were not obtained. This appeared to be attributable to a response bias in the articulatory-suppression conditions of these experiments. Under articulatory suppression, subjects performed worse on the incorrect sentences, which were all

---

[3] Note that Caplan and coworkers assume that with reversible combinations, syntactic processing may not be carried out at all (Caplan, personal communication, April, 1999), thus predicting that reversible judgments should be easier. Effects of Clause Type in Experiments 1 and 2 in this book suggest otherwise.

filler sentences due to a bias for *yes* responses. This bias would lead to an apparently correct response to the target sentences, but not to the fillers that were implausible.

In languages that always have the verb in second position like English, speeding up does not have a function, whereas a slow-down is not unexpected for the articulatory-suppression conditions because dual tasks are generally more demanding than single tasks. This is exactly the pattern that was obtained by Waters et al. (1987, exp. 3), who conducted an experiment that had the same design as Experiment 3, using English materials.

The Phonological Loop and Clause Type

The Main Effect of Clause Type. The model predicts that subject-relative clauses yield faster RTs and lower error rates than object-relative clauses from a structural point of view. However, in terms of checking of subject-verb agreement features, the opposite prediction is made: subject-verb agreement must be checked over a longer distance for subject- than for object-relative clauses.

Let us consider the effects found in Experiments 1 to 4 (which were the experiments that manipulated Clause Type). Faster RTs for subject-relative clauses were obtained in Experiments 1 to 3. This suggests that in most cases, building the clause structure is more demanding than the checking of subject-verb agreement features.

This is not to say that feature checking plays no role at all: in Experiment 2, faster RTs for subject relatives were accompanied by higher error proportions for these constructions (a speed-accuracy trade-off). This seems to be due to the fact that participants had more difficulty retrieving the subject-verb agreement information in the subject relatives than in the object relatives. This suggestion receives even more substance when one takes into account that this effect in the error rates was part of a super-additive two-way interaction of Clause Type x Articulatory Suppression.

There is a second trade-off with respect to the manipulation of Clause Type: in Experiment 4, the subject-relative clauses were judged more slowly and less accurately. The difference between Experiments 2 and 4 is the reversibility of the materials. In Experiment 2, reversible materials and lexicosemantic distractors were used and in Experiment 4 the materials were nonreversible and had θ-role inversions as distractors.

The fact that the trade-offs went in different directions in Experiments 2 and 4 suggests that the task demands exert considerable influence on the syntactic processor and its use of the phonological back-up. In Experiment 2, processing of agreement features was important for interpretation of the sentence, but not for the judgment since judgment relied on lexical semantics, and not on agreement information. Participants had no need to process the agreement features consciously, which lead to decreased accuracy in feature checking and worse performance for subject-relative clauses.

However, the opposite seems to be the case in Experiment 4. There, subjects had to process the agreement information thoroughly because this was explicitly needed for the judgment. Therefore, subjects paid more attention to the agreement features of subject-relative clauses, which lead to a slow-down.

An Interaction of Clause Type x Articulatory Suppression. The trade-off in Experiment 2, with faster RTs, but decreased accuracy for subject- compared to object-relative clauses, was accompanied by a two-way interaction of Clause Type x Articulatory Suppression. Subject-relative clauses were processed less accurately under articulatory suppression only. In subject-relative clauses, subject-verb agreement must be checked over a longer distance than in object-relative clauses. The increased error rates for subject-relatives suggest that this process is hampered by articulatory suppression. They furthermore suggest that the participants seem unaware of this detrimental influence on their performance, probably because they do not consciously need the agreement features in order to make the right judgment. (Remember: the distractor sentences had lexical violations rather than θ-role inversions.)

The Phonological Loop and Embedding Type

No Main Effect of Embedding Type. It is proposed that center-embedded structures are structurally as complex as right-embedded ones. The fact that no main effects of Embedding Type were found in any of the experiments supports this assumption. Center-embedded structures are more demanding than right-embedded sentences only in those cases where subject-verb agreement features must be checked over a long distance, and these effects are only expected to show up in articulatory-suppression conditions. More attention will be devoted to this phenomenon in the discussion below about interactions.

Interactions of Embedding Type x Articulatory Suppression. Effects of Embedding Type are not present in no-load conditions, but only arise under articulatory suppression. This superadditive interaction of Embedding Type x

Articulatory Suppression suggests that the checking of agreement features is hindered by articulatory suppression, but that it can be accomplished without any trouble in the no-load conditions.

Subject-Verb Agreement, Relative-Clause Complexity, and Embedding. Why do interactions with Articulatory Suppression show up consistently with Embedding Type, but not with Clause Type? I believe that this has to do with the complexity of the two integrations. In order to check the agreement features of the main clause that contains the center-embedded clause, a larger part of the back-up must be addressed than for the checking of the agreement features within subject-relative clauses. This is illustrated in (12):

(12) De *fotograaf* die de uitgevers *accepteert, prijst* de studenten
The photographer that the publishers accepts, praises the students
'The photographer who accepts the publishers praises the students'

If we go back to Sternberg's (1969) interaction logic, we can observe that super-additive interactions only occur when a particular system is taxed to its limits.[4] Because the integration of subject and verb in the main clause containing a center-embedded clause is more complex than that of the subject and the verb within a subject-relative clause, Embedding Type is more likely to interact with Articulatory Suppression than Clause Type is, since the distance in the phonological representation is greater.

In the context of Sternberg's (1969) interaction logic, it is noteworthy that a superadditive interaction of Clause Type x Articulatory Suppression occurs only in the most complex experiment, Experiment 2. This experiment used reversible materials, which are more demanding to the participants than nonreversible sentences. Additionally, the materials were somewhat longer than in the other reversible experiment, Experiment 1.

Thus, it seems that feature checking in subject-relative clauses is less taxing than in center-embedded constructions and that this checking is only causing interactions if the subject-relative clause is really complex.

---

[4] For those who are not familiar with Sternberg's (1969) interaction logic, the reader is referred to the subsection on Sternberg's Interaction Logic on page 18 of this book.

The Phonological Loop and Adjectival Load

With the integration of adjectives, the checking of agreement features is not necessary for structural disambiguation; therefore, no interactions with Articulatory Suppression are predicted. In Experiment 5b, we saw that Adjectival Load produced independent effects with Embedding Type and with Articulatory Suppression. Experiment 5a did produce a three-way interaction of these factors, but this was apparently the result of task strategies. For a detailed coverage of this phenomenon, the reader is referred to the Discussion of Experiment 5a on page 111.

The fact that Articulatory Suppression did not produce consistent superadditive interactions with Adjectival Load is in line with the assumption that feature checking is not done during all integration types that involve agreement features.[5] If feature checking is not crucial for sentence interpretation and ambiguity resolution, it is not carried out according to the model.

The Phonological Loop and Propositional Complexity

Because propositional and syntactic processing are independent in the model, there should not be any superadditive interactions of Propositional Complexity with syntactic manipulations, nor with Articulatory Suppression because the articulatory-suppression manipulation is also taken to affect only the syntactic stage. This is indeed what was found: no superadditive interactions with syntactic or articulatory-suppression manipulations were obtained in either of the experiments.

Summary

In the current subsection, the potential interactions predicted by the model will be summarized. Table 1 gives an overview of predicted and obtained superadditive two-way interactions in Experiments 1 to 5b.

---

[5] Adjectives and nouns have an agreement relation in Dutch, just as subjects and verbs. For more information on adjective-noun agreement in Dutch, the reader is referred to the subsection titled Adjective-Noun Agreement in Dutch on page 105.

Table 1
Overview of Predicted and Obtained Superadditive Two-Way Interactions in Experiments 1 to 5b

Interaction	Predicted	Obtained
Clause Type x Propositional Complexity	no	Exp. 3x[ac]
Clause Type x Embedding Type	no	never
Clause Type x Articulatory Suppression	yes	Exp. 2[b]
Articulatory Suppression x Propositional Complexity	no	never
Articulatory Suppression x Embedding Type	yes	Exp. 2[a] Exp. 4[bc]
Articulatory Suppression x Adjectival Load	no	Exp. 5a[bc]
Adjectival Load x Embedding Type	no	Exp. 5a[b]

[a]in error data.
[b]in RT data.
[c] speed-accuracy trade off.

Relating the Model to Language Pathology Data

So far, we have only covered the model and its relation to the WM resource discussion. However, a number of phenomena observed in people suffering from language impairment are important for the current model. These phenomena will be discussed in the current section. The focus will be on the insensitivity to phonologically nonsalient information in children suffering from SLI and on the inability to process center-embedded constructions in a patient suffering from aphasia who exhibits decreased functioning of the phonological loop.

The Model and the Phonological Salience Hypothesis of SLI

SLI occurs in children who develop normally, have normal intelligence, but do not develop age-appropriate language skills. Different proposals have been made about the nature of SLI. Some researchers have argued for a genetic basis of certain components of grammar (e.g., Gopnik, 1997; Pinker, 1989), whereas others adhere the view that perceptual or phonological deficits are the main cause of SLI (e.g., Bishop, 1997; Joannisse & Seidenberg, 1998; Leonard, 1997).

Those who advocate that SLI is the result of a phonological deficit have proposed that SLI children have difficulty processing phonologically nonsalient elements such as past tenses and plurals (Joanisse & Seidenberg, 1998). If these children have trouble with these elements, they may exhibit syntactic problems consequently because syntactic analysis partly depends on such phonologically subtle morphosyntactic information (Bishop, 1997).

If we assume that (at least in some SLI children) SLI is indeed the result of a phonological deficit, how would this relate to the model of WM and sentence processing outlined above? It seems that the locus of this impairment could be in two places.

First, it may be that the perception of the morphosyntactic information is hampered. This will result in a suboptimal phonological back-up representation.

Alternatively, these children may suffer from a phonological memory deficit. In that case, they process the phonology correctly, but have difficulty with the storage of the phonological back-up. In either case, the deficit will result in a degraded phonological back-up representation of the sentence. This situation is depicted in Figure 2.

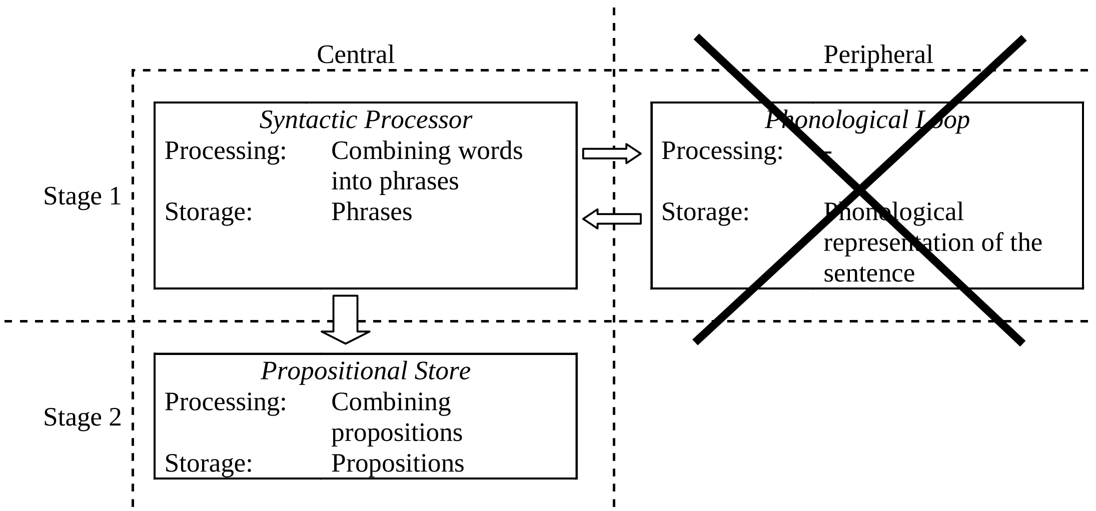


Figure 2. The locus of impairment in SLI children with phonological problems.

In summary, the current model offers an explanation of processing phenomena in healthy volunteers that is similar to the account of the nature of SLI proposed by for example. Bishop (1997), Joannisse & Seidenberg (1998), and Leonard (1997). Both accounts assume that syntactic difficulties may arise as a consequence of phonological processing difficulty. In the current research, this difficulty was provoked by articulatory suppression, whereas in SLI children, a phonological processing deficit may be the cause of the difficulty.

Evidence from Aphasia

In addition to the SLI data, there is also evidence from aphasia that suggests a role for phonological memory during long-distance subject-verb integration. Bastiaanse (1993) reported a Dutch patient, S.L., who had impaired functioning of the phonological loop component of WM, as shown by his performance on a nonsense word repetition task (Bastiaanse, unpublished). In that particular task, S.L. showed good performance on real words (8 out of 9 items repeated correctly) and extremely poor performane on the nonwords (only 1 out of 9 items repeated correctly). S.L.'s sentence comprehension abilities were relatively spared, but when tested on a sentence repetition task (Graetz et al., 1992), he proved unable to repeat long center-embedded sentences.

This pattern of performance suggests, like the SLI data, that phonological memory is used during the long-distance integration of subject and verb. When phonological memory is impaired, performing that integration becomes impossible.

Open Issues and Suggestions for Future Research

A number of issues are still unresolved with respect to the current proposal of a partition of WM resources deployed in sentence processing. These will be discussed in the current section.

The first subsection will deal with the nature of the phonological back-up representation. Waters et al. (1987) have proposed that it is stored in an articulatory format. This means that in terms of Baddeley's (1986) WM model, the back-up should reside in the articulatory rehearsal mechanism. However, Baddeley et al. (1987) have proposed that it is kept in the phonological store component of the phonological loop. These two hypotheses will be explored in more detail and related to the data of Experiments 1 to 5b.

The second subsection will discuss whether the syntactic processor has only one additional memory store (i.e., the phonological loop) or maybe more. Martin & Romani (1994) have proposed that an additional lexicosemantic memory store is deployed during sentence processing. Their proposal will be evaluated based on Experiments 1 to 5b.

Finally, it will be discussed whether the lack of super-additive interactions of Adjectival Load x Articulatory Suppression in Experiments 5a and b is due to the fact that feature checking is not carried out (as proposed in the model). An alternative explanation could be that these features are checked locally in the three-adjective sentences.

Storing the Back-Up: Articulatory Rehearsal or Phonological Storage?

In Chapters 3 and 4, there were two interactions with Articulatory Suppression: in the reversible sentences of Experiment 2, there was an interaction with Clause Type in which subject-relative clauses yielded higher error rates than their object-relative counterparts. Additionally, in both center-embedding experiments, there were interactions with Embedding Type. These results suggest that under some circumstances, the phonological loop is addressed in syntactic processing.

At present, there are two possibilities with respect to the nature of that role. One is that the phonological loop is used for the storage of unintegrated material, for example the matrix subject of center-embedded structures (Withaar, Stowe, & Hoeks, 2000). The alternative hypothesis is that during integration processes, a phonological back-up representation of the sentence is addressed in order to check the agreement features of the different nouns and verbs in a particular clause (cf. Baddeley et al., 1987).

The Articulatory Rehearsal Mechanism and the Back-Up. Let us examine the first hypothesis in more detail. The interaction of Embedding Type x Articulatory Suppression suggests that the matrix subject, which cannot be integrated into the existing structure before the appearance of the matrix verb, is kept active in the phonological loop. If this idea is correct, the phonological loop would have to keep the material active using articulatory rehearsal, because the phonological store is taken to be a passive storage component (Baddeley, 1986).

Neuro-imaging findings presented by Caplan, Alpert, & Waters (1998) and by Paulesu, Frith, & Frackowiak (1993) also point in this direction. Caplan et al. found that during syntactic processing, as tested by comparing subject- and object-relative clauses, yielded increased regional cerebral bloodflow in Broca's area (BA 44). Paulesu et al. report that the same area shows increased activation during articulatory rehearsal. Thus, it seems likely that during syntactic processing, articulatory rehearsal is deployed.

However, if unintegrated material or θ-less NPs are kept in an articulatory format, one would expect there to be no differences between subject- and object-relative constructions with respect to their processing under articulatory suppression. This difference is expected because the two sentence types have the same amount of NPs that must be stored before they can be integrated with the verb.

The Phonological Store and the Back-Up. The fact that there is a difference between subject and object relatives and their reactions to articulatory suppression suggests that a back-up representation of the sentence is addressed during feature checking at the embedded verb. Baddeley et al. (1987) have suggested that during sentence processing, the phonological store acts as a mnemonic window in which a back-up is kept for reference during syntactic processing.

At first sight, the assumption that the phonological store is addressed during comprehension is somewhat unexpected. Those familiar with the Baddeley (1986) model may object to this idea on the grounds that articulatory suppression is a manipulation that loads the rehearsal mechanism, and not the phonological store.

However, it is possible that articulatory suppression affected the phonological store in the present experiments. First, Gupta & MacWhinney (1995) pointed out that articulatory suppression also has an additional effect on the phonological store. The phonological store can be taxed by irrelevant speech (Baddeley, 1986). In irrelevant-speech conditions, participants have to perform a certain task while listening to unrelated speech sounds. Gupta & MacWhinney posited that during articulatory-suppression tasks, participants hear the irrelevant speech sounds they produce themselves. Therefore, articulatory suppression does not seem to be such a pure method of taxing the articulatory rehearsal mechanism only. Thus, interactions with articulatory suppression turn out to be compatible with Baddeley et al.'s (1987) proposal of a phonological back-up kept in the phonological store component of WM.

Second, one task of the rehearsal mechanism is to convert printed text to an acoustic format that is accessible to the phonological store. If functioning of the rehearsal mechanism is hampered by articulatory suppression, the acoustic representation may be distorted. Thus, readers who perform an articulatory-suppression task can only maintain a degraded representation of the sentence in their phonological stores.

Evidence for a Separate Semantic WM Component?

It may be that during the integration of stacked adjectives, participants used different cues than during θ-role assignment in Experiments 1 to 4. In the latter experiments, integration of the matrix subject and verb depended on the availability of agreement information. In Experiments 5a and b, the adjective-noun integration was semantic in nature: subjects had to judge whether certain adjective-noun combinations were compatible or not.

Martin & Romani (1994) have argued for a dissociation of phonological and semantic memory based on the performance of two brain-damaged patients. One patient, E.A., had a disturbance in phonological memory, but intact semantic memory: she did much worse on a nonword span than on a word span task. A different patient, A.B., had the opposite pattern of impairment, as shown by his poor performance on a rhyme probe task and his good performance on a category probe task.

On the basis of these data, Martin & Romani (1994) proposed an extension of the number of slave systems in Baddeley's (1986) WM model. Baddeley's model consists of a central executive, which is involved in the processing and storage of information, and two so-called slave systems. These slave systems are units that can be deployed when the central executive needs additional storage capacity. One of these systems is tuned to the processing of visual information: the visuo-spatial sketchpad. The other is involved in the storage of phonological information, the phonological loop. In addition to these two systems, which have been proposed by Baddeley, Martin & Romani assume a semantic slave system as well.

The assumption of a separate memory store for semantic information fits very nicely with the data from the Experiments 1 to 5b. It may be that no interactions of Adjectival Load x Articulatory Suppression were found, because the two complexities draw on different resource pools. The suppression manipulation taxes the phonological loop, whereas during the adjective-noun integration, the semantic memory system is addressed.

Interrupting the Adjective-Noun Agreement Relation

A third issue that must be considered more closely is the assumption that the phonological loop is not involved in the checking of agreement features during adjective-noun integration. I proposed that this was not the case because feature checking is only carried out in cases where it may guide structural ambiguity resolution.

This proposal was made because long-distance adjective-noun integration as tested by comparing sentences with one adjective and sentences with three adjectives, as shown in (13a and b) respectively:

(13) a. De barkeeper heeft de chips opgegeten terwijl de verregende klant niets doorhad
The barkeeper has the chips eaten while the drenched customer nothing noticed
'The barkeeper ate the chips while the drenched customer did not notice anything'

b. De barkeeper heeft de chips opgegeten terwijl de verregende hongerige trouwe klant niets doorhad
The barkeeper has the chips eaten while the drenched hungry loyal customer nothing noticed
'The barkeeper ate the chips while the drenched hungry loyal customer did not notice anything'

The adjectives in (13b) must be integrated with the head noun of the phrase over a longer distance.

However, given the fact that all adjectives agree with the head noun, it is not entirely clear whether this is a local or a long-distance feature checking process. It is true that the first adjective and the head noun are far apart, but the first adjective has the same agreement features as the second and the third adjective. It may well be that the agreement features of the first and second adjective are checked locally. Likewise, the agreement features of the second and third adjective may be checked locally. Then finally, there is only local feature checking of adjective-noun agreement between the third adjective and the head noun.

Because this interpretation of the data cannot be ruled out on the basis of the data in Experiments 5a and b, this issue requires further exploration. In Dutch, it is possible to disrupt the agreement chain of adjectives by adding a modifying clause with a perfect participle as its head, as shown in (14):

(14) De schilder heeft, toen hij het mooie, ver van de snelweg gelegen huis verfde, z'n been gebroken
The painter has, while he the beautiful, far from the freeway located house painted, his leg broken
'The painter broke his leg while he was painting the beautiful house, which was located far away from the freeway'

A suggestion for future research would be to compare the kind of constructions in (15) with sentences that lack such an intervening clause with a perfect participle. An example of such a sentence is given in (15):

(15) De schilder heeft, toen hij het mooie huis verfde, z'n been gebroken
The painter has, while he the beautiful house painted, his leg broken
'The painter broke his leg while he was painting the beautiful house'

Summary

Despite the fact that some issues need further investigation, this study produced several interesting insights. First, the data presented in this study show that sentences are processed in two stages, a syntactic and a propositional stage (in line with Berwick & Weinberg, 1984 and Caplan & Waters, 1999). Second, it showed that during the first stage, but not the second, a phonological back-up representation is addressed (contra Caplan & Waters, 1999).

The fact that this back-up is addressed becomes apparent when participants read sentences with long-distance subject-verb integrations under articulatory suppression. The availability of the phonological back-up representation is reduced under articulatory suppression. The finding that reduced availability of phonological information may result in syntactic processing problems is also in line with the data from the SLI literature (e.g., Bishop, 1997; Joannisse & Seidenberg, 1998; Leonard, 1997).

References

Babcock, R., & Salthouse, T. (1990). Effects of increased processing demands on age differences in working memory. *Psychology and Aging, 5,* 421-428.

Baddeley, A.D. (1986). *Working Memory.* Oxford, England: Clarendon Press.

Baddeley, A.D. & Hitch, G.J. (1974). Working memory. In: G. Bower (Ed.), *The Psychology of Learning and Motivation, 8* (pp. 47-90). New York, NY: Academic Press.

Baddeley, A.D., Vallar, G., & Wilson, B. (1987). Sentence comprehension and phonological memory: some neuropsychological evidence. In: Coltheart, M. (Ed.), *Attention and Performance XII: The Psychology of Reading* (pp. 509-529). Hillsdale, NJ.: LEA Press.

Bastiaanse, Y.R.M. (1993). *Studies in Aphasia.* University of Groningen, The Netherlands: PhD Thesis.

Berwick, R.C., & Weinberg, A.S. (1984). *The Grammatical Basis of Linguistic Performance: Language Use and Acquisition.* Cambridge, MA: MIT Press.

Bever, T.G. (1970). The cognitive basis for linguistic structure. In: Hayes, J.R. (Ed.), *Cognition and the Development of Language* (pp. 279-362). New York, NY: John Wiley.

Bishop, D.V.M. (1997). *Uncommon Understanding: Development and Disorders of Language Comprehension in Children.* Hove, England: Psychology Press.

Brooks, J.O., III & Watkins, M.J. (1990). Further evidence of the intricacy of memory span. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1134-1141.

Burnage, G. (1990). *CELEX, A Guide for Users.* Nijmegen, The Netherlands: Drukkerij SSN.

Caplan, D., Alpert, N., & Waters, G. (1998). Effects of syntactic structure and propositional number on patters of regional blood flow. *Journal of Cognitive Neuroscience, 10,* 541-552.

Caplan, D., & Hildebrandt, N. (1988). *Disorders of Syntactic Comprehension.* Cambridge, MA: MIT Press.

Caplan, D., & Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*, 77-94.

Caramazza, A., Basili, A.G., Koller, J., & Berndt, R.S. (1981). An investigation of repetition and language processing in a case of conduction aphasia. *Brain and Language*, *14*, 235-271.

Chomsky, N.A. (1981). *Lectures on Government and Binding.* Cambridge, MA: MIT Press.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450-466.

Daneman, M., & Merikle, P.M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review, 3,* 422-433.

Frazier, L., (1979). *On Comprehending Sentences: Syntactic Parsing Strategies.* Bloomington, IN: Indiana University Linguistics Club.

Frazier, L., (1987). Sentence processing: A tutorial review. In: Coltheart, M. (Ed.), *Attention and Performance XII: The Psychology of Reading* (pp. 601-681). Hillsdale, NJ.: LEA Press.

Frazier, L., & Clifton, C. (1996). *Construal.* Cambridge, MA: MIT Press.

Frazier, L., & Flores d'Arcais, G. (1989). Filler-driven parsing: A study of gap-filling in Dutch. *Journal of Memory and Language, 28,* 331-344.

Frazier, L., & Rayner K. (1987). Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. *Journal of Memory and Language, 26,* 505-526.

Frazier, L., & Rayner, K. (1988). Parameterizing the language processing system: Left-vs. right-branching within and across languages. In: Hawkins, J. (Ed.), *Explaining Language Universals* (pp. 247-279). Oxford, England: Blackwell.

Freedman, D., Pisani, R., & Purves, R. (1998). *Statistics* (2nd ed.). New York, NY: Norton.

Gathercole, S.E., & A.D. Baddeley (1993). *Working Memory and Language.* Hillsdale, NJ: LEA Press.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68,* 1-76.

Gopnik, M. (1997). Language deficits and and genetic factors. *Trends in Cognitive Neurosciences, 1,* 5-9.

Graetz, P., de Bleser, R., & Willmes, K. (1992). *Akense Afasietest* [Aachen Aphasia Test]. Lisse, The Netherlands: Swets & Zeitlinger.

Gunter, T. (1995). *Aging and the Processing of Semantic Information: an Electrophysiological Approach.* University of Groningen, The Netherlands: PhD Thesis.

Gupta, P., & MacWhinney, B. (1995). Is the articulatory loop articulatory or auditory? Reexamining the effects of concurrent articulation on immediate serial recall. *Journal of Memory and Language, 34*, 63-88.

Joanisse, M.F., & Seidenberg, M.S. (1998). Specific language impairment: A deficit in grammar or processing? *Trends in Cognitive Sciences, 2,* 240-247.

Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122-149.

Just, M.A., Carpenter, P.A., Keller, T.A., Eddy, W.F., & Thulborn, K.R. (1996). Brain activation modulated by sentence comprehension. *Science, 274,* 114-116.

Kaan, E. (1997). *Processing Subject-Object Ambiguities in Dutch.* University of Groningen, The Netherlands: PhD Thesis.

King, J., & Just, M.A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30,* 580-602.

La Pointe, L., & Engle, R.(1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16(6),* 1118-1133.

Lamers, M.J.A. (2001). *Sentence Processing Using Syntactic, Semantic, and Thematic Information*. University of Groningen, The Netherlands: PhD Thesis.

Leonard, L. (1997). *Children with Specific Language Impairments.* Cambrigde, MA: MIT Press.

MacDonald, M.C., & Christiansen, M. (in press). Individual differences without working memory: A reply to Just & Carpenter and Waters & Caplan. To appear in: *Psychological Review.*

MacDonald, M.C., Pearlmutter, N.J., & Seidenberg, M.S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676-703.

Mak, W.M. (2001). *Processing Relative Clauses: Effects of Pragmatic, Semantic, and Syntactic Variables.* University of Nijmegen, The Netherlands: PhD Thesis.

Marcus, P.M. (1980). *A Theory of Syntactic Recognition for Natural Language.* Cambridge, MA: MIT Press.

Martin, R.C. (1993). Short-term memory and sentence processing: Evidence from neuropsychology. *Memory & Cognition, 21,* 176-183.

Martin, R.C. & Romani, C. (1994). Verbal working memory and sentence processing: a multiple component view. *Neuropsychology, 8,* 506-523.

Martin, R.C., Wogalter, M.S., & Forlano, J.G. (1988). Reading comprehension in the presence of unattended speech and music. *Journal of Memory and Language, 27,* 382-398.

Miller, G. & Isard, S. (1964). Free recall of self-embedded English sentences. *Information and Control, 7,* 292-303.

Miyake, A., Carpenter, P.A., & Just, M.A. (1996). A capacity approach to syntactic comprehension disorders: Making normal adults perform like aphasic patients. *Cognitive Neuropsychology, 11*, 671-717.

Pinker, S. (1994). *The Language Instinct: The New Science of Language and Mind.* New York, NY: Morrow.

Saffran, E.M., & Marin, O.S.M. (1975). Immediate memory ofr word lists and sentences in a patient with deficient auditory short-term memory. *Brain and Language, 2,* 420-433.

Salthouse, T. (1990). Working memory as a processing resource in cognitive aging. *Developmental Review, 10,* 101-124.

Salthouse, T. (1994). The aging of working memory. *Neuropsychology, 8,* 535-543.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125,* 4-27.

Stowe, L.A. (1986). Parsing WH-constructions: Evidence for on-line gap location. *Language and Cognitive Processes, 1,* 227-245.

Stowe, L.A. (1991). Ambiguity resolution: Behavioral evidence for a delay. In: Hammond, K.J., & Gentner, D. (Eds.), *Proceedings Thirteenth Annual Conference of the Cognitive Science Society.* (pp.257-262). Hillsdale, NJ: LEA Press.

Stowe, L.A., Withaar, R.G., Wijers, A.A., Broere, C.A.J., & Paans, A.M.J. (2002). Encoding and storage in working memory during sentence comprehension. In: Merlo, P., & Stevenson, S. (Eds.), *Proceedings of the Eleventh CUNY Conference on Sentence Processing* (pp. 181-205). Philadelphia, PA: John Benjamins.

Talairach, J., & Tournoux, P. (1988). *A Co-planar Stereotactic Atlas of the Human Brain.* Stuttgart, Germany: Thieme.

Tanenhaus, M.K., & Trueswell, J.C. (1995). Sentence comprehension. In: Miller, J.S., & Eimas, P.D. (Eds.), *Speech, Language, and Communication.* (pp. 217-262). San Diego, CA: Academic Press.

Vos, S.H. (1999). *Verbal Working Memory and Sentence Processing: An Electrophysiological Investigation.* University of Nijmegen, The Netherlands: PhD Thesis.

Wanner, E., & Maratsos, M. (1978). An ATN approach to comprehension. In: Halle, M., Bresnan, J., & Miller, G., (Eds.), *Linguistic Theory and Psychological Reality.* (pp. 119-161). Cambridge, MA: MIT Press.

Waters, G., Caplan, D., & Hildebrandt, N. (1987). Working memory and written sentence comprehension. In: Coltheart, M. (Ed.), *Attention and Performance XII: The Psychology of Reading* (pp. 531-555). Hillsdale, NJ: LEA Press.

Withaar, R.G. (2001a). *Thematic Role Assignment: When is it Vulnerable to Articulatory Suppression?* Neurological Basis of Language Conference, Groningen, The Netherlands. Retrieved February, 12, 2002 from Homepage Rienk Withaar: http://www.let.rug.nl/~withaar/research/pdf/nbl.pdf.

Withaar, R.G. (2001b). *The Role of Inner Speech in Sentence Processing.* Colloquium Center for Language and Cognition Groningen, The Netherlands, February 23rd, 2001. Retrieved February, 12, 2002 from Homepage Rienk Withaar: http://www.let.rug.nl/~withaar/research/tc2001.html

Withaar, R.G., & Stowe, L.A. (1999a). *One vs. Two Sentence Processing Working Memories.* Sixth Annual Meeting of the Cognitive Neuroscience Society. Washington, DC. Retrieved February, 12, 2002 from Homepage Rienk Withaar: http://www.let.rug.nl/~withaar/research/pdf/cns1999a.pdf.

Withaar, R.G., & Stowe, L.A., (1999b). *Re-examining Evidence for Separate Sentence Processing Resources.* AMLaP 1999, Edinburgh, Scotland. Retrieved February, 12, 2002 from Homepage Rienk Withaar: http://www.let.rug.nl/~withaar/research/pdf/amlap1999.pdf.

Withaar, R.G., Stowe, L.A., & Hoeks, J.C.J. (2000). *Toward a Two-Stage Storage Model of Sentence Processing.* CUNY 2000, San Diego, CA. Retrieved February, 12, 2002 from Homepage Rienk Withaar: http://www.let.rug.nl/~withaar/research/pdf/cuny2000.pdf.

Appendices

Materials: Experiment 1

Het waren de broeders die het liefje beminden/beminde (1PSR/1POR)

Het waren de misdadigers die de priester beïnvloedden/beïnvloedde (1PSR/1POR)

Het zijn de dieven die de commissaris onderschatten/onderschat (1PSR/1POR)

Het was de gemeente die de besturen vernederde/vernederden (1PSR/1POR)

Het was de matroos die de vissers hinderde/hinderden (1PSR/1POR)

Het is de fabrikant die de koopmannen ondersteunt/ondersteunen (1PSR/1POR)

Het waren de dames die de minnaar ontwaarden/ontwaarde (1PSR/1POR)

Het waren de ridders die de reus velden/velde (1PSR/1POR)

Het zijn de diplomaten die de admiraal identificeren/identificeert (1PSR/1POR)

Het was de genie die de Grieken bezocht/bezochten (1PSR/1POR)

Het was de ober die de bakkers misleidde/misleidden (1PSR/1POR)

Het is de portier die de gezelschappen plaagt/plagen (1PSR/1POR)

Het waren de duivels die de schepper vervloekten/vervloeken (1PSR/1POR)

Het waren de ruiters die de krijger verjoegen/verjoeg (1PSR/1POR)

Het zijn de professoren die de officier groeten/groet (1PSR/1POR)

Het was de hoofdpersoon die de prinsessen zoende/zoenden (1PSR/1POR)

Het was de zendeling die de rechercheurs haatte/haatten (1PSR/1POR)

Het is de vreemdeling die de experts verdraagt/verdragen

Het waren de filosofen die de theoloog ontweken/ontweek (1PSR/1POR)

Het zijn de bewakers die de luitenant kwellen/kwelt (1PSR/1POR)

Het zijn de waarnemers die de psychiater kwetsen/kwetst (1PSR/1POR)

Het was de buurman die de bezoekers doodde/doodden (1PSR/1POR)

Het was de maagd die de engelen omhelsde/omhelsden (1PSR/1POR)

Het is de weduwe die de poezen verrast/verrassen (1PSR/1POR)

Ik omschreef de cadet die de opvolgers benoemde/benoemden (2PSR/2POR)

Ik wantrouw de brouwer die de handelaars registreert/registreren (2PSR/2POR)

Jij lokte de rat die de katers beet/beten (2PSR/2POR)

Jij inspireert de notarissen die de deskundige veroordelen/veroordeelt (2PSR/2POR)

Jij ontslaat de pater die de bruiden mist/missen (2PSR/2POR)

Jij verontschuldigt de toeristen die de kelner benaderen/benadert (2PSR/2POR)

Ik ondervroeg de functionarissen die de producent vervoerden/vervoerde (2PSR/2POR)

Ik beschuldig de kunstenaar die de componisten citeert/citeren (2PSR/2POR)

Ik wreek de non die de pastoors berooft/beroven (2PSR/2POR)

Jij keurde de therapeut die de sociologen rook/roken (2PSR/2POR)

Jij schokt de politiemannen die de baas waarschuwen/waarschuwt (2PSR/2POR)

Jij verraadt de consumenten die de majoor schoppen/schopt (2PSR/2POR)

Ik bevrijdde de arts die de architecten verborg/verborgen (2PSR/2POR)

Ik benauw de gasten die het individu beschermen/beschermt (2PSR/2POR)

Ik passeer de leerling die de tegenstanders verslaat/verslaan (2PSR/2POR)

Jij bedreigde de verenigingen die de wethouder waardeerden/waardeerde (2PSR/2POR)

Jij verwijderde de ingenieurs die het comité verwarden/verwarde (2PSR/2POR)

Jij straft de juffrouw die de kappers irriteert/irriteren (2PSR/2POR)

Ik noteer de geleerden die de rector typeren/typeert (2PSR/2POR)

Ik test de kerels die de zakenman troosten/troost (2PSR/2POR)

Jij verzorgde de patiëntes die de boerin bedrogen/bedroog (2PSR/2POR)

Jij bewondert de Engelsman die de kennissen vermoordt/vermoorden (2PSR/2POR)

Jij onderbreekt de betrokkene die de deelnemers achterhaalt/achterhalen (2PSR/2POR)

Jij verdedigt de communisten die het regime isoleren/isoleert (2PSR/2POR)

Materials: Experiment 2

Het bezoek vermoeit de chauffeurs die de passagier bedanken/bedankt (RESR/REOR)
De chauffeurs die de passagier bedanken/bedankt, vermoeien het bezoek (CESR/CEOR)

De koning ontmoedigt de hertogen die het koor vrezen/vreest (RESR/REOR)
De hertogen die het koor vrezen/vreest, ontmoedigen de koning (CESR/CEOR)

De Vlaming verwent de dames die de minnaar ontwaren/ontwaart (RESR/REOR)
De dames die de minnaar ontwaren/ontwaart, verwennen de Vlaming (CESR/CEOR)

De directeur krabt de secretarissen die de spreker beledigen/beledigt (RESR/REOR)
De secretarissen die de spreker beledigen/beledigt, krabben de directeur (CESR/CEOR)

Het congres schaadt de getuigen die de burgemeester ondermijnen/ondermijnt (RESR/REOR)
De getuigen die de burgemeester ondermijnen/ondermijnt, schaden het congres (CESR/CEOR)

De farao benijdt de genieën die de Griek bezoeken/bezoekt (RESR/REOR)
De genieën die de Griek bezoeken/bezoekt, benijden de farao (CESR/CEOR)

De miss versiert de broeders die het liefje beminnen/bemint (RESR/REOR)
De broeders die het liefje beminnen/bemint, versieren de miss (CESR/CEOR)

De hoogleraar karakteriseert de diplomaten die de admiraal identificeren/identificeert (RESR/REOR)
De diplomaten die de admiraal identificeren/identificeert, karakteriseren de hoogleraar (CESR/CEOR)

Het publiek wantrouwt de brouwers die de handelaar registeren/registreert (RESR/REOR)
De brouwers die de handelaar registeren/registreert, wantrouwen het publiek (CESR/CEOR)

Het college verbiedt de commissies die de socialist corrigeren/corrigeert (RESR/REOR)
De commissies die de socialist corrigeren/corrigeert, verbieden het college (CESR/CEOR)

De kolonels overheersen de marine die de generaals verdrijft (RESR/REOR)
De generaals die de kolonel verdrijven/verdrijft, overheersen de marine (CESR/CEOR)

De mijnheer snapt de knapen die het mannetje redden/redt (RESR/REOR)
De knapen die het mannetje redden/redt, snappen de mijnheer (CESR/CEOR)

De medewerker grijpt de gekken die de bejaarde belemmeren/belemmert (RESR/REOR)
De gekken die de bejaarde belemmeren/belemmert, grijpen de medewerker (CESR/CEOR)

De regent verdenkt de acteurs die de collega ontstellen/ontstelt (RESR/REOR)
De acteurs die de collega ontstellen/ontstelt, verdenken de regent (CESR/CEOR)

De manager stimuleert de adviseurs die het ministerie steunen/steunt (RESR/REOR)
De adviseurs die het ministerie steunen/steunt, stimuleren de manager (CESR/CEOR)

De zendeling wreekt de nonnen die de pastoor beroven/berooft (RESR/REOR)
De nonnen die de pastoor beroven/berooft, wreken de zendeling (CESR/CEOR)

De graaf verbant de bewakers die de luitenant kwellen/kwelt (RESR/REOR)
De bewakers die de luitenant verbannen/verbant, kwellen de graaf (CESR/CEOR)

De voorganger weerstaat de duivels die de schepper vervloeken/vervloekt (RESR/REOR)
De duivels die de schepper vervloeken/vervloekt, weerstaan de voorganger (CESR/CEOR)

De zeeman imponeert de fabrikanten die de koopman ondersteunen/ondersteunt (RESR/REOR)
De fabrikanten die de koopman ondersteunen/ondersteunt, imponeren de zeeman (CESR/CEOR)

De massa stoort de filosofen die de theoloog ontwijken/ontwijkt (RESR/REOR)
De filosofen die de theoloog ontwijken/ontwijkt, storen de massa (CESR/CEOR)

De gestalte beangstigt de matrozen die de visser hinderen/hindert (RESR/REOR)
De matrozen die de visser hinderen/hindert, beangstigen de gestalte (CESR/CEOR)

De huisvrouw verwelkomt de reizigers die de bediende vertrouwen/vertrouwt (RESR/REOR)
De huisvrouwen die de bediende vertrouwen/vertrouwt, verwelkomen de reiziger (CESR/CEOR)

De werkgroep begeleidt de leveranciers die de econoom beoordelen/beoordeelt (RESR/REOR)
De leveranciers die de econoom beoordelen, begeleiden de werkgroep (CESR/CEOR)

De liberaal verdedigt de communisten die het regime isoleren/isoleert (RESR/REOR)
De communisten die het regime isoleren/isoleert, verdedigen de liberaal (CESR/CEOR)

De vertellers inspireren de notaris die de deskundigen veroordeelt/veroordelen (RESR/REOR)
De notaris die de deskundigen veroordeelt/veroordelen, inspireert de vertellers (CESR/CEOR)

De curatoren onderzoeken de voogd die de cliëntes nadert/naderen (RESR/REOR)
De voogd die de cliëntes nadert/naderen, onderzoekt de curatoren (CESR/CEOR)

De neven bevrijden de arts die de architecten verbergt/verbergen (RESR/REOR)
De arts die de architecten verbergt/verbergen, bevrijdt de neven (CESR/CEOR)

De tandartsen ontbieden de ober die de bakkers misleidt/misleiden (RESR/REOR)
De ober die de bakkers misleidt/misleiden, ontbiedt de tandartsen (CESR/CEOR)

De verkopers onderbreken de intellectueel die de advocaten verbijstert/verbijsteren (RESR/REOR)
De intellectueel die de advocaten verbijstert/verbijsteren, onderbreekt de verkopers (CESR/CEOR)

De notabelen bespieden de arbeidersklasse die de democraten toespreekt/toespreken (RESR/REOR)
De arbeidersklasse die de democraten toespreekt/toespreken, bespiedt de notabelen (CESR/CEOR)

De geallieerden introduceren de buitenlander die de kiezers erkent/erkennen (RESR/REOR)
De buitenlander die de kiezers erkent/erkennen, introduceert de geallieerden (CESR/CEOR)

De studenten prijzen de fotograaf die de uitgevers accepteert/accepteren (RESR/REOR)
De fotograaf die de uitgevers accepteert/accepteren, prijst de studenten (CESR/CEOR)

De werknemers ondervragen de functionaris die de producenten vervoert/vervoeren (RESR/REOR)
De functionaris die de producenten vervoert/vervoeren, ondervraagt de werknemers (CESR/CEOR)

De monniken begeren de meid die de dominees aait/aaien (RESR/REOR)
De meid die de dominees aait/aaien, begeert de monniken (CESR/CEOR)

De monseigneurs ontvangen de hoofdredacteur die de drukkers begunstigt/begunstigen (RESR/REOR)
De hoofdredacteur die de drukkers begunstigt/begunstigen, ontvangt de monseigneurs (CESR/CEOR)

De commissies peilen de betrokkene die de deelnemers achterhaalt/achterhalen (RESR/REOR)
De betrokkene die de deelnemers achterhaalt/achterhalen, peilt de commissies (CESR/CEOR)

De kijkers noteren de geleerde die de rectoren typeert/typeren (RESR/REOR)
De geleerde die de rectoren typeert/typeren, noteert de kijkers (CESR/CEOR)

De lords minachten de gemeente die de besturen vernedert/vernederen (RESR/REOR)
De gemeente die de besturen vernedert/vernederen, minacht de lords (CESR/CEOR)

De Romeinen versterken de ruiter die de krijgers verjaagt/verjagen (RESR/REOR)
De ruiter die de krijgers verjaagt/verjagen, versterkt de Romeinen (CESR/CEOR)

De waarnemers omschrijven de cadet die de opvolgers benoemt/benoemen (RESR/REOR)
De cadet die de opvolgers benoemt/benoemen, omschrijft de waarnemers (CESR/CEOR)

De lieden passeren de leerling die de tegenstanders verslaat/verslaan (RESR/REOR)
De leerling die de tegenstanders verslaat/verslaan, passeert de lieden (CESR/CEOR)

De redacteurs raadplegen de assistent die de bewoners ontzet/ontzetten (RESR/REOR)
De assistent die de bewoners ontzet/ontzetten, raadpleegt de redacteurs (CESR/CEOR)

De bankiers vleien de schoonheid die de baronnen verleidt/verleiden (RESR/REOR)
De schoonheid die de baronnen verleidt/verleiden, vleit de bankiers (CESR/CEOR)

De Nederlanders kussen de Fransman die de Italianen scheert/scheren (RESR/REOR)
De Fransman die de Italianen scheert/scheren, kust de Nederlanders (CESR/CEOR)

De knapen aanbidden de auteur die de lezers treft/treffen (RESR/REOR)
De auteur die de lezers treft/treffen, aanbidt de knapen (CESR/CEOR)

De conducteurs omarmen de apotheker die de verantwoordelijken opwacht/opwachten (RESR/REOR)
De apotheker die de verantwoordelijken opwacht/opwachten, omarmt de conducteurs (CESR/CEOR)

De detectives bewapenen de bezitter die de inbrekers onthutst/onthutsen (RESR/REOR)
De bezitter die de detectives bewapent/bewapenen, onthutst de inbrekers (CESR/CEOR)

De barkeepers stompen de gangster die de advocates tart/tarten (RESR/REOR)
De gangster die de advocates tart/tarten, stompt de barkeepers (CESR/CEOR)

Materials: Experiment 3

Dit zijn de politici die het voorstel afwezen (1PSR)
Dit is het voorstel dat de politici afwezen (1POR)

Dit zijn de mariniers die de landing uitvoerden (1PSR)
Dit is de landing die de mariniers uitvoerden (1POR)

Dit zijn de managers die het filiaal overnamen (1PSR)
Dit is het filiaal dat de managers overnamen (1POR)

Dit is de secretaresse die de notulen uittypte (1PSR)
Dit zijn de notulen die de secretaresse uittypte (1POR)

Dit is de knecht die de strobalen opstapelde (1PSR)
Dit zijn de strobalen die de knecht opstapelde (1POR)

Dit is de cineast die de reportages maakte (1PSR)

Dit zijn de reportages die de cineast maakte (1POR)

Dit zijn de adviseurs die de beursgang begeleidden (1PSR)
Dit is de beursgang die de adviseurs begeleidden (1POR)

Dit zijn de mannequins die de lingerie showden (1PSR)
Dit is de lingerie die de mannequins showden (1POR)

Dit zijn de onderwijzers die de staking uitriepen (1PSR)
Dit is de staking die de onderwijzers uitriepen (1POR)

Dit is de verkoper die de Cd-rom's inkocht (1PSR)
Dit zijn de Cd-rom's die de verkoper inkocht (1POR)

Dit is de journalist die de artikelen publiceerde (1PSR)
Dit zijn de artikelen die de journalist publiceerde (1POR)

Dit is de dominee die de preken voorbereidde (1PSR)

Dit zijn de preken die de dominee voorbereidde (1POR)

De vandaal ontliep de agenten die het stadion beveiligden (2PSR)
De vandalen ontliepen het stadion dat de agenten beveiligden (2POR)

De suppoost vond de inbrekers die het beeld meenamen (2PSR)
De suppoosten vonden het beeld dat de inbrekers meenamen (2POR)

De puber schopte de kleuters die de bal gooiden (2PSR)
De pubers schopten de bal die de kleuters gooiden (2POR)

De meiden wasten de baby die de luiers vervuilde (2PSR)
De meid waste de luiers die de baby vervuilde (2POR)

De kranten bespraken de president die de verkiezingen verloor (2PSR)
De krant besprak de verkiezingen die de president verloor (2POR)

De feestgangers ontweken de ober die de borden afruimde (2PSR)
De feestganger ontweek de borden die de ober afruimde (2POR)

De body guard hoorde de fans die de limousine tegenhielden (2PSR)
De body guards hoorden de limousine die de fans tegenhielden (2POR)

De chauffeur vervoerde de koerier die de drugs afleverde (2PSR)
De chauffeurs vervoerden de drugs die de koerier afleverde (2POR)

De raddraaier bespiedde de ME'ers dat het pand ontruimden (2PSR)
De raddraaiers bespiedden het pand dat de ME'ers ontruimden (2POR)

De onverlaten grepen de postbode die de brieven bezorgde (2PSR)
De onverlaat greep de brieven die de postbode bezorgde (2POR)

De experts beoordeelden de juwelier die de horloges repareerde (2PSR)
De expert beoordeelde de horloges die de juwelier repareerde (2POR)

De kinderen waardeerden de oppas die de eieren bakte (2PSR)
Het kind waardeerde de eieren die de oppas bakte (2POR)

Materials: Experiment 4

De bejaarde vergiftigde de buurvrouwen die de sherry opdronken (RESR)
De bejaarden vergiftigden de sherry die de buurvrouwen opdronken (REOR)
De buurvrouwen die de sherry opdronken, prikkelden de bejaarde (CESR)
De sherry die de buurvrouwen opdronken, prikkelde de bejaarden (CEOR)

De bisschop herkende de priesters die het gif mengden (RESR)
De bisschoppen herkenden het gif dat de priesters mengden (REOR)
De priesters die het gif mengden, bedwelmden de bisschop (CESR)
Het gif dat de priesters mengden, bedwelmde de bisschoppen (CEOR)

De conducteur overzag de reizigers die de trein uitstapten (RESR)
De conducteurs overzagen de trein die de reizigers uitstapten (REOR)
De reizigers die de trein uitstapten, intrigeerden de conducteur (CESR)
De trein die de reizigers uitstapten, intrigeerde de conducteurs (CEOR)

De presentatoren kleineerden de producent die de spelshows vertoonde (RESR)
De presentator kleineerde de spelshows die de producent vertoonde (REOR)
De producent die de spelshows vertoonde, motiveerde de presentatoren (CESR)
De spelshows die de producent vertoonde, motiveerden de presentator (CEOR)

De rechercheurs onderzochten de bende die de geldtransporten overviel (RESR)
De rechercheur onderzocht de geldtransporten die de bende overviel (REOR)
De bende die de geldtransporten overviel, verraste de rechercheurs (CESR)
De geldtransporten die de bende overviel, verrasten de rechercheur (CEOR

De sheriffs raadpleegden de pionier die de kaarten optekende (RESR)
De sheriff raadpleegde de kaarten die de pionier optekende (REOR)
De pionier die de kaarten optekende, bevreemdde de sheriffs (CESR)
De kaarten die de pionier optekende, bevreemdden de sheriff (CEOR)

De cameraman filmde de acteurs die het kasteel bezichtigden (RESR)
De cameramannen filmden het kasteel dat de acteurs bezichtigden (REOR)
De acteurs die het kasteel bezichtigden, ontstemden de cameraman (CESR)
Het kasteel dat de acteurs bezichtigden, ontstemde de cameramannen (CEOR)

De conciërge aanschouwde de leraren die de taart opaten (RESR)
De conciërges aanschouwden de taart die de leraren opaten (REOR)
De leraren die de taart opaten, amuseerden de conciërge (CESR)
De taart die de leraren opaten, amuseerde de conciërges (CEOR)

De directie ondersteunde de ondernemers die het bedrijf saneerden (RESR)
De directies ondersteunden het bedrijf dat de ondernemers saneerden (REOR)
De ondernemers die het bedrijf saneerden, verontrustten de directie (CESR)
Het bedrijf dat de ondernemers saneerden, verontrustte de directies (CEOR)

De nieuwslezers hekelden de weerman die de depressies voorspelde (RESR)
De nieuwslezer hekelde de depressies die de weerman voorspelde (REOR)
De weerman die de depressies voorspelde, bedrukte de nieuwslezers (CESR)
De depressies die de weerman voorspelde, bedrukten de nieuwslezer (CEOR)

De correspondenten verafschuwden de prins die de steekpenningen kreeg (RESR)
De correspondent verafschuwde de steekpenningen die de prins kreeg (REOR)
De prins die de steekpenningen kreeg, verwonderde de correspondenten (CESR)
De steekpenningen die de prins kreeg, verwonderden de correspondent (CEOR)

De veeboeren bekeken de monteur die de tractors controleerde (RESR)
De veeboer bekeek de tractors die de monteur controleerde (REOR)
De monteur die de tractors controleerde, tergde de veeboeren (CESR)
De tractors die de monteur controleerde, tergden de veeboer (CEOR)

De columnist veroordeelde de commando's die de kaping beëindigden (RESR)
De columnisten veroordeelden de kaping die de commando's beëindigden (REOR)

De commando's die de kaping beëindigden, schokten de columnist (CESR)
De kaping die de commando's beëindigden, schokte de columnisten (CEOR)

De douanier onderschepte de smokkelaars die de cocaïne verscheepten (RESR)
De douaniers onderschepten de cocaïne die de smokkelaars verscheepten (REOR)
De smokkelaars die de cocaïne verscheepten, verstoorden de douanier (CESR)
De cocaïne die de smokkelaars verscheepten, verstoorde de douaniers (CEOR)

De graaf naderde de bedienden die de vloer schrobden (RESR)
De graven naderden de vloer die de bedienden schrobden (REOR)
De bedienden die de vloer schrobden, verblijdden de graaf (CESR)
De vloer die de bedienden schrobden, verblijdde de graven (CEOR)

De hoogleraren financierden de onderzoeker die de experimenten opzette (RESR)
De hoogleraar financierde de experimenten die de onderzoeker opzette (REOR)
De onderzoeker die de experimenten opzette, vermaakte de hoogleraren (CESR
De experimenten die de onderzoeker opzette, vermaakten de hoogleraar (CEOR)

De Indianen schuwden de cowboy die de dorpen verwoestte (RESR)
De Indiaan schuwde de dorpen die de cowboy verwoestte (REOR)
De cowboy die de dorpen verwoestte, ontmoedigde de Indianen (CESR)
De dorpen die de cowboy verwoestte, ontmoedigden de Indiaan (CEOR)

De scherpschutters beschoten de strijder die de bunkers passeerde (RESR)
De scherpschutter beschoot de bunkers die de strijder passeerde (REOR)
De strijder die de bunkers passeerde, ontgoochelde de scherpschutters (CESR)
De bunkers die de strijder passeerde, ontgoochelden de scherpschutter (CEOR)

De bakker minachtte de gezellen die het brood verkochten (RESR)
De bakkers minachtten het brood dat de gezellen verkochten (REOR)
De gezellen die het brood verkochten, bedroefden de bakker (CESR)
Het brood dat de gezellen verkochten, bedroefde de bakkers (CEOR)

De barones meed de schilders die de kamer witten (RESR)
De baronessen meden de kamer die de schilders witten (REOR)
De schilders die de kamer witten, deprimeerden de barones (CESR)
De kamer die de schilders witten, deprimeerde de baronessen (CEOR)

De impresario verachtte de filmsterren die de rol weigerden (RESR)
De impresario's verachtten de rol die de filmsterren weigerden (REOR)
De filmsterren die de rol weigerden, mishaagden de impresario (CESR)
De rol die de filmsterren weigerden, mishaagde de impresario's (CEOR)

De generaals zonden de piloot die de vliegtuigen bestuurde (RESR)
De generaal zond de vliegtuigen die de piloot bestuurde (REOR)
De piloot die de vliegtuigen bestuurde, alarmeerde de generaals (CESR)
De vliegtuigen die de piloot bestuurde, alarmeerden de generaal (CEOR)

De kandidaten bestudeerden de conrector die de examens binnenbracht (RESR)
De kandidaat bestudeerde het examens dat de conrector binnenbracht (REOR)
De conrector die de examens binnenbracht, demotiveerde de kandidaten (CESR)
De examens die de conrector binnenbracht, demotiveerden de kandidaat (CEOR)

De optometristen corrigeerden de opticien die de monturen verboog (RESR)
De optometrist corrigeerde de monturen die de opticien verboog (REOR)
De opticien die de monturen verboog, irriteerde de optometristen (CESR)
De monturen die de opticien verboog, irriteerden de optometrist (CEOR)

De didacticus evalueerde de docenten die de cursus bedachten (RESR)
De didactici evalueerden de cursus die de docenten bedachten (REOR)
De docenten die de cursus bedachten, interesseerden de didacticus (CESR)
De cursus die de docenten bedachten, interesseerde de didactici (CEOR)

De regisseur vervloekte de spelers die het toneelstuk afraffelden (RESR)
De regisseurs vervloekten het toneelstuk dat de spelers afraffelden (REOR)
De spelers die het toneelstuk afraffelden, ergerden de regisseur (CESR)
Het toneelstuk dat de spelers afraffelden, ergerde de regisseurs (CEOR)

De amanuensis ontdekte de biologen die de rat ontleedden (RESR)
De amanuensissen ontdekten de rat die de biologen ontleedden (REOR)
De biologen die de rat ontleedden, ontstelden de amanuensis (CESR)
De rat die de biologen ontleedden, ontstelde de amanuensissen (CEOR)

De stedelingen verwensten de toerist die de hapjes bestelde (RESR)
De stedeling verwenste de hapjes die de toerist bestelde (REOR)
De toerist die de hapjes bestelde, behaagde de stedelingen (CESR)
De hapjes die de toerist bestelde, behaagden de stedeling (CEOR)

De zakenlieden begrepen de stewardess die de bladen uitreikte (RESR)
De zakenman begreep de bladen die de stewardess uitreikte (REOR)
De stewardess die de bladen uitreikte, onderhield de zakenlui (CESR)
De bladen die de stewardess uitreikte, onderhielden de zakenman (CEOR)

De redacteuren verdedigden de rebel die de aanslagen pleegde (RESR)
De redacteur verdedigde de aanslagen die de rebel pleegde (REOR)
De rebel die de aanslagen pleegde, verbitterde de redacteurs (CESR)
De aanslagen die de rebel pleegde, verbitterden de redacteur (CEOR)

De leerling bejubelde de recensenten die de film afkraakten (RESR)
De leerlingen bejubelden de film die de recensenten afkraakten (REOR)
De recensenten die de film afkraakten, ontroerden de leerling (CESR)
De film die de recensenten afkraakten, ontroerde de leerlingen (CEOR)

De dorpeling begeerde de boerinnen die de vrucht plukten (RESR)
De dorpelingen begeerden de vrucht die de boerinnen plukten (REOR)
De boerinnen die de vrucht plukten, bekoorden de dorpeling (CESR)
De vrucht die de boerinnen plukten, bekoorde de dorpelingen (CEOR)

De huisgenoot fotografeerde de corpsballen die het krat leegdronken (RESR)
De huisgenoten fotografeerden het krat dat de corpsballen leegdronken (REOR)
De corpsballen die het krat leegdronken, shockeerden de huisgenoot (CESR)
Het krat dat de corpsballen leegdronken, shockeerde de huisgenoten (CEOR)

De voorbijgangers roken de vuilnisman die de biobakken leegde (RESR)
De voorbijganger rook de biobakken die de vuilnisman leegde (REOR)
De vuilnisman die de biobakken leegde, belemmerde de voorbijgangers (CESR)
De biobakken die de vuilnisman leegde, belemmerden de voorbijganger (CEOR)

De voorzitters wantrouwden de burgemeester die de declaraties indiende (RESR)
De voorzitter wantrouwde de declaraties die de burgemeester indiende (REOR)
De burgemeester die de declaraties indiende, frustreerde de voorzitters (CESR)
De declaraties die de burgemeester indiende, frustreerden de voorzitter (CEOR

De historici citeerden de nazi die de boeken verbrandde (RESR)
De historicus citeerde de boeken die de nazi verbrandde (REOR)
De nazi die de boeken verbrandde, verwarde de historici (CESR)
De boeken die de nazi verbrandde, verwarden de historicus (CEOR)

De kenner prees de koks die het deeg kneedden (RESR)
De kenners prezen het deeg dat de koks kneedden (REOR)
De koks die het deeg kneedden, verrukten de kenner (CESR)
Het deeg dat de koks kneedden, verrukte de kenners (CEOR)

De kwajongen verfoeide de meesters die het strafwerk uitdeelden (RESR)
De kwajongens verfoeiden het strafwerk dat de meesters uitdeelden (REOR)
De meesters die het strafwerk uitdeelden, kwelden de kwajongen (CESR)
Het strafwerk dat de meesters uitdeelden, kwelde de kwajongens (CEOR)

De minister roemde de accountants die de offerte opstelden (RESR)
De ministers roemden de offerte die de accountants opstelden (REOR)
De accountants die de offerte opstelden, bevielen de minister (CESR)
De offerte die de accountants opstelden, beviel de ministers (CEOR)

De rectoren ontvingen de scholier die de roosters ophaalde (RESR)
De rector ontving de roosters die de scholier ophaalde (REOR)
De scholier die de roosters ophaalde, belastte de rectoren (CESR)
De roosters die de scholier ophaalde, belastten de rector (CEOR)

De verslaafden misten de dokter die de recepten uitschreef (RESR)
De verslaafde miste de recepten die de dokter uitschreef (REOR)
De dokter die de recepten uitschreef, verheugde de verslaafden (CESR)
De recepten die de dokter uitschreef, verheugden de verslaafde (CEOR)

De wetenschappers onderschatten de statisticus die de problemen oploste (RESR)
De wetenschapper onderschatte de problemen die de statisticus oploste (REOR)
De statisticus die de problemen oploste, verveelde de wetenschappers (CESR)
De problemen die de statisticus oploste, verveelden de wetenschapper (CEOR)

De kanonnier bestookte de soldaten die de vliegbasis aanvielen (RESR)
De kanonniers bestookten de vliegbasis die de soldaten aanvielen (REOR)
De soldaten die de vliegbasis aanvielen, imponeerden de kanonniers (CESR)
De vliegbasis die de soldaten aanvielen, imponeerde de kanonniers (CEOR)

De kunstenaar schilderde de meisjes die de koek serveerden (RESR)
De kunstenaars schilderden de koek die de meisjes serveerden (REOR)
De meisjes die de koek serveerden, plezierden de kunstenaar (CESR)
De koek die de meisjes serveerden, plezierde de kunstenaars (CEOR)

De professor behandelde de dichters die het pamflet schreven (RESR)
De professoren behandelden het pamflet dat de dichters schreven (REOR)
De dichters die het pamflet schreven, beledigden de professor (CESR)
Het pamflet dat de dichters schreven, beledigde de professoren (CEOR)

De premiers verguisden de boef die de moorden beraamde (RESR)
De premier verguisde de moorden die de boef beraamde (REOR)
De boef die de moorden beraamde, beangstigde de premier (CESR)'
De moorden die de boef beraamde, beangstigden de premier (CEOR)

De vakkenvullers vergoeilijkten de caissière die de kasverschillen veroorzaakte (RESR)
De vakkenvuller vergoeilijkte de kasverschillen die de caissière veroorzaakte (REOR)
De caissière die de kasverschillen veroorzaakte, verbaasde de vakkenvuller (CESR)
De kasverschillen die de caissière veroorzaakte, verbaasden de vakkenvuller (CEOR)'

De handelaars verlieten de boer die de stallen uitmestte (RESR)
De handelaar verliet de stallen die de boer uitmestte (REOR)
De boer die de stallen uitmestte, benauwde de handelaren (CESR)
De stallen die de boer uitmestte, benauwden de handelaar (CEOR)

Materials: Experiment 5

De barkeeper heeft de chips opgegeten terwijl de verregende (hongerige trouwe) klant niets doorhad (RE1A/RE3A)
De barkeeper heeft, terwijl de verregende (hongerige trouwe) klant niets doorhad, de chips opgegeten (CE1A/CE3A)

De professor heeft de brief aangereikt toen de charmante (lange blonde) secretaresse terloops opkeek (RE1A/RE3A)
De professor heeft, toen de charmante (lange blonde) secretaresse terloops opkeek, de brief aangereikt (CE1A/CE3A)

De journalist heeft de overval beschreven hoewel de ontsnapte (gestoorde wrede) misdadiger hem bedreigde (RE1A/RE3A)
De journalist heeft, hoewel de ontsnapte (gestoorde wrede) misdadiger hem bedreigde, de overval beschreven (CE1A/CE3A)

De musicus heeft de sonate vertolkt voordat de statige (wijze deftige) koningin iedereen toesprak (RE1A/RE3A)
De musicus heeft, voordat de statige (wijze deftige) koningin iedereen toesprak, de sonate vertolkt (CE1A/CE3A)

De familie heeft de begrafenis georganiseerd nadat de zieke (demente oude) grootvader rustig overleed (RE1A/RE3A)
De familie heeft, nadat de zieke (demente oude) grootvader rustig overleed, de begrafenis georganiseerd (CE1A/CE3A)

De revolutionair heeft een aanslag beraamd omdat de strenge (paranoïde sadistische) dictator morgen trouwt (RE1A/RE3A)
De revolutionair heeft, omdat de strenge (paranoïde sadistische) dictator morgen trouwt, een aanslag beraamd (CE1A/CE3A)

De chimpansee heeft zijn kooi verlaten toen de zenuwachtige (nieuwe jonge) verzorger onvoldoende oplette (RE1A/RE3A)
De chimpansee heeft, toen de zenuwachtige (nieuwe jonge) verzorger onvoldoende oplette, zijn kooi verlaten (CE1A/CE3A)

De technicus heeft het gereedschap neergegooid nadat zijn chagrijnige (onattente dominante) collega luidkeels foeterde (RE1A/RE3A)
De technicus heeft, nadat zijn chagrijnige (onattente dominante) collega luidkeels foeterde, het gereedschap neergegooid (CE1A/CE3A)

De timmerman heeft het dekzeil gespannen terwijl de koude (striemende gure) regen keihard neerkletterde (RE1A/RE3A)
De timmerman heeft, terwijl de koude (striemende gure) regen keihard neerkletterde, het dekzeil gespannen (CE1A/CE3A)

De prostituee heeft de deur geopend toen de onzekere (slungelige zwetende) puber steeds langsliep (RE1A/RE3A)
De prostituee heeft, toen de onzekere (slungelige zwetende) puber steeds langsliep, de deur geopend (CE1A/CE3A)

De bruidegom heeft de taart aangesneden voordat zijn ongemanierde (arrogante gierige) schoonvader tenslotte arriveerde (RE1A/RE3A)
De bruidegom heeft, voordat zijn ongemanierde (arrogante gierige) schoonvader tenslotte arriveerde, de taart aangesneden (CE1A/CE3A)

De bezoeker heeft een roos aangeboden toen de aantrekkelijke (lenige slanke) danseres sierlijk boog (RE1A/RE3A)
De bezoeker heeft, toen de aantrekkelijke (lenige slanke) danseres sierlijk boog, een roos aangeboden (CE1A/CE3A)

De predikant heeft een preek gehouden omdat de dronken (luidruchtige agressieve) jongelui luid boerden (RE1A/RE3A)
De predikant heeft, omdat de dronken (luidruchtige agressieve) jongelui luid boerden, een preek gehouden (CE1A/CE3A)

De prinses heeft een wijntje gedronken voordat de beleefde (bescheiden hulpvaardige) lakei alles opruimde (RE1A/RE3A)
De prinses heeft, voordat de beleefde (bescheiden hulpvaardige) lakei alles opruimde, een wijntje gedronken (CE1A/CE3A)

De terrorist heeft de explosieven geplaatst voordat het grote (moderne Israëlische) vliegtuig 's morgens vertrok (RE1A/RE3A)
De terrorist heeft, voordat het grote (moderne Israëlische) vliegtuig 's morgens vertrok, de explosieven geplaatst (CE1A/CE3A)

De postbode heeft het hek gesloten toen de enorme (grommende wilde) rottweiler blaffend aansnelde (RE1A/RE3A)
De postbode heeft, toen de enorme (grommende wilde) rottweiler blaffend aansnelde, het hek gesloten (CE1A/CE3A)

De pony heeft het hooi geproefd nadat de lieve (bedaarde zorgzame) oppasser stilletjes aaide (RE1A/RE3A)
De pony heeft, nadat de lieve (bedaarde zorgzame) oppasser stilletjes aaide, het hooi geproefd (CE1A/CE3A)

De piraat heeft de aanval ingezet toen het kolossale (Portugese houten) schip langzaam naderde (RE1A/RE3A)
De piraat heeft, toen het kolossale (Portugese houten) schip langzaam naderde, de aanval ingezet (CE1A/CE3A)

De wandelaar heeft zijn route gewijzigd omdat het kronkelige (smalle steile) pad onbegaanbaar werd (RE1A/RE3A)
De wandelaar heeft, omdat het kronkelige (smalle steile) pad onbegaanbaar werd, zijn route gewijzigd (CE1A/CE3A)

De pelgrim heeft zijn mantel uitgedaan toen de vervelende (moeilijke barre) tocht eindelijk afliep (RE1A/RE3A)
De pelgrim heeft, toen de vervelende (moeilijke barre) tocht eindelijk afliep, zijn mantel uitgedaan (CE1A/CE3A)

De dief heeft de grendel geforceerd terwijl zijn luie (gemene onberekenbare) handlanger schichtig rondkeek (RE1A/RE3A)
De dief heeft, terwijl zijn luie (gemene onberekenbare) handlanger schichtig rondkeek, de grendel geforceerd (CE1A/CE3A)

De dolfijn heeft de diepte opgezocht toen de kille (zwarte kolkende) oceaan wild golfde (RE1A/RE3A)
De dolfijn heeft, toen de kille (zwarte kolkende) oceaan wild golfde, de diepte opgezocht (CE1A/CE3A)

De weduwe heeft de lever klaargemaakt terwijl de kleine (bruine keffende) pekinees hongerig wachtte (RE1A/RE3A)
De weduwe heeft, terwijl de kleine (bruine keffende) pekinees hongerig wachtte, de lever klaargemaakt (CE1A/CE3A)

De monteur heeft zijn neus gesnoten toen de bebrilde (dikke kale) miljonair minachtend keek (RE1A/RE3A)
De monteur heeft, toen de bebrilde (dikke kale) miljonair minachtend keek, zijn neus gesnoten (CE1A/CE3A)

De wesp heeft het jongetje gestoken toen de fruitige (zoete rode) jam helemaal opging (RE1A/RE3A)
De wesp heeft, toen de fruitige (zoete rode) jam helemaal opging, het jongetje gestoken (CE1A/CE3A)

Het personeel heeft een bonus ontvangen toen het dynamische (vlotte grensverleggende) bedrijf aardig uitbreidde (RE1A/RE3A)
Het personeel heeft, toen het dynamische (vlotte grensverleggende) bedrijf aardig uitbreidde, een bonus ontvangen (CE1A/CE3A)

De pianist heeft de partituur dichtgeslagen nadat de snelle (ingewikkelde barokke) étude klaar was (RE1A/RE3A)
De pianist heeft, nadat de snelle (ingewikkelde barokke) étude klaar was, de partituur dichtgeslagen (CE1A/CE3A)

De pionier heeft zijn ogen afgedekt toen het gelige (fijne scherpe) woestijnzand woest stoof (RE1A/RE3A)
De pionier heeft, toen het gelige (fijne scherpe) woestijnzand woest stoof, zijn ogen afgedekt (CE1A/CE3A)

De rover heeft de postkoets tegengehouden voordat de dravende (rossige dampende) paarden voorbij waren (RE1A/RE3A)
De rover heeft, voordat de dravende (rossige dampende) paarden voorbij waren, de postkoets tegengehouden (CE1A/CE3A)

De scheidsrechter heeft de wedstrijd stopgezet nadat de woedende (fanatieke racistische) supporter vervaarlijk schreeuwde (RE1A/RE3A)
De scheidsrechter heeft, nadat de woedende (fanatieke racistische) supporter vervaarlijk schreeuwde, de wedstrijd stopgezet (CE1A/CE3A)

De veldwachter heeft de aanhouding verricht toen de stinkende (laveloze arme) zwerver 's nachts rondliep (RE1A/RE3A)
De veldwachter heeft, toen de stinkende (laveloze arme) zwerver 's nachts rondliep, de aanhouding verricht (CE1A/CE3A)

De zakenman heeft zijn aandelen verkocht voordat de onverwachte (zware wereldwijde) beurskrach zich aandiende (RE1A/RE3A)
De zakenman heeft, voordat de onverwachte (zware wereldwijde) beurskrach zich aandiende, zijn aandelen verkocht (CE1A/CE3A)

De cowboy heeft zijn viool gestemd toen de vrolijke (feestelijke muzikale) avond bijna begon (RE1A/RE3A)
De cowboy heeft, toen de vrolijke (feestelijke muzikale) avond bijna begon, zijn viool gestemd (CE1A/CE3A)

De voorman heeft zijn gal gespuid omdat de ziekelijke (ongemotiveerde slappe) ploeg vaak pauzeerde (RE1A/RE3A)
De voorman heeft, omdat de ziekelijke (ongemotiveerde slappe) ploeg vaak pauzeerde, zijn gal gespuid (CE1A/CE3A)

De aannemer heeft zijn biezen gepakt toen de bouwvallige (tochtige vervallen) woning schrikbarend kraakte (RE1A/RE3A)
De aannemer heeft, toen de bouwvallige (tochtige vervallen) woning schrikbarend kraakte, zijn biezen gepakt (CE1A/CE3A)

De dirigent heeft de concertmeester begroet voordat het beroemde (gedisciplineerde Duitse) orkest daverend inzette (RE1A/RE3A)
De dirigent heeft, voordat het beroemde (gedisciplineerde Duitse) orkest daverend inzette, de concertmeester begroet (CE1A/CE3A)

De diplomaat heeft zijn verontwaardiging uitgesproken toen de kwaadaardige (bloeddorstige Servische) troepen dreigend optrokken (RE1A/RE3A)
De diplomaat heeft, toen de kwaadaardige (bloeddorstige Servische) troepen dreigend optrokken, zijn verontwaardiging uitgesproken (CE1A/CE3A)

De marinier heeft het dynamiet aangebracht terwijl de dreigende (glanzende metalen) kanonnen voortdurend vuurden (RE1A/RE3A)
De marinier heeft, terwijl de dreigende (glanzende metalen) kanonnen voortdurend vuurden, het dynamiet aangebracht (CE1A/CE3A)

De wielrenner heeft zijn voorsprong behouden hoewel het fanatieke (getrainde ervaren) peloton enigszins inliep (RE1A/RE3A)
De wielrenner heeft, hoewel het fanatieke (getrainde ervaren) peloton enigszins inliep, zijn voorsprong behouden (CE1A/CE3A)

De maîtresse heeft de slaapkamer betreden hoewel de achterdochtige (jaloerse hartvochtige) keizerin argwanend toekeek (RE1A/RE3A)
De maîtresse heeft, hoewel de achterdochtige (jaloerse hartvochtige) keizerin argwanend toekeek, de slaapkamer betreden (CE1A/CE3A)

De chirurg heeft de operatie uitgevoerd toen de aanstellerige (veeleisende verwende) patiënt nauwelijks opknapte (RE1A/RE3A)
De chirurg heeft, toen de aanstellerige (veeleisende verwende) patiënt nauwelijks opknapte, de operatie uitgevoerd (CE1A/CE3A)

De nieuwkomer heeft zijn baan opgezegd omdat de harteloze (egoïstische verwaande) werknemers nooit meewerkten (RE1A/RE3A)
De nieuwkomer heeft, omdat de harteloze (egoïstische verwaande) werknemers nooit meewerkten, zijn baan opgezegd (CE1A/CE3A)

De profeet heeft zijn toespraak beëindigd omdat de ongelovige (minachtende vijandige) menigte niet luisterde (RE1A/RE3A)
De profeet heeft, omdat de ongelovige (minachtende vijandige) menigte niet luisterde, zijn toespraak beëindigd (CE1A/CE3A)

De stadhouder heeft de stadsmuren versterkt omdat de plunderende (brandschattende Spaanse) legers dichterbij kwamen (RE1A/RE3A)
De stadhouder heeft, omdat de plunderende (brandschattende Spaanse) legers dichterbij kwamen, de stadsmuren versterkt (CE1A/CE3A)

De president heeft zijn ontslag ingediend omdat de omstreden (Oostenrijkse fascistische) partij overtuigend won (RE1A/RE3A)
De president heeft, omdat de omstreden (Oostenrijkse fascistische) partij overtuigend won, zijn ontslag ingediend (CE1A/CE3A)

De schildwacht heeft het zoeklicht aangezet toen de lugubere (duistere winterse) nacht geleidelijk inviel (RE1A/RE3A)
De schildwacht heeft, toen de lugubere (duistere winterse) nacht geleidelijk inviel, het zoeklicht aangezet (CE1A/CE3A)

De schipper heeft zijn reis afgebroken omdat de krakkemikkige (ouderwetse stalen) sluizen slecht functioneerden (RE1A/RE3A)
De schipper heeft, omdat de krakkemikkige (ouderwetse stalen) sluizen slecht functioneerden, zijn reis afgebroken (CE1A/CE3A)

De restaurateur heeft het ijzer schoongemaakt omdat de smerige (antieke verweerde) sloten nogal roestten (RE1A/RE3A)
De restaurateur heeft, omdat de smerige (antieke verweerde) sloten nogal roestten, het ijzer schoongemaakt (CE1A/CE3A)

Span Data

Span Data Normally Distributed

In this thesis, Pearson's correlation coefficients were used. One important assumption that underlies this type of correlation is that the memory span data are normally distributed. In Figures 1 to 4, QQ of the memory span data are given. The fact that the data points fit the solid lines quite well shows that this assumption is legitimate for the span data at hand.

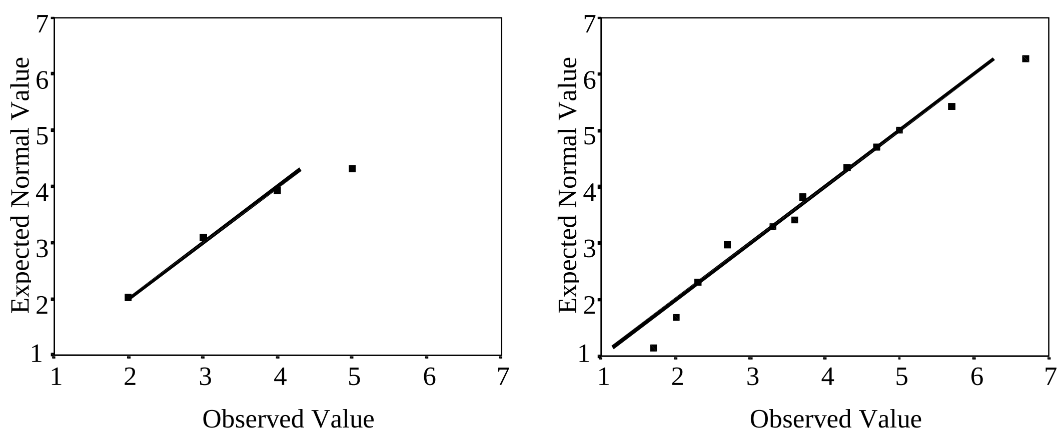


Figure 1. QQ Plots for the nonword (left) and Salthouse Listening Span data (right) in Experiment 1.

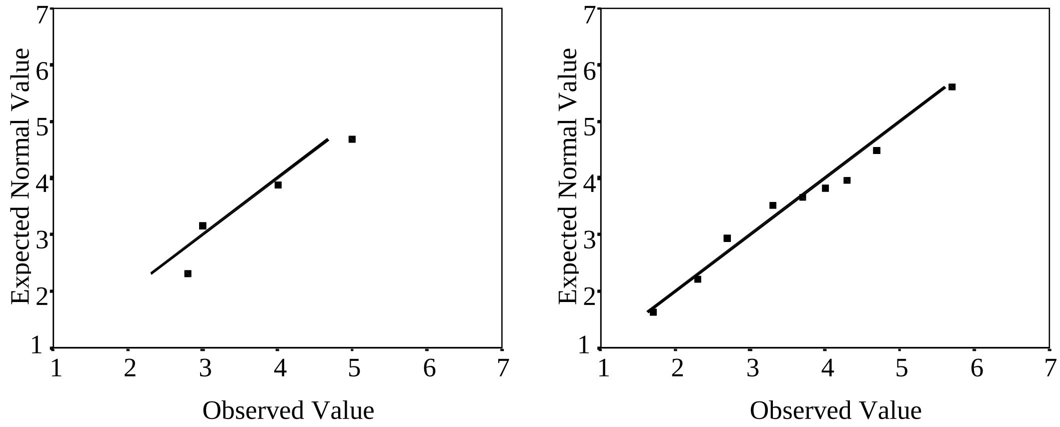


Figure 2. QQ Plots for the nonword (left) and Salthouse Listening Span data (right) in Experiment 2.

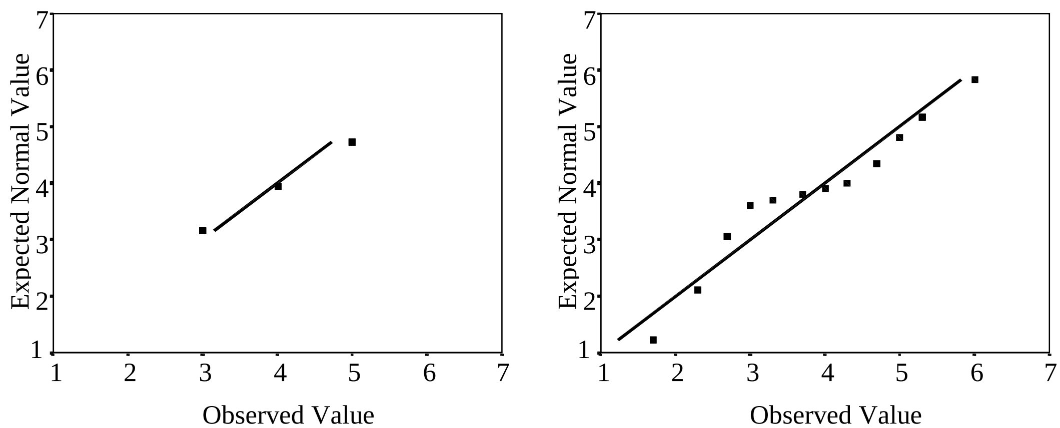


Figure 3. QQ Plots for the nonword (left) and Salthouse Listening Span data (right) in Experiment 2,4, and 5b.

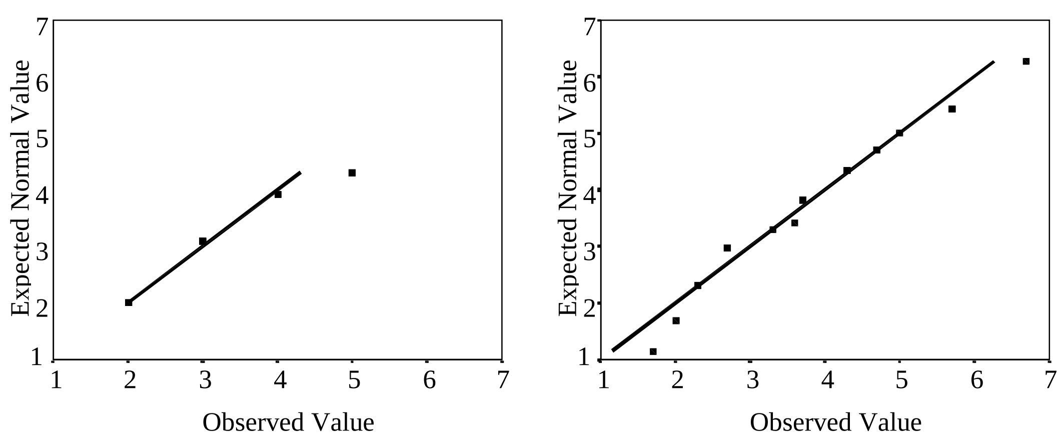


Figure 4. QQ Plots for the nonword (left) and Salthouse Listening Span data (right) in Experiment 5a.

Correlations with Nonword Span

Correlations with Effects in the RT Analysis. In the RT analysis of Experiment 1, there was a significant negative correlation of nonword span and performance under the condition of Articulatory Suppression (r = -.40, $p < .05$). Figure 5 shows the mean effect size of Articulatory Suppression in ms by nonword span.

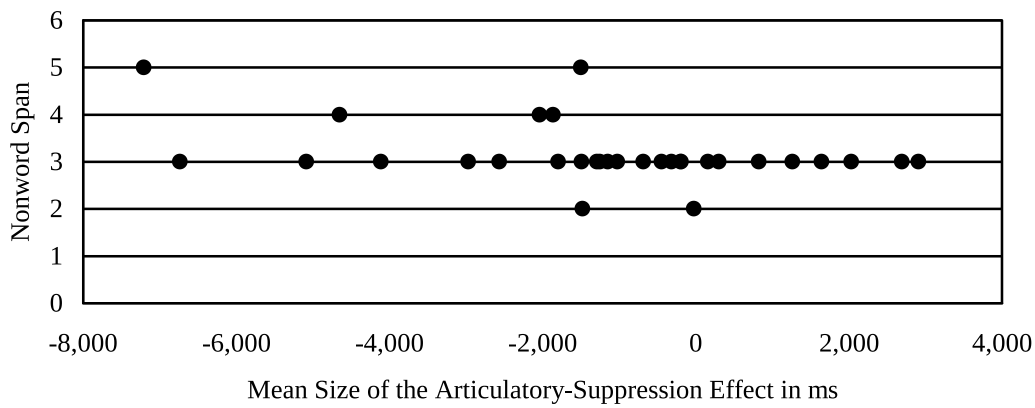


Figure 5. Mean effect size of Articulatory Suppression in ms by nonword span in Experiment 1.

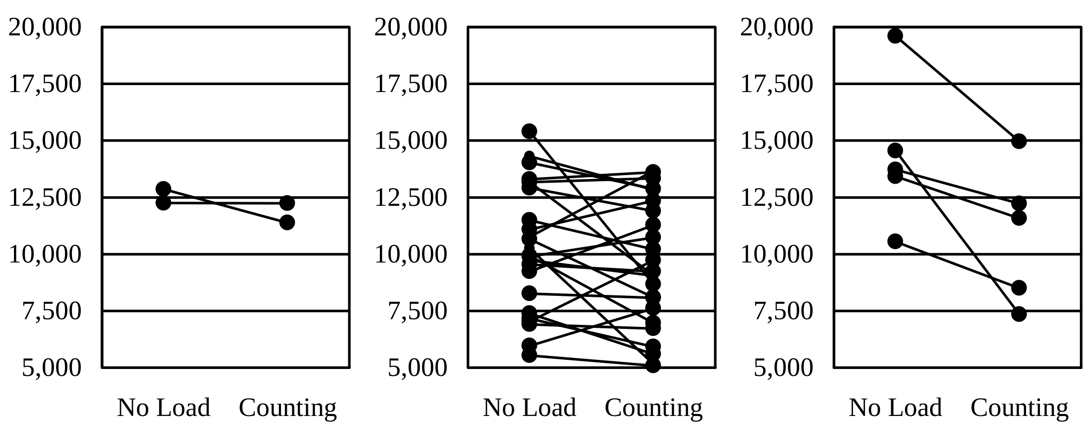


Figure 6. Mean RTs in ms for the no load and counting conditions for nonword low (n = 2, span = 2 [left]), medium (n = 25, span = 3 [middle]), and high spanners (n = 5, span = 4 or 5 [right]) in Experiment 1.

It is hard to interpret the nature of the effect size just by looking at the scatterplot in Figure 5 because the data points are subtractions of the mean RTs to the counting conditions and those of the no-load conditions. One cannot tell whether a negative effect size should be attributed to low RTs to the counting conditions or to high RTs to the no-load condition.

The correlation of nonword span and Articulatory Suppression suggests that participants with high nonword spans read more quickly under articulatory suppression than in the no-load condition, but that low spanners did not show such a tendency.

Figure 6 gives an overview of the mean response latencies of the no-load and counting conditions per subject. Participants were divided in three span groups so that one can see how the different span sizes behaved. Participants with span size 2 were included in the low-span group ($n = 2$). Those with span size 3 were included in the middle-span group ($n = 25$) and those with span sizes 4 and 5 were included in the high-span group ($n = 5$).

Correlations with Effects in the Error Analysis. The error analysis of Experiment 3 showed a significant correlation of nonword span and the two-way interaction of Clause Type x Propositional Complexity ($r = .38$, $p < .05$). It turned out that the nonword span correlated significantly with Clause Type in two-proposition sentences ($r = .37$, $p < .05$), but not in one-proposition sentences ($r = -.09$, $p > .1$).

Subjects were assigned to span groups: those with memory spans of 3 or lower were included in the low-span group ($n = 14$); those with higher spans were included in the high-span group ($n = 18$). The correlation of nonword span with the simple effect of Clause Type in the two-proposition sentences is shown in Figure 7.

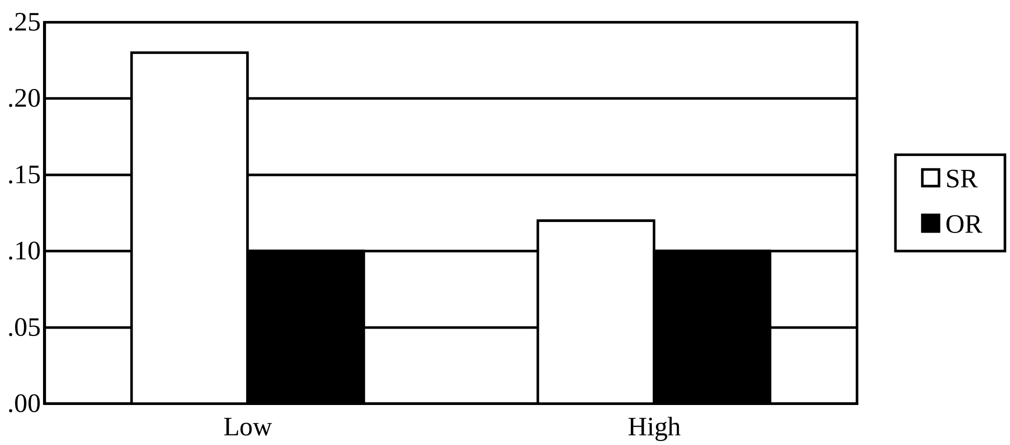


Figure 7. Mean error proportions for subject (SR) and object relatives (OR) of the two-proposition sentences for nonword low (span ≤ 3, $n = 14$ [left]) and high spanners (span ≥ 4, $n = 18$ [right]) in Experiment 3.

As can be seen in this figure, low span subjects were worse reading object-relative two-proposition sentences than subject-relative two-action sentences, but high spanners showed virtually no effect.

Second, Experiment 5a produced a significant correlation of nonword span and the two-way interaction of Embedding Type x Articulatory Suppression in the error rates (r = -.37, p < .05). There were no significant correlations with nonword span in the center-embedded conditions (r = .15, p > .1), but the right-embedded conditions yielded a marginally significant correlation of nonword span with Articulatory Suppression (r = -.35, p = .057).

Figure 8 shows the mean RTs for the right-embedded no-load and counting conditions for each participant; participants are divided in three span groups: low (span = 2, n = 2), medium (span = 3, n = 25), and high (span ≥ 4, n = 5). There is general tendency for participants to make more errors in the counting condition than in the no-load condition of the right-embedded sentences only. In the middle- and high-span group however, some individuals show an opposite tendency.

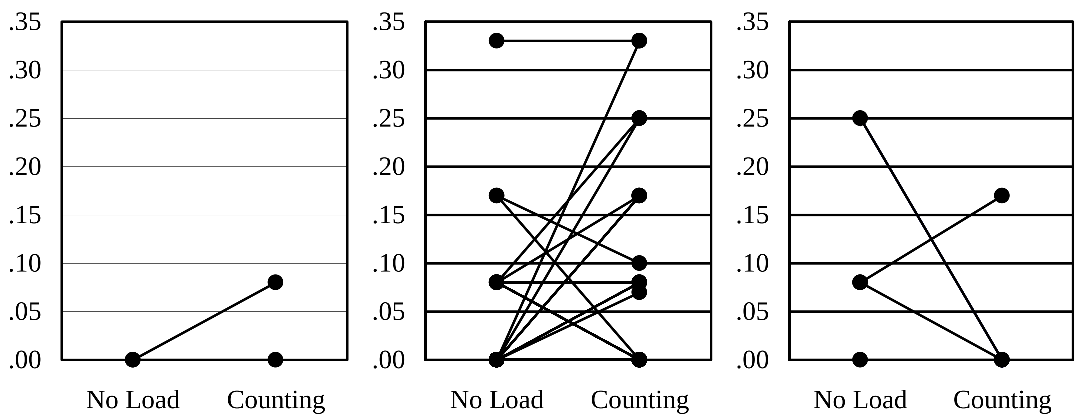


Figure 8. Mean RTs in ms for the right-embedded no load and counting conditions for nonword low (n = 2, span = 2 [left]), medium (n = 25, span = 3 [middle]), and high spanners (n = 5, span = 4 or 5 [right]) in Experiment 5a.

Correlations with Salthouse Listening Span

Correlations with Effects in the RT Analysis. Experiment 2 produced a significant correlation of Salthouse listening span and performance under the condition of Articulatory Suppression (r = .41, p < .05) in the RT analysis.

In order to get more insight in the exact nature of the span correlation, participants were divided over two span groups, low (n = 17) and high spanners (n = 15). The low-span group had memory spans of 3 or lower; the high spanners had spans of 3.3 or higher. Figure 9 shows the mean RTs for the no-load and counting conditions for the two span groups. The correlation reflects the fact that participants with low listening spans read more quickly in the articulatory-suppression condition, but that high-span readers did not show this effect.

Second, the Salthouse listening span produced a significant correlation with the two-way interaction of Adjective x Embedding Type in the RT analysis of Experiment 5b ($r = .35$, $p < .05$). There were no simple effects of Embedding Type, neither in the one- ($r = -.26$, $p > .1$) nor in the three-adjective sentences ($r = .13$, $p > .1$). Therefore, it is not possible to explain this correlation.

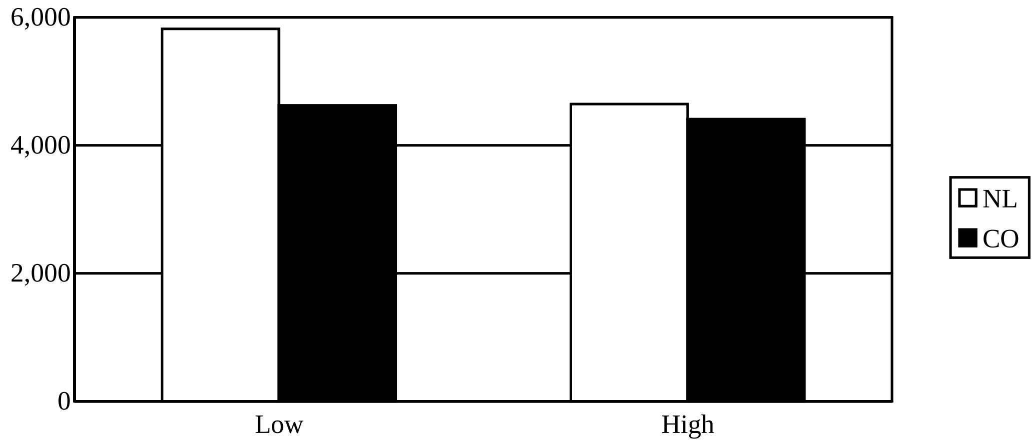


Figure 9. Mean RTs in ms for the no load (NL) and counting (CO) conditions for Salthouse low (span ≤ 3, $n = 17$ [left]) and high spanners (span ≥ 4, $n = 15$ [right]) in Experiment 2.

Third, Experiment 4 produced a significant correlation of Salthouse listening span with the three-way interaction of Clause Type x Embedding Type x Articulatory Suppression in the RT analysis ($r = -.44$, $p < .05$). In the right-embedded sentences, there was no significant correlation with span ($r = .125$, $p > .1$), but in the center-embedded sentences, there was a significant correlation with the two-way interaction of Clause Type x Articulatory Suppression (Pearson = -.45, $p < .05$).

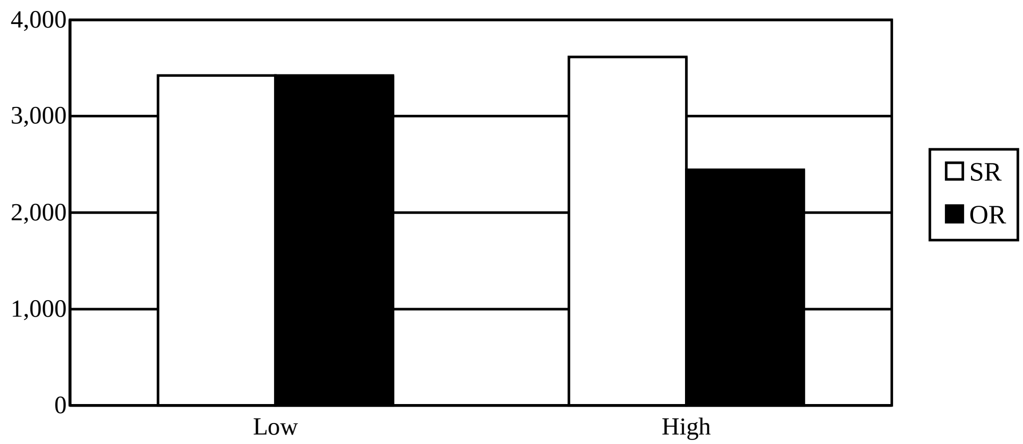


Figure 10. Mean RTs in ms for the subject (SR) and object relatives (OR) in the articulatory-suppression condition of the center-embedded sentences for Salthouse low (span ≤ 3, *n* = 19 [left]) and high spanners (span ≥ 4, *n* = 13 [right]) in Experiment 4.

It turned out that there were no significant correlations with Clause Type in the no-load condition of the center-embedded sentences, but that there was in the articulatory-suppression condition ($r = -.37$, $p < .05$). Subjects were divided over two span groups: subjects with listening spans of 3 or lower were considered low spanners ($n = 19$); the other participants were included in the high-span group ($n = 13$). The correlation with simple effect of Clause Type in the articulatory-suppression condition of the center-embedded sentences is shown in Figure 10.
This figure demonstrates that high spanners were faster in the articulatory-suppression condition of the center-embedded constructions reading object- than subject-relative clauses. This tendency was not present in the low spanners, who did not show any difference between the two conditions.

Correlations with Effects in the Error Analysis. The error analysis of Experiment 5b showed a significant correlation of the Salthouse listening span with Adjectival Load ($r = .38$, $p < .05$). In order to get more insight into the nature of this correlation, span scores were divided over two groups. Participants with a memory span of 3 or lower were assigned to the low-span group ($n = 19$); the others were in the high-span group ($n = 13$). The effect sizes of the one and three-adjective conditions are shown in Figure 11. Participants with low memory spans were more accurate reading one- than three-adjective sentences; high spanners showed the opposite effect.

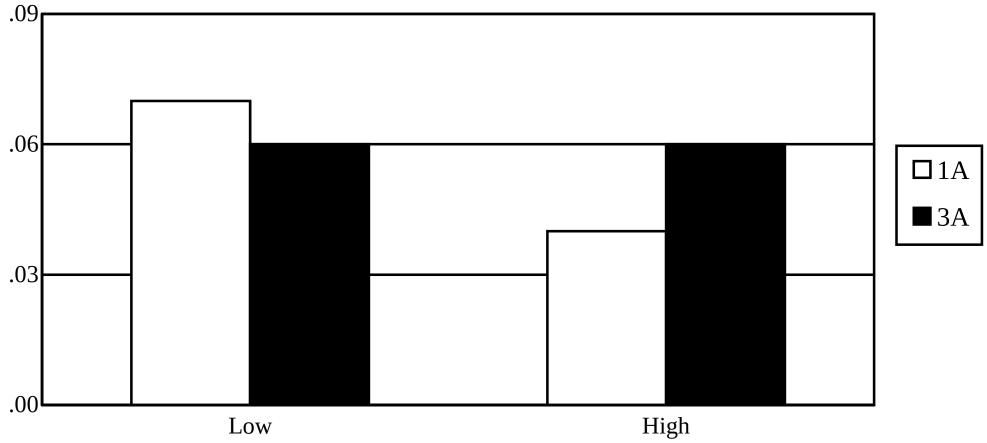


Figure 11. Mean error proportions for the one- (1A) and three-adjective (3A) sentences for Salthouse low (span ≤ 3, *n* = 19) and high spanners (span ≥ 4, *n* = 13) in Experiment 5b.

Second, there was a significant correlation of the Salthouse listening span and Articulatory Suppression ($r = -.37$, $p < .05$) in Experiment 5b. In order to study this correlation, participants were divided into low and high spanners (low spanners had memory spans of 3 or lower; the other participants were assigned to the high-span group. The mean RTs to the no-load and counting conditions for each span group are shown in Figure 12. Low-span participants showed higher accuracy in the no-load conditions, but high spanners in the counting condition.

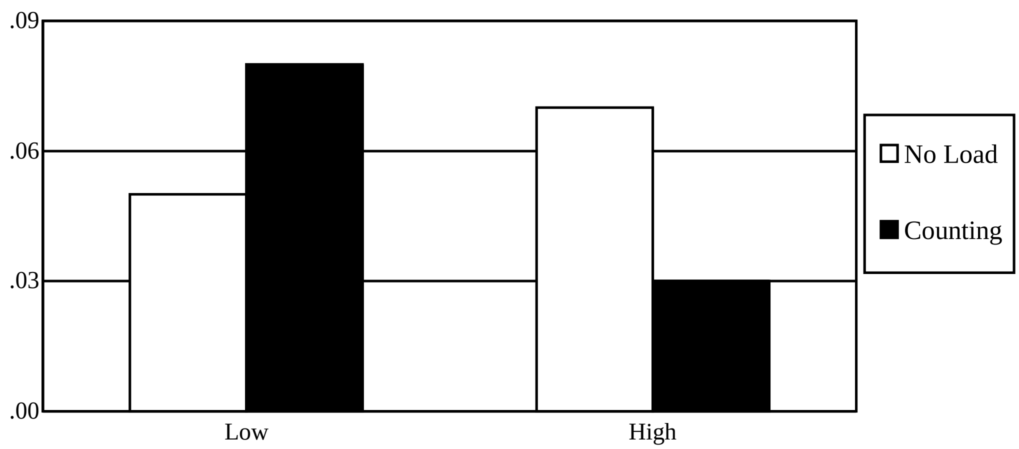


Figure 12. Mean error proportions for the no load and counting conditions for Salthouse low (span ≤ 3, *n* = 19) and high spanners (span ≥ 4, *n* = 13) in Experiment 5b.

Third, in the error rates of Experiment 5a, there was a significant correlation between the Salthouse listening span and the interaction of Embedding Type x Articulatory Suppression ($r = -.39$, $p < .05$). There was no significant correlation of

Salthouse listening span with Articulatory Suppression in the center-embedded conditions (r = .07, p > .1). The right-embedded sentences, however, produced a significant correlation of Articulatory Suppression with Salthouse listening span (r = -.43, p < .05).

For interpretation of this correlation, span data were divided over three span groups: low (span sizes up to 3), middle (spans between 3 and 4), and high (spans of 4 or higher).

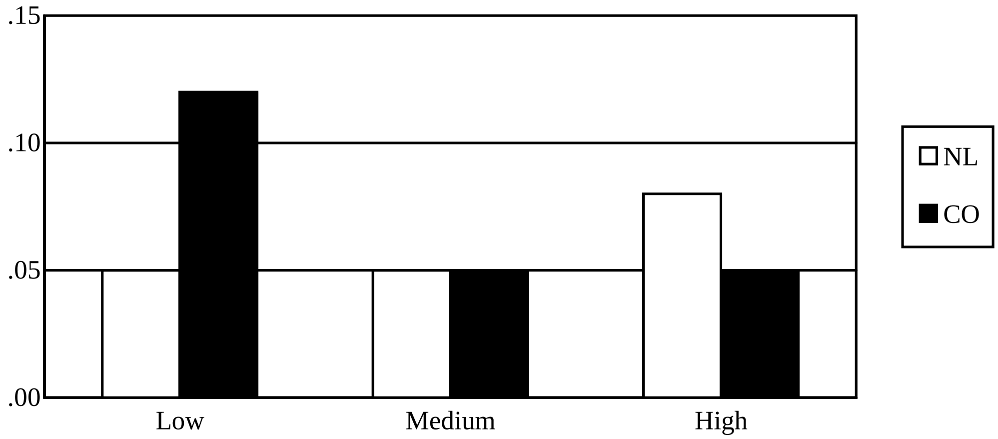


Figure 13. Mean error proportions for the no load (NL) and counting (CO) conditions of the right-embedded sentences by Salthouse listening span group (low [n = 12], medium [n = 15], and high [n = 5]) in Experiment 5a.

Figure 13 shows the mean RTs of the no load and articulatory-suppression conditions of the right-embedded sentences for the three Salthouse listening span groups (low [n = 12], medium [n = 15], and high [n = 5]). In the one-adjective sentences, low-span subjects tend to make more errors under articulatory suppression, whereas medium spanners show no difference and high-span participants show a slight effect in the opposite direction.

Fourth, there was a significant correlation in Experiment 5a of Salthouse listening span with the two-way interaction of Adjectival Load x Embedding Type in the error rates (r = .35, p = .05). No correlations with the Embedding Type effect were found in the three-adjective sentences (r = -.171, p > .1). In the one-adjective sentences, however, there was a marginally significant correlation of Salthouse span with Embedding Type (r = .33, p < .1). This correlation is shown in Figure 14. Subjects were assigned to three groups: low (span < 3, n = 12), medium (3 < span < 4, n = 15), and high (span ≥ 4, n = 5). Low- and medium-span participants tend to have lower error rates for the center-embedded constructions than for right-embedded ones, whereas high spanners show an opposite tendency.

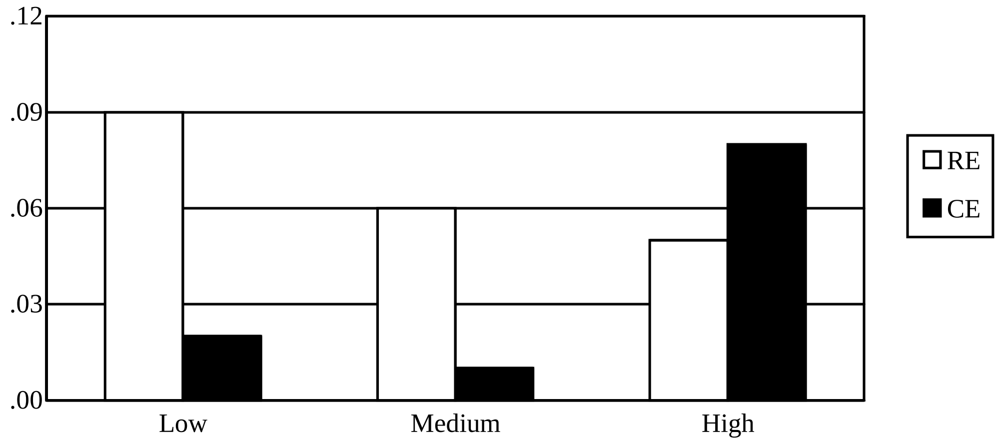


Figure 14. Mean error proportions for the right- (RE) and center-embedded (CE) conditions of the one-adjective sentences by Salthouse listening span group (low [span < 3, n = 12], medium [3 < span < 4, n = 15], and high [span ≥ 4, n = 5]) in Experiment 5a.

Samenvatting

Dit proefschrift gaat over de rol van het werkgeheugen bij het begrijpen van zinnen. Dat werkgeheugen een rol speelt bij zinsbegrip is niet zo verwonderlijk. Woorden kunnen immers niet allemaal tegelijk gehoord of gelezen worden. Als iemand bij het eind van een zin niet meer weet wat het begin ervan was, wordt het begrip van zo'n zin een onmogelijke opgave.

Daarom nemen veel onderzoekers aan dat (één of andere vorm van) werkgeheugen betrokken is bij het begijpen van zinnen. Alhoewel er brede consensus bestaat over die betrokkenheid, zijn veel onderzoekers het er niet over eens hoe zo'n werkgeheugensysteem nu precies in elkaar zit.

In dit proefschrift wordt een model gepresenteerd van de werkgeheugensystemen die betrokken zijn bij zinsbegrip. In dat model staat met name de rol van de *fonologische loop* centraal. Dit is een component van Baddeley's (1986) werkgeheugenmodel. Baddeley's model bestaat uit een centraal controleapparaat, de *central executive,* en twee hulpsystemen, één voor opslag van visuele informatie, het *visuospatial sketchpad,* en één voor klanken, de fonologische loop. De fonologische loop zelf bestaat weer uit een klankgeheugendeel waarin spraak binnenkomt en uit een articulatorische herhaalmodule waarmee de akoestische sporen die in het klankgeheugendeel zijn opgeslagen, kunnen worden herhaald. Dit herhalen voorkomt dat de sporen vergeten worden.

Het model dat gepresenteerd wordt in dit proefschrift is een aanpassing van het werkgeheugenmodel van Waters, Caplan & Hildebrandt (1987), dat ervan uitgaat dat zinsverwerking in twee stadia plaatsvindt.

Hoofdstuk 1 en 2. Het eerste hoofdstuk van het proefschrift schetst het wetenschappelijk kader van de experimenten die later in het boek besproken worden. In het tweede hoofdstuk van dit proefschrift wordt het model van Waters e.a. (1987) beschreven. In het eerste stadium van dat model vindt de syntactische verwerking plaats. Daarna wordt de opgebouwde zinsstructuur overgeheveld naar een tweede stadium, dat van de propositionele analyse.

Waters e.a. (1987) nemen aan dat zowel de syntactische als de propositionele verwerking zich beide in de central executive van Baddeley's (1986) werkgeheugenmodel afspelen. Verder veronderstellen zij dat tijdens de propositionele analyse de fonologische loop een rol speelt. Deze laatste veronderstelling is enigszins problematisch omdat ze theoretisch onvoldoende onderbouwd is en gebaseerd op experimenten waarin propositionele complexiteit en inbedding vervlochten zijn.

Waters e.a. (1987) hebben hun model gebaseerd op de uitkomst van het volgende experiment. Ze hebben de syntactische verwerkingsfase getest in een plausibiliteitsbeoordelingstaak door subject- (zie 1c) en objectrelatieve zinnen (zie 1d) te vergelijken en de propositionele fase door zinnen met één (zie 1a en b) en twee handelingen (zie 1c en d) te contrasteren.

(1) a. It was the thief that broke into the warehouse
    b. It was the broken clock that the jeweller adjusted
    c. The man hit the landlord that requested the money
    d. The meat that the butcher cut delighted the customer

Proefpersonen werd gevraagd om de zin te beoordelen op semantische acceptabiliteit. Dit deden zij in twee taakcondities: één zonder extra taak en één waarin gevraagd werd de zinnen te beoordelen terwijl ze hardop van 1 tot en met 6 telden. Het hardop tellen was bedoeld om de fonologische loop uit te schakelen. Waters e.a. vonden dat propositionele complexiteit wel, maar syntactische complexiteit niet interageerde met het hardop tellen.

Het is opvallend aan deze zinnen dat bij sommige van de twee-actiezinnen de bijzin in het midden van de hoofdzin staat, terwijl hij bij de een-actiezinnen altijd aan het eind staat. Er wordt aangenomen dat centraal ingebedde zinnen lastiger te begrijpen zijn dan rechtsingebedde zinnen (Gunter, 1995; Miller & Isard, 1964). Nu kan het best zo zijn dat (een deel van) de effecten van propositionele complexiteit die door Waters e.a. (1987) zijn gevonden, toe te schijven zijn aan het gebruik van centraal ingebedde zinnen.

Hoofdstuk 3. Om te onderzoeken in hoeverre dat het geval was, is Experiment 1 van dit proefschrift uitgevoerd. Het had hetzelfde paradigma en dezelfde experimentele condities als het experiment van Waters e.a. (1987), behalve dat (a) alleen rechtsingebedde zinnen gebruikt werden en (b) de materialen in het Nederlands waren.

Experiment 1 leverde twee belangrijke resultaten op: in de eerste plaats deden de proefpersonen de taak sneller in de hardop-tellenconditie. Deze versnelling ging gepaard met een verslechtering van hun prestatie op semantisch implausibele zinnen.

De proefpersonen van Waters e.a. (1987) lieten juist een vertraging zien tijdens het hardop tellen. Gezien het feit dat in Nederlandse bijzinnen het werkwoord achteraan staat, duidt dit effect in Experiment 1 erop dat proefpersonen snel naar het einde van die zin willen komen omdat bij het werkwoord allerhande integraties plaatsvinden, zoals integratie van het subject en het werkwoord en θ-roltoekenning. Als tweede werd de interactie van propositionele complexiteit met hardop tellen van Waters e.a. (1987) niet gevonden. Dit wijst erop dat de interactie van Waters e.a. (1987) in feite een interactie van centrale inbedding met hardop tellen zou zijn.

Om deze laatste suggestie nader te onderzoeken is Experiment 2 uitgevoerd. Dit experiment vergeleek subject- en objectrelatieve zinnen en rechts- en centraal ingebedde zinnen in twee verschillende taakcondities (geen dubbeltaak tegenover hardop tellen). Deze taak leverde een aantal belangrijke effecten op.

Ten eerste presteerden proefpersonen net als in Experiment 1 sneller als ze hardop moesten tellen dan als ze geen extra taak uitvoerden. Deze versnelling ging wederom gepaard met een hoger aantal fouten voor implausibele zinnen tijdens het hardop tellen.

Ten tweede werd er geen hoofdeffect van centrale inbedding gevonden. Deze complexiteit vertoonde echter een interactie met het hardop tellen: zonder dubbeltaak vertoonden proefpersonen geen effect van centrale inbedding, maar tijdens het hardop tellen duurde het lezen van centraal ingebedde zinnen langer dan dat van rechts ingebedde zinnen.

In de derde plaats liet de foutenanalyse een tweeweginteractie zien van bijzinstype met het hardop tellen. Subject-relatieve zinnen leverdeen meer fouten op dan object-relatieve zinnen. Dit effect trad alleen op tijdens het hardop tellen en niet in de conditie zonder dubbeltaak.

De uitkomsten van Experiment 2 suggereren dat de interactie van propositionele complexiteit en hardop tellen van Waters e.a. (1987) geïnterpreteerd moet worden als een interactie van inbedding en hardop tellen.

Hoofdstuk 4. Experimenten 3 en 4 hadden hetzelfde paradigma als Experimenten 1 en 2, maar gebruikten niet-reversibele zinnen in plaats van reversibele zinnen. Zij lieten grofweg hetzelfde interactiepatroon zien als Experimenten 1 en 2, behalve dat Experiment 4 een speed-accuracy trade-off liet zien voor bijzinstype: langere reactietijden maar lagere foutenaantallen voor subjectrelatieve zinnen dan voor objectrelatieve zinnen. Dit verschil kan verklaard worden door het feit dat proefpersonen de agreementinformatie van onderwerp en werkwoord bewuster gebruiken in de niet-reversibele experimenten dan in de reversibele experimenten.

De vraag die rijst naar aanleiding van deze vier experimenten is welke rol de fonologische loop nu precies speelt bij het verwerken van zinnen. Interacties met inbedding suggereren dat de fonologische loop het subject van de hoofdzin actief houdt totdat het geïntegreerd kan worden met het hoofdwerkwoord van die zin.

Een alternatieve verklaring zou kunnen zijn dat de fonologische loop een akoestische representatie van de zin bewaart, zodat de lezer bepaalde delen van die representatie kan raadplegen tijdens de integratie van het subject en het werkwoord. De interactie met hardop tellen wijst er dan op dit proces moeilijker verloopt naarmate er een groter deel van die representatie geraadpleegd moet worden.

Deze tweede verklaring lijkt op basis van de data van Experimenten 1 tot en met 4 geschikter, omdat ze niet alleen de interactie met inbedding kan verklaren, maar ook de interactie van bijzinstype met hardop tellen in Experiment 2. Als de functie van de fonologische loop bestond uit het actief houden van zinsdelen die (voorlopig) nog niet geïntegreerd kunnen worden, zou die laatste interactie niet verwacht worden. Immers, subject- en objectrelatieve zinnen hebben dezelfde hoeveelheid niet-integreerbaar materiaal. Dit is te zien in (2a en b): in beide zinnen kunnen de woorden *die de uitgevers* niet worden geïntegreerd voordat het werkwoord bekend is.

(2) a. De studenten prijzen de fotograaf die de uitgevers accepteert
b. De studenten prijzen de fotograaf die de uitgevers accepteren

Maar als het controleren van congruentiekenmerken gebeurt met behulp van een fonologische geheugenrepresentatie in de fonologische loop, wat gebeurt er dan precies? Het kan zijn dat deze representatie wordt geraadpleegd bij alle congruentierelaties. Verder kan het zijn dat ze alleen wordt gebruikt bij de integratie van subject en werkwoord. In dat laatste geval zijn er twee mogelijkheden met betrekking tot de functie van dat proces: de representatie kan opgeroepen worden tijdens subject-werkwoordintegratie of tijdens het toekennen van θ-rollen. De θ-rollen geven aan wie wat doet in de zin.

Hoofdstuk 5. Deze verschillende ideeën over de rol van de fonologische loop bij zinsverwerking zijn verder onderzocht in Experiment 5. In dit experiment werd de hypothese dat de fonologische loop wordt gebruikt bij het controleren van alle mogelijke agreementrelaties getoetst door langeafstands- en locale integraties van adjectiva en zelfstandige naamwooden te vergelijken (in zinnen met 1 tegenover 3 adjectiva, zoals te zien is in [3a en b]).

(3) a. De barkeeper heeft, terwijl de verregende klant niets doorhad, de chips opgegeten
b. De barkeeper heeft, terwijl de verregende hongerige trouwe klant niets doorhad, de chips opgegeten

De vraag of de fonologische loop nu actief is tijdens θ-roltoekenning of tijdens de integratie van subject en werkwoord werd getoetst door te kijken of interacties met hardop tellen ook optreden in centraal ingebedde zinnen waar er een agreementrelatie bestaat tussen het onderwerp en een naburig hulpwerkwoord (tussen *barkeeper* en *heeft* in [3a en b]) en waar het onderwerp zijn θ-rol ontvangt door een lexicaal werkwoord aan het eind van de zin (*opgegeten* in [3a en b]).

Experiment 5 liet geen enkele interactie met hardop tellen zien. Dit wijst erop dat de fonologische loop wordt geraadpleegd om de agreement kenmerken van subject en werkwoord te controleren ten behoeve van de integratie van beide woordgroepen.

Hoofdstuk 6. Binnen de werkgeheugenliteratuur wordt veel gebruik gemaakt van werkgeheugentaken. Sommige onderzoekers nemen aan dat werkgeheugencapaciteit zoals die door dergelijke taken gemeten wordt een goede voorspeller is van welke lezers goed en welke minder goed lezen. De vrijwilligers die participeerden in Experiment 1 tot en met 5b werden aan twee geheugentaken onderworpen, aan een onzinwoordenherhaaltaak en aan een reading span taak (vgl. Daneman & Carpenter, 1980). Hun geheugenspans correleerden bijzonder slecht met de complexiteitsfactoren die werden getest in Experiment 1 tot en met 5b. Dit suggereert dat de gebruikte geheugentaken weinig tot geen voorspellende waarde hebben voor zinsverwerkingstaken.

Deze bevindingen zijn in strijd met voorspellingen van Just & Carpenter (1992), die aannemen dat geheugentaken, en met name de reading span taak, juist uitstekende voorspellers zijn van taalverwerking. Verder zijn ze in strijd met voorspellingen van Caplan & Waters (1999). Deze laatste onderzoekers stellen namelijk dat geheugentaken goede voorspellers van propositionele verwerkingscapaciteit zijn; dit kon niet in de huidige experimenten worden teruggevonden.

Hoofdstuk 7. In het laatste hoofdstuk wordt een model van het gebruik van werkgeheugensystemen gepresenteerd, waarin de zinsverwerking in twee stadia plaatsvindt. Gedurende de eerste fase vindt syntactische analyse plaats. Tijdens deze analyse kan de syntactische verwerker gebruik maken van een fonologische geheugenrepresentatie die actief wordt gehouden in de fonologische loop.

Na de syntactische analyse vindt de propositionele analyse plaats. Tijdens dit stadium worden verschillende proposities met elkaar gelinkt en vinden processen plaats zoals het interpreteren van pronomina.

Groningen Dissertations in Linguistics
(GRODIL)

1. Henriette de Swart (1991). *Adverbs of Quantification: A Generalized Quantifier Approach.*
2. Eric Hoekstra (1991). *Licensing Conditions on Phrase Structure.*
3. Dicky Gilbers (1992). *Phonological Networks. A Theory of Segment Representation.*
4. Helen de Hoop (1992). *Case Configuration and Noun Phrase Interpretation.*
5. Gosse Bouma (1993). *Nonmonotonicity and Categorial Unification Grammar.*
6. Peter I. Blok (1993). *The Interpretation of Focus.*
7. Roelien Bastiaanse (1993). *Studies in Aphasia.*
8. Bert Bos (1993). *Rapid User Interface Development with the Script Language Gist.*
9. Wim Kosmeijer (1993). *Barriers and Licensing.*
10. Jan-Wouter Zwart (1993). *Dutch Syntax: A Minimalist Approach.*
11. Mark Kas (1993). *Essays on Boolean Functions and Negative Polarity.*
12. Ton van der Wouden (1994). *Negative Contexts.*
13. Joop Houtman (1994). *Coordination and Constituency: A Study in Categorial Grammar.*
14. Petra Hendriks (1995). *Comparatives and Categorial Grammar.*
15. Maarten de Wind (1995). *Inversion in French.*
16. Jelly Julia de Jong (1996). *The Case of Bound Pronouns in Peripheral Romance.*
17. Sjoukje van der Wal (1996). *Negative Polarity Items and Negation: Tandem Acquisition.*
18. Anastasia Giannakidou (1997). *The Landscape of Polarity Items.*
19. Karen Lattewitz (1997). *Adjacency in Dutch and German.*
20. Edith Kaan (1997). *Processing Subject-Object Ambiguities in Dutch.*
21. Henny Klein (1997). *Adverbs of Degree in Dutch.*
22. Leonie Bosveld-de Smet (1998). *On Mass and Plural Quantification: The Case of French 'des'/'du'-NPs.*
23. Rita Landeweerd (1998). *Discourse semantics of perspective and temporal structure.*
24. Mettina Veenstra (1998). *Formalizing the Minimalist Program.*
25. Roel Jonkers (1998). *Comprehension and Production of Verbs in aphasic Speakers.*
26. Erik Tjong Kim Sang (1998). *Machine Learning of Phonotactics.*
27. Paulien Rijkhoek (1998). *On Degree Phrases and Result Clauses.*
28. Jan de Jong (1999). *Specific Language Impairment in Dutch: Inflectional Morphology and Argument Structure.*
29. H. Wee (1999). *Definite Focus.*
30. Eun-Hee Lee (2000). *Dynamic and Stative Information in Temporal Reasoning: Korean Tense and Aspect in Discourse.*
31. Ivilin Stoianov (2001). *Connectionist Lexical Processing.*
32. Klarien van der Linde (2001). *Sonority Substitutions.*
33. Monique Lamers (2001). *Sentence processing: Using Syntactic, Semantic, and Thematic Information.*
34. Shalom Zuckerman (2001). *The Acquisition of "Optimal" Movement.*
35. Rob Koeling (2002) *Dialogue-Based Disambiguation: Using Dialogue Status to Improve Speech Understanding.*
36. Esther Ruigendijk (2002). Case Assignment in Agrammatism: A Cross-Linguistic Study.
37. Anthony Mullen (2002). An Investigation into Compositional Features and Feature Merging for Maximum Entropy-Based Parse Selection.
38. Nanette Bienfait (2002). Grammatica-onderwijs aan allochtone jongeren.

39. Dirk-Bart den Ouden (2002). Phonology in Aphasia: Syllables and Segments in Level-Specific Deficits.

Grodil, secretary department of general linguistics
P.O.Box 716
9700 AS Groningen
The Netherlands